Constraining large-scale photosynthesis and transpiration fluxes in a biosphere model through data assimilation of in situ flux measurements including carbonyl sulfide

Camille Abadie

THÈSE DE DOCTORAT

NNT : 2024UPASJ027

université PARIS-SACLAY

Constraining large-scale photosynthesis and transpiration fluxes in a biosphere model through data assimilation of in situ flux measurements including carbonyl sulfide

Contraindre les flux de photosynthèse et de transpiration à grande échelle dans un modèle de biosphère par assimilation de mesures de flux in situ dont l'oxysulfure de carbone

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°129, Sciences de l'Environnement d'Ile de France (SEIF)
Spécialité de doctorat : Géosciences
Graduate School : Géosciences, climat, environnement et planètes
Référent : Université de Versailles-Saint-Quentin-en-Yvelines

Thèse préparée dans l'unité de recherche LSCE (Université Paris-Saclay, CNRS, CEA, UVSQ), sous la direction de **Philippe PEYLIN**, chercheur CNRS, et la co-direction de **Fabienne MAIGNAN**, chercheuse CEA.

Thèse soutenue à Paris-Saclay, le 5 novembre 2024, par

Camille ABADIE

Composition du Jury

Membres du jury avec voix délibérative

Erwan PERSONNE Professeur des Universités, AgroParisTech, Université Paris-Saclay - INRAE Île-de-France Versailles-Saclay	Président
Tristan QUAIFE Professeur - University of Reading	Rapporteur & Examinateur
Lisa WINGATE Directrice de Recherche - INRAE Bordeaux	Rapportrice & Examinatrice
Maarten KROL Professeur associé - Wageningen University, Utrecht University	Examinateur
Georg WOHLFAHRT Professeur associé - University of Innsbruck	Examinateur

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences de l'Environnement d'Île-de-France (SEIF)

Titre : Contraindre les flux de photosynthèse et de transpiration à grande échelle dans un modèle de biosphère par assimilation de mesures de flux *in situ* dont l'oxysulfure de carbone

Mots clés : Photosynthèse, Transpiration, Oxysulfure de carbone, Modélisation, Assimilation de données

Résumé : La production primaire brute (GPP), absorption photosynthétique du CO_2 atmosphérique par la végétation, joue un rôle crucial dans l'atténuation du changement climatique. La transpiration, émission d'eau par les plantes concomitante à la GPP, renvoie une part significative des précipitations terrestres dans l'atmosphère. Bien que la GPP soit le plus grand flux du cycle du carbone et que la transpiration soit la principale composante de l'évapotranspiration terrestre, leurs estimations globales et leurs réponses au changement climatique demeurent incertaines. Cette recherche vise donc à améliorer la simulation de la GPP et de la transpiration dans un modèle de surfaces continentales, ORCHIDEE. J'ai exploré l'utilisation de mesures d'oxysulfure de carbone (COS), un gaz atmosphérique absorbé par les plantes de manière similaire au CO_2, pour contraindre la GPP et la transpiration dans ORCHIDEE. J'ai tout d'abord implémenté un modèle des échanges de COS par les sols, complétant le modèle existant d'absorption de COS par la végétation. Ce développement a permis de simuler les flux de COS de l'écosystème à différentes échelles et de fournir de nouvelles estimations des contributions de la végétation et des sols au budget global du COS. Cette étude a montré l'importance de considérer la capacité des sols oxiques à émettre du COS en plus de leur absorption, et que les sols anoxiques peuvent produire des quantités significatives de COS (96 GgS/an), compensant en grande partie l'absorption nette de COS par les sols oxiques (-126 GgS/an), résultant en une absorption nette totale des sols de -30 GgS/an. Avec ces développements, des mesures *in situ* de flux de COS de l'écosystème ont pu être utilisées pour optimiser les paramètres d'ORCHIDEE via des techniques d'assimilation de données (DA). J'ai assimilé la plus longue série temporelle de flux de COS de l'écosystème de la forêt boréale de Hyytiälä en Finlande, afin d'optimiser les paramètres impliqués dans la simulation de la GPP et de la transpiration. La comparaison entre une assimilation conjointe ou indépendantes des données de COS et de GPP a montré que l'assimilation des données de COS en plus de celles de GPP améliore à la fois le flux de chaleur latente (LE) et la GPP simulés, ce que ne permet pas l'assimilation de la GPP seule. L'application des paramètres optimisés à l'ensemble de ce biome boréal a augmenté l'absorption de COS dans les hautes latitudes, en accord avec des études indépendantes. Ensuite, en collaboration avec d'autres chercheurs, j'ai travaillé sur les concentrations atmosphériques de COS et de CO_2 pour contraindre les flux de COS de surface, GPP, et respiration. Enfin, l'évaluation de la contrainte apportée par le COS sur la GPP et le LE simulés lors d'une sécheresse ayant révélé des erreurs structurelles non corrigées par la DA, la dernière partie de ma thèse s'est concentrée sur l'amélioration de la représentation de la réponse de la végétation à un stress hydrique. J'ai assimilé des données de GPP et de LE sur plus de 40 sites enregistrant les récentes années de sécheresse en Europe. Ce travail a montré que le paramètre déterminant la vitesse de fermeture stomatique pendant un stress hydrique pouvait être défini en fonction des conditions à long terme de déficit de pression de vapeur. L'implémentation de cette réponse dans ORCHIDEE permet de mieux prendre en compte la diversité de réponse de la végétation aux sécheresses et sa capacité d'acclimatation. Enfin, j'ai réalisé des projections pour évaluer l'impact de cette nouvelle réponse dans un climat futur. En conclusion, cette recherche encourage l'utilisation du COS comme proxy pour la GPP et la transpiration, et préconise davantage de mesures de flux et de concentration de COS pour affiner la paramétrisation des modèles. Elle suggère également des pistes pour améliorer la GPP et la transpiration simulées dans les modèles en tenant compte de la capacité des plantes à s'acclimater face au changement climatique.

université PARIS-SACLAY | ÉCOLE DOCTORALE Sciences de l'Environnement d'Île-de-France (SEIF)

Title: Constraining large-scale photosynthesis and transpiration fluxes in a biosphere model through data assimilation of *in situ* flux measurements including carbonyl sulfide

Keywords: Photosynthesis, Transpiration, Carbonyl sulfide, Modeling, Data assimilation

Abstract: Gross primary production (GPP), the photosynthetic absorption of atmospheric CO_2 by continental vegetation, plays a critical role in climate mitigation. In parallel, transpiration, the emission of water by plants, returns a significant portion of land precipitation to the atmosphere. Despite GPP being the largest carbon cycle flux and transpiration the largest component of terrestrial evapotranspiration, their global estimates and responses to climate change remain uncertain. Therefore, this research aims to enhance the simulation of GPP and plant transpiration in a land surface model (LSM), ORCHIDEE. First, I investigated the use of carbonyl sulfide (COS) measurements, an atmospheric gas absorbed by plants similarly to CO_2, to constrain GPP and plant transpiration in ORCHIDEE. This involved implementing a model for soil COS exchanges, complementing the existing vegetation COS uptake model. This development enabled the simulation of ecosystem COS fluxes from the site to the global scale and provided new estimates of vegetation and soil contributions to the global atmospheric COS budget. Notably, this study highlighted the importance of considering that oxic soils can emit COS in addition to absorbing COS, and that anoxic soils can produce significant amounts of COS (96 GgS y^{-1}), offsetting much of the net oxic soil COS uptake (-126 GgS y^{-1}), and resulting in a total global net soil uptake of -30 GgS y^{-1}. With these developments, *in situ* measurements of ecosystem COS fluxes have been used to optimize the parameters of ORCHIDEE through data assimilation (DA) techniques. Therefore, in a second phase, I assimilated the longest time series of ecosystem COS flux from the Hyytiälä boreal evergreen needleleaf forest in Finland to optimize parameters involved in GPP and plant transpiration simulation. Comparing a joint assimilation or independent assimilations of COS and GPP data showed that assimilating COS along with GPP data improves the simulated latent heat flux (LE) as well as GPP, unlike the GPP-only assimilation. Upscaling the optimized parameters across the boreal evergreen needleleaf forest biome increased COS uptake in high latitudes, aligning with independent studies. Then, in collaboration with other researchers, I worked on atmospheric concentrations of COS and CO_2 to constrain surface COS fluxes, GPP, and respiration. Finally, as the evaluation of COS constraint on the simulated GPP and LE during a drought event revealed structural errors in the simulated fluxes uncorrected by DA, the final part of my PhD focused on improving the representation of vegetation response to drought events in ORCHIDEE. To this end, I assimilated GPP and LE data at over 40 sites capturing recent drought years across Europe. This work demonstrated that the parameter determining stomatal closure speed during soil moisture stress, influencing both GPP and transpiration, could be defined as a function of long-term vapor pressure deficit conditions. Implementing this response in ORCHIDEE allows better consideration of site-specific vegetation responses to droughts and vegetation acclimation capacity. Finally, I performed projections to assess the impact of this new response under future climate. Overall, this research supports using COS as a proxy for GPP and transpiration, advocating for more COS flux and concentration measurement campaigns to refine LSM parameterization. It also suggests future avenues to improve GPP and plant transpiration representations in LSMs by accounting for plants' ability to acclimate in their response to climate change.

Publications

In this section, I present the publications I produced as a first author, followed by those where I am a co-author. The first publication listed under first authorship will soon be submitted to *Global Biogeochemical Cycles*, and the second publication listed under co-authorship is to be submitted to *Atmospheric Chemistry and Physics* in early September 2024.

Abadie, C., Maignan, F., Bacour, C., Peylin, P.: Vegetation acclimation to atmospheric moisture stress during drought events over Europe. (*to be submitted*)

Abadie, C., Maignan, F., Remaud, M., Kohonen, K.-M., Sun, W., Kooijmans, L., Vesala, T., Seibt, U., Raoult, N., Bastrikov, V., Belviso, S., Peylin, P.: 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 (2023)

Abadie, C., Maignan, F., Remaud, M., Ogée, J., Campbell, J. E., Whelan, M. E., Kitz, F., Spielmann, F. M., Wohlfahrt, G., Wehr, R., Sun, W., Raoult, N., Seibt, U., Hauglustaine, D., Lennartz, S., Belviso, S., Montagne, D., Peylin, P.: Global modelling of soil carbonyl sulfide exchanges. Biogeosciences, 19, 2427–2463. https://doi.org/10.5194/bg-19-2427-2022 (2022)

De Vries, A., Wohlfahrt, G., Kohonen, K.-M., **Abadie, C.**, Remaud, M., Kesselmeier, J., Korrensalo, A., Mammarella, I., Whelan, M., Vesala, T.: On the contribution of boreal wetlands to the Northern hemisphere carbonyl sulfide sink (*in prep.*)

Peylin, P., **Abadie, C.**, Haslehner, M., Belviso, S., Cadule, P., Remaud, M., Maignan, F.: Evaluation of CMIP/TRENDY model gross primary productivity and terrestrial ecosystem respiration using atmospheric COS and CO_2 data. (*to be submitted*)

Berchet, A., Pison, I., Huselstein, C., Narbaud, C., Remaud, M., Belviso, S., **Abadie, C.**, Maignan, F.: Can we gain knowledge on COS anthropogenic and biogenic emissions from a single atmospheric mixing ratios measurement site? Atmospheric Chemistry and Physics [preprint]. https://doi.org/10.5194/egusphere-2024-549 (*In review*)

Remaud, M., Ma, J., Krol, M., **Abadie, C.**, Cartwright, M. P., Patra, P., Niwa, Y., Rodenbeck, C., Belviso, S., Kooijmans, L., Lennartz, S., Maignan, F., Chevallier, F., Chipperfield, M. P., Richard, J. P., Harrison, J. J., Vimont, I., Wilson, C., Peylin, P.: 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, e2022JD037817. https://doi.org/10.1029/2022JD037817 (2023)

Belviso, S., **Abadie, C.**, Montagne, D., Hadjar, D., Tropée, D., Vialettes, L., Kazan, V., Delmotte, M., Maignan, F., Remaud, M., Ramonet, M., Lopez, M., Yver-Kwok, C., Ciais, P.: Carbonyl sulfide (COS)

emissions in two agroecosystems in central France. PLOS ONE, 17(12): e0278584. https://doi.org/10.1371/journal.pone.0278584 (2022)

Belviso, S., Remaud, M., **Abadie, C.**, Maignan, F., Ramonet, M., Peylin, P.: 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 (2022)

Remaud, M., Chevallier, F., Maignan, F., Belviso, S., Berchet, A., Parouffe, A., **Abadie, C.**, Bacour, C., Lennartz, S., Peylin, P.: Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling. Atmospheric Chemistry and Physics, 22, 2525–2552. https://doi.org/10.5194/acp-22-2525-2022 (2022)

Maignan, F., **Abadie, C.**, Remaud, M., Kooijmans, L., Kohonen, K.-M., Commane, R., Wehr, R., Campbell, J. E., Belviso, S., Montzka, S. A., Raoult, N., Seibt, U., Shiga, Y. P., Vuichard, N., Whelan, M. E., Peylin, P.: Carbonyl sulfide: comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach. Biogeosciences, 18, 2917–2955. https://doi.org/10.5194/bg-18-2917-2021 (2021)

Acknowledgments

My deepest thanks go to my two PhD supervisors, Fabienne Maignan and Philippe Peylin, for their support, guidance, and continuous encouragement in promoting my work. I am especially thankful for their confidence in me, which allowed me to grow and improve in areas where I felt least comfortable. I am extremely grateful for the opportunity to work with them over these past years, benefiting from their high and diverse levels of expertise, and for the rich and healthy work environment they have always provided.

I would like to express my gratitude to Lisa Wingate, Georg Wohlfahrt, Tristan Quaife, Maarten Krol, and Erwan Personne, who have accepted to take part in the jury commission of this PhD thesis.

Special recognition is also due to the COS team at LSCE, particularly Marine Remaud, for the fruitful discussions and whose collaboration has been a pleasure. I am also thankful to Sauveur Belviso for his constant motivation to integrate experimental and modeling efforts, and for inviting me on the experimental side to gather a few COS concentration flask samples. I would also like to extend my thanks to the members of the MOSAIC team of which I was a part, and to Cédric Bacour for his numerous pieces of advice and significant help in improving my work.

I am grateful to all the past and present postdocs, PhD students, and interns at LSCE whom I have had the pleasure of knowing, and who collectively contribute to making LSCE such a nice place to work. Special thanks to Julien, Lucas, Aya, Simon, Vincent, Mandresy, Luis, Mojtaba, Karine, Amélie, Zacharie, Maureen, Valentin, Benoit, Guillermo, Maud, Coralie, and many others. Julien and Lucas deserve a particular mention for their constant support, (unexpected and well hidden) wisdom, (sometimes questionable) humor, but undoubtedly valuable friendships. I am also grateful for Aya's positivity, kindness, and shared laughter, which have uplifted me daily. Also a special thanks to the motivated people of the climbing group, Aya, Vincent, Simon, Mojtaba, Karine, Amélie, Guillermo, and Maud, with whom suffering together has turned out to be so enjoyable.

I am also thankful for the wonderful people in my life, Alexandra, Viviane, Antoine, Mathias, Thomas, Hugo, Léa, Charlotte, and many others.

Finally, I would like to convey my appreciation to my family, especially my mother and brother, for their constant support and belief in me through the years.

Contents

Abbreviations

ABA	Abscisic Acid
ALT	Alert station, Canada
APAR	Absorbed Photosynthetically Active Radiation
ATM	Atmospheric Transport Model
ATP	Adenosine triphosphate
BEPS	Boreal Ecosystem Productivity Simulator
C_a	Atmospheric concentration
C_c	Chloroplast concentration
C_i	Internal concentration
C_s	Leaf surface concentration
CA	Carbonic Anhydrase
CIF	Community Inversion Framework
CMIP	Coupled Model Intercomparison Projects
CO_2	Carbon dioxide
COS/OCS	Carbonyl sulfide
CS_2	Carbon disulfide
DA	Data Assimilation
DMS	Dimethyl sulfide
ENVISAT	European Space Agency Environmental Satellite
ESA	European Space Agency
ESM	Earth System Model
EVI	Enhanced Vegetation Index
FACE	Free-Air CO_2 Enrichment
FAO	Food and Agriculture Organization
FLEX	Fluorescence Explorer
FLEXPART	Flexible Particle model
g_b	Boundary layer conductance
g_c	Rubisco reaction rate equivalent conductance for CO_2
g_i	Internal conductance for COS
g_m	Mesophyll conductance for CO_2
g_s	Stomatal conductance
GA	Genetic Algorithm
GLEAM	Global Land Evaporation Amsterdam Model
GPP	Gross Primary Production
H_2O	Water
HIPPO	HIAPER Pole-to-Pole Observations
ICOS	Integrated Carbon Observation System
ITRDB	International Tree-ring Data Bank
JULES	Joint United Kingdom Land Environment Simulator
LAI	Leaf Area Index
L-BFGS-B	Limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm with Bound constraints
LE	Latent heat flux
LMDZ	Laboratoire de Météorologie Dynamique atmospheric circulation model
LRU	Leaf Relative Uptake
LSM	Land Surface Model

LST	Land Surface Temperature
LUE	Light Use Efficiency
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLO	Mauna Loa station, United States
MODIS	Moderate Resolution Imaging Spectroradiometer
$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Hydrogen
NASA	National Aeronautics and Space Administration
NBP	Net Biome Productivity
NDVI	Normalized Difference Vegetation Index
NEE	Net Ecosystem Exchange
NOAA	National Oceanic and Atmospheric Administration
NPQ	Non-Photochemical Quenching
NUCAS	Nanjing University Carbon Assimilation System
O_2	Oxygen
ORCHIDAS	Organising Carbon and Hydrology In Dynamic Ecosystems Data Assimilation System
ORCHIDEE	Organising Carbon and Hydrology In Dynamic Ecosystems
PAR	Photosynthetically Active Radiation
PFT	Plant Functional Type
PRI	Photochemical Reflectance Index
^{222}Rn	^{222}Radon
R_{aero}	Aerodynamic resistance
$R_{boundary}$	Boundary layer resistance
$R_{internall}$	Internal resistance
$R_{mesophyll}$	Mesophyll resistance
$R_{stomata}$	Stomatal resistance
RMSD	Root Mean Squared Deviation
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	Ribulose-1,5-bisphosphate
SA	Sensitivity Analysis
SiB	Simple Biosphere model
SIF	Solar-Induced Fluorescence
SSP	Shared Socioeconomic Pathways
T_{air}	Air temperature
TransCom	Atmospheric tracer transport model intercomparison project for COS
TRENDY	Trends in Net Land-Atmosphere Carbon Exchange
TROPOMI	Tropospheric Monitoring Instrument
USDA	United Nations-United States Department of Agriculture
VIs	Vegetation Indices
VOD	Vegetation Optical Depth
VPD	Vapor Pressure Deficit
WUE	Water Use Efficiency

1 INTRODUCTION

1.1 Context and overview: Challenges of modeling vegetation responses to climate change

The ability of vegetation to face climate change is a fundamental issue for societies that rely on the services provided by ecosystems, which represent all the benefits derived from natural ecosystems, essential for human physical, social, and economic well-being (Costanza et al. 1997). These ecosystem services encompass provisioning services such as the supply of food and wood fuel, regulating services like water purification and carbon storage, and cultural services (Daily 1997; MA 2005). All of these depend on underlying supporting services, including nutrient cycling and water cycling. Terrestrial vegetation acts as a major carbon sink, playing a part in climate mitigation by absorbing about 30% of the atmospheric CO_2 released by anthropogenic activities, while the oceans take up about 25%, and the remainder stays in the atmosphere (Friedlingstein et al., 2022). In 2021, 3.5 ± 0.9 GtC y^{-1} were absorbed by the terrestrial land sink, and 2.9 ± 0.4 GtC y^{-1} by the ocean sink, out of a total anthropogenic emission of 10.9 ± 0.8 GtC y^{-1} (Friedlingstein et al., 2022). In parallel to CO_2 uptake through photosynthesis, plants emit water vapor to the atmosphere during transpiration, sending back 40% of the total precipitation over lands (Oki & Kanae, 2006).

However, the increasing atmospheric CO_2 concentration significantly alters ecosystem functioning by enhancing plant carbon assimilation mainly for plants with C_3 photosynthetic metabolism, thereby promoting plant growth, a phenomenon known as the CO_2 fertilization effect (Chen et al., 2022). Indeed, under current atmospheric CO_2 levels, plant growth is limited by CO_2 availability in the absence of other limiting factors such as soil nutrients or water availability (Boretti et al., 2019). This rise in atmospheric CO_2 concentration, as well as shifts in precipitation and temperature patterns, also influences the phenology and length of the growing season (Keenan et al., 2014), and modifies vegetation carbon uptake interannual variability (Zhang et al., 2016). Climate change also exacerbates the frequency and intensity of extreme events such as droughts and heatwaves, inducing further stress on vegetation (Crausbay et al., 2017). Due to the strong interactions between vegetation and the atmosphere through mass and energy exchanges, climate change stresses on ecosystems also influences the future climate (Bonan & Doney, 2018). For example, increases in vegetation leaf area index (LAI) in regions where there is winter snow can cause a positive feedback and warming through reductions in the winter snow-cover albedo (Bonan et al., 1992). On the other hand, an increase in LAI enhances the cooling effect of evapotranspiration, leading to negative feedback (Woodward et al., 1998). Therefore, vegetation responses and feedback to present and future climate is still highly uncertain.

Because both photosynthesis and plant transpiration cannot be directly measured at the ecosystem scale (Damm et al., 2010), various approaches have been developed to quantify these fluxes from the ecosystem to the global scale. Data-driven methods, leveraging increasingly available Earth observations and machine learning techniques, have been widely adopted to

provide global scale estimates. Additionally, land surface models (LSMs) have been developed to quantify photosynthesis and plant transpiration across different spatial and temporal scales, facilitating the study of future vegetation-climate interactions and feedback through projections. The accuracy of LSM simulations and projections, however, strongly depends on our ability to accurately represent exchange processes between the land surface and the atmosphere. Consequently, LSMs are continually being refined to incorporate more complexity, enhancing the realism of vegetation responses to climate change. A major challenge, therefore, lies in improving the modeling of photosynthesis and plant transpiration to accurately represent vegetation responses to climate change and extreme events.

In this context, different proxies have emerged, corresponding to a measurement of one physical quantity used as a substitute for another quantity that is too challenging or costly to measure directly (Bailey et al., 2018). These proxies can be used to inform the representation of photosynthesis and plant transpiration in LSMs.

1.2 GPP and plant transpiration: mechanisms, interactions, and climate sensitivity

1.2.1 GPP and plant transpiration: uncertainties in two key processes of plant functioning

Vegetation carbon uptake through photosynthesis, called gross primary production (GPP), is the largest and the most uncertain flux in the carbon cycle (Anav et al., 2015). There is no consensus on the global total GPP, with estimates ranging from 112 to 169 PgC y^{-1} (Anav et al., 2015). While observation-based and LSM GPP estimates typically align on the phase of the GPP seasonal cycle, there are significant discrepancies regarding the seasonal amplitude (Anav et al., 2015). Additionally, GPP interannual variability and trends differ between observation-based databases and LSM estimates from 1990 to 2009, with LSMs showing stronger increasing trends, ranging from 0.28 PgC y^{-1} to 0.62 PgC y^{-1} (Anav et al., 2015).
Similarly, plant transpiration is the largest component of total terrestrial evapotranspiration (Jasechko et al., 2013; Wei et al., 2017). Using a combination of remote sensing and modeling approaches, Vicente-Serrano et al. (2022) estimated global plant transpiration to be 50.5 10^3 km^3 y^{-1} between 1980 and 2020, accounting for 74.7% of total terrestrial evapotranspiration, with a positive trend of 0.28 km^3 $decade^{-1}$ over this period. However, transpiration estimates vary significantly, ranging from 24% to 90% of total land evapotranspiration, depending on the estimation method, including satellite-based estimations, reanalysis, LSMs, and isotopic measurements (Wei et al., 2017).
Therefore, both photosynthesis and plant transpiration estimates still have large uncertainties. This highlights the need to better constrain these two key processes of plant functioning, which are interconnected through the diffusion pathway of carbon and water vapor via stomata and are strongly influenced by climate drivers.

Quantifying GPP is essential because it represents the initial influx of material and energy into a terrestrial ecosystem, directly indicating its productivity. Plants absorb light energy to produce organic molecules such as glucose from CO_2 and H_2O, while releasing molecular O_2 into the atmosphere. These organic molecules can then be used for plant metabolism and growth (Alberts et al., 2002). Photosynthesis consists of two interconnected parts, called light and dark reactions (Stirbet et al., 2020). The light reaction takes place in the thylakoid membranes of the chloroplasts, where the light energy absorbed by the photosystems containing chlorophyll pigments is mainly converted into redox chemical energy. In this phase, H_2O is oxidized to O_2, adenosine triphosphate (ATP) is produced, which is a molecule that stores energy used for cellular processes, and nicotinamide adenine dinucleotide phosphate ($NADP^+$) is reduced to NADPH, an electron donor that provides energy for cellular reactions. Then, both NADPH and ATP are required for CO_2 assimilation during the dark part of photosynthesis in the chloroplast stroma. This part includes the Calvin cycle reactions with the Ribulose-1,5-bisphosphate carboxylase/oxygénase (Rubisco) enzyme that initiates CO_2 fixation onto Ribulose 1,5-bisphosphate (RuBP), which will further result in the production of C_3 or C_4 sugars that the plant uses for metabolism.

This carbon gain through photosynthesis occurs alongside water loss through transpiration. Due to its large contribution, transpiration determines global water availability for human and ecological uses. The spatial variability in transpiration is influenced by multiple factors, such as precipitation, soil water availability, or vegetation density and physiological mechanisms (Sitch et al., 2003). By emitting water vapor to the atmosphere, plant transpiration represents one of the main components of the water cycle and contributes to moisture recycling (Oki & Kanae, 2006). Plant transpiration is also a key process in the soil-plant-atmosphere continuum. Indeed, it drives the water transport from the roots to the leaves, allowing the distribution of the nutrients in the plant cells. In addition, by converting liquid water to water vapor in the stomata, plant transpiration has a cooling effect as the energy used for this conversion would otherwise be used to heat the leaves (Grossiord et al., 2020).

1.2.2 The coupling between GPP and plant transpiration through stomatal conductance

The stomata play a crucial role in vegetation and atmosphere interplays, allowing the exchanges of carbon and water vapor between these two components. Because carbon gain and water loss are coupled through the regulation of stomatal opening, Briggs & Shantz (1913) introduced the notion of water use efficiency (WUE) as a measure of the amount of carbon assimilated as biomass per unit of water used by a plant. Then, stomatal regulation balances the risk of carbon starvation and the risk of hydraulic failure. Indeed, while maintaining the stomata open benefits CO_2 uptake for photosynthesis, an excessive water loss through plant transpiration can lead to soil water depletion, a decrease in plant water potential, xylem cavitation, and a risk of mortality (Drake et al., 2017; Martin-StPaul et al., 2017; Dewar et al., 2018).

Considering this tradeoff, optimality theories suggested that plants respond to environmental changes by regulating stomatal opening to maximize the difference between carbon gain and water loss (Prentice et al., 2014; Wang et al., 2020). The concepts of isohydricity and anisohydricity have also emerged to describe the different plant water use strategies (Hochberg et al., 2018). These concepts link stomatal regulation to soil and leaf water potentials (Konings et al., 2017; Martínez-Vilalta et al., 2014; Meinzer et al., 2016). When soil water potential drops due to water depletion, while isohydric species close their stomata to maintain the leaf water potential with a risk of carbon starvation, anisohydric species maintain their stomata open with a risk of hydraulic failure (Martínez-Vilalta et al., 2014).

In addition to soil moisture status affecting stomatal conductance, vapor pressure deficit (VPD), defined as the difference between saturation and actual air vapor pressure, also regulates stomatal opening (Grossiord et al., 2020). A rapid increase in VPD typically causes plants to close their stomata to minimize water loss. However, stomatal sensitivity to VPD varies significantly between and within species (Creese et al., 2014; Gao et al., 2015). The exact mechanisms driving stomatal closure in response to increased VPD are not fully understood but likely involve active sensing of water status within leaf cells, mediated by hormonal signals such as abscisic acid (ABA). Stomatal responses are influenced by leaf water potential and hydraulic conductance, linking them to the plant's overall hydraulic characteristics (Sharpe et al., 1987). Therefore, variations in stomatal conductance in response to VPD determine plant transpiration rates. When VPD is low and stomata are fully open, plant transpiration increases with VPD. However, above a certain VPD threshold, stomatal closure reduces plant transpiration to limit water loss (Cunningham, 2004; Whitley et al., 2013). Finally, short-term stomatal responses to high VPD can differ from long-term impacts, as stomata can acclimate to prolonged high VPD exposure, enhancing stomatal conductance and plant transpiration (Marchin et al., 2016).

1.2.3 GPP and plant transpiration dynamics under climate change

Ecosystems are changing in response to rising CO_2 concentrations driving climate change, which influences global shifts such as temperature increases, and altered precipitation patterns, while also exacerbating the frequency and intensity of extreme events (Malhi et al., 2020). However, the sensitivity and responses of ecosystems to climate change vary across different regions and are influenced by species' abilities to adapt, resist, and recover, as well as by additional disturbances, including those caused by humans, and interactions among organisms.
Several studies have projected an increase in GPP in the Northern high latitudes, where temperature typically limits growth. Global warming is expected to enhance photosynthesis in these regions (Wu et al., 2015) and lead to earlier growing season start and later senescence for temperate and boreal vegetation (Keenan et al., 2014; Jeong et al., 2011; Delbart et al., 2008). Conversely, GPP is anticipated to decrease in tropical regions, related to a contraction of tropical humid regions, and expansion of areas with intense dry periods (Uribe et al., 2023). In arid and semi-arid regions, where precipitation is often a limiting factor for vegetation growth (Snyder & Tartowski, 2006), global warming may exacerbate droughts by increasing evapotranspiration (Miao et al., 2020). Additionally, in the Mediterranean, Southern North America, Southern Africa,

Australia, and Amazonia, earth system model (ESM) intercomparison exercises project reduced plant transpiration due to decreased precipitation, resulting in lower soil moisture availability (Douville et al., 2021).

The increase in atmospheric CO_2 concentration is expected to enhance photosynthesis driven by the higher availability of CO_2 for fixation by Rubisco (Drake et al., 1997). However, the degree of these responses and spatial patterns are unclear due to the influence of other factors on photosynthesis, such as water and nutrient availability (Leakey et al., 2009; Zhu et al., 2017).
Regarding plant transpiration, increased CO_2 typically reduces stomatal opening, thereby decreasing transpiration. However, this effect is countered by the increase in transpiration due to CO_2 fertilization and enhanced foliage cover (Cowling & Field, 2003).

The increase in CO_2 concentration also induces an increase in global temperature, another extensively studied driver of photosynthesis. Photosynthesis critically depends on air temperature, however, photosynthesis temperature dependence varies between or within species, also depending on different growing conditions (Yamori et al., 2014; Kumarathunge et al., 2019). Indeed, it has been shown that the optimal temperature that maximizes the photosynthetic rate can increase with growth temperature, which is the temperature of the environment where plants grow, due to acclimation mechanisms (Berry & Björkman, 1980). However, global temperature increase is expected to have distinct effects on biomes, being mainly beneficial for plants in cooler habitats, while more stressful in warmer habitats.

The global rise in temperature leads to an increase in VPD (Hatfield & Prueger, 2015), which is a key driver of plant functioning. Indeed, high VPD corresponds to a strong atmospheric water demand that has been found to increase plant transpiration up to a threshold (Franks et al., 1997). This enhancement of plant transpiration can result in soil moisture depletion and increase the risk of plant water stress (Dai, 2013; Grossiord et al., 2020). To reduce the risk of hydraulic failure, plants tend to close their stomata under high VPD conditions (Running, 1976). While further research is needed on the impact of high VPD on photosynthetic capacity, Flexas et al. (2006) found that Rubisco activity can be downregulated with lower carboxylation and electron transport rates following stomatal closure induced by high VPD. Under soil moisture stress conditions, plant carbon assimilation has also been found to decrease due to non-stomatal limitations with a reduction in mesophyll conductance (Flexas et al., 2012).

The complexity of GPP and plant transpiration responses to climate change associated to different timescales makes it challenging to accurately project vegetation functioning under future climate. Moreover, it is difficult to disentangle the effect of each driver on GPP and plant transpiration with combined effects of co-existing global changes such as a rise in atmospheric CO_2, air temperature, and VPD (Novick et al., 2016), but also co-occurring extreme events with heatwaves and droughts that often happen alongside. Therefore, it is essential to test hypotheses using LSMs and validate these models against experimental data to better understand the interactions between these complex factors and improve the reliability of climate projections.

1.3 Challenges in estimating GPP and plant transpiration

1.3.1 Approaches to estimating GPP and LE from the ecosystem to the global scale

There is no direct measurement of GPP above the leaf scale because net CO_2 exchange entails both photosynthesis and respiration, which must be partitioned to estimate GPP (Damm et al., 2010). Similarly, for plant transpiration, at the ecosystem scale the water flux involves other components, such as soil evaporation, snow sublimation, evaporation from canopy and interception, or floodplain evaporation, which collectively contribute to evapotranspiration or latent heat flux (LE). Consequently, at global scales, our ability to estimate GPP and plant transpiration rely on data-driven approaches or simulated fluxes in LSMs.

At the ecosystem scale, since the 1990s, eddy-covariance tower networks like FLUXNET (https://fluxnet.org/), including Ameriflux (https://ameriflux.lbl.gov/) and the Integrated Carbon Observation System (ICOS) (https://www.icos-cp.eu/), have been developed to provide *in situ* LE and GPP estimates at hundreds of sites across the globe. The eddy-covariance technique provides continuous measurements of GPP or LE by analyzing the turbulent exchange of CO_2 or water vapor between the ecosystem and the atmosphere, using high-frequency data from sensors placed on a tower to capture the vertical wind speed and gas concentrations (Burbat, 2013). GPP estimates are derived by partitioning the net ecosystem exchange (NEE) into its GPP and ecosystem respiration components based on day-time or night-time separation algorithms (Lasslop et al., 2010; Reichstein et al., 2005). However, regions such as North America and Europe are disproportionately represented in terms of site numbers compared to other regions like the tropics, which have a scarcity of sites. Moreover, the eddy-covariance tower footprint, defined as the temporally dynamic areas that contribute to measured fluxes, typically ranges from a few hundred meters to a few kilometers (Chu et al., 2021). Consequently, *in situ* eddy-covariance measurements are insufficient to represent the spatiotemporal patterns of global GPP and LE and the diversity of terrestrial ecosystems.
On the model of the FLUXNET network, Poyatos et al. (2021) recently released SAPFLUXNET, the first global database of sap flow measurements (https://sapfluxnet.creaf.cat/). Sap flow measurements can provide information on plant transpiration dynamics and its response to environmental changes. The SAPFLUXNET database encompasses 202 datasets from 2714 plants of 174 species, predominantly trees. Covering diverse biomes, particularly woodlands, shrublands and temperate forests, the dataset span from 1995 to 2018 and include sub-daily sap flow and hydrometeorological data, along with metadata on stand characteristics and plant attributes. However, similar to FLUXNET, SAPFLUXNET suffers from the sparsity of sites, with Europe, the Eastern USA, and Australia particularly well represented.

Data-driven GPP and LE products, such as FLUXCOM (https://www.fluxcom.org/) (Jung et al. 2019; Tramontana et al. 2016; Nelson et al., 2024), can be derived from machine learning methods that upscale *in situ* eddy-covariance measurements using remote sensing and meteorological data. However, this approach is constrained by the limited number of *in situ* sites and the uncertainties

involved in upscaling these local measurements to a global scale. Consequently, while these machine learning-based products are extensively used for benchmarking, significant biases in GPP estimation tend to occur in regions with sparse flux sites (Jung et al., 2020).

Other data-driven approaches do not rely on an upscaling of *in situ* eddy-covariance measurements, but uses models and remote sensing data to provide large scale GPP and LE estimates. This includes for example GPP products that are based on light use efficiency (LUE) models, such as FLUXSAT (Joiner et al. 2018) (https://daac.ornl.gov/VEGETATION/guides/FluxSat_GPP_FPAR.html), and LE estimates from models like the Global Land Evaporation Amsterdam Model (GLEAM) (Miralles et al., 2011) (https://www.gleam.eu/), which computes the different components of evapotranspiration from satellite data. In the LUE approach, GPP is a function of vegetation absorbed photosynthetically active radiation (APAR) and the efficiency of light absorption for carbon fixation. Note that LUE products can still rely on *in situ* eddy covariance measurements for calibration. While LUE products effectively capture the spatial distribution patterns of GPP, they tend to overestimate GPP under dry and cold conditions (Ryu et al., 2011; Wei et al., 2017).

All these data driven methods present limitations related to their underlying assumptions, parametrization, and uncertainties in the *in situ* eddy-covariance and remote sensing data. Consequently, various data-driven products yield different GPP and LE estimates, as illustrated in Figure 1 for GPP, highlighting the need for further constraints on these fluxes to reduce the associated uncertainties.

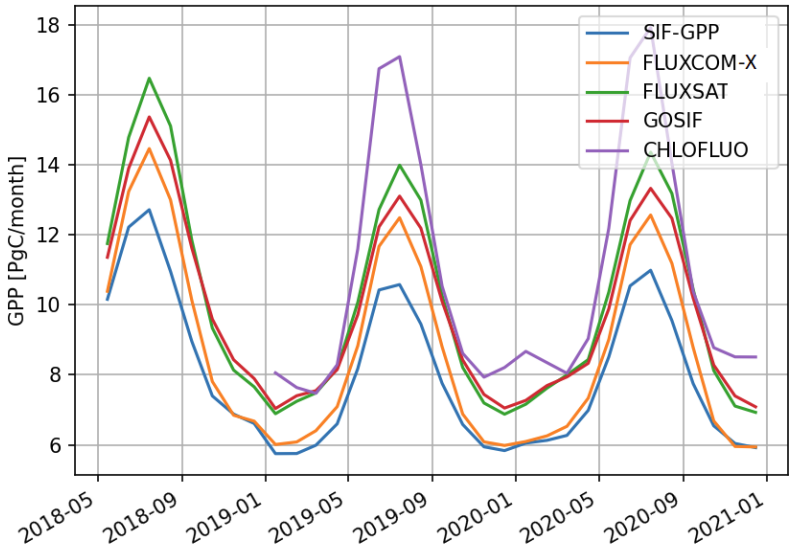

Figure 1: Seasonal cycles of total GPP estimates (PgC month^{-1}) from 5 data-driven global products from May 2018 to December 2020. The 5 data-driven products are SIF-GPP (Duveiller et al., 2023), FLUXCOM-X (Nelson et al., 2024), FLUXSAT (Joiner et al., 2018), GOSIF (Li and Xiao, 2019), and CHLOFLUO (Doughty et al., 2024).

Another approach to estimate GPP and LE or plant transpiration at large scales relies on their simulation in LSMs such as the Organising Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) model. This approach is not independent from the *in situ* eddy covariance measurements and data-driven products described above, as these data are used for LSM

parameter calibration and simulated flux evaluation. In addition to providing global GPP and LE estimates, projections by LSMs also enable to investigate future flux spatial and temporal dynamics. A diverse range of LSMs have been developed, increasing in complexity to improve process representation and accuracy. However, estimated GPP and LE vary significantly between LSMs, as illustrated in Figure 2, which shows the simulated GPP by 18 LSMs from the Trends in Net Land-Atmosphere Carbon Exchange (TRENDY) project. This variability in simulated fluxes arises from differences in parametrization, including process representation and parameter values that depend on calibration accuracy (Anav et al., 2015). Additionally, uncertainties in simulated fluxes also come from LSM input data, such as climate forcing data.

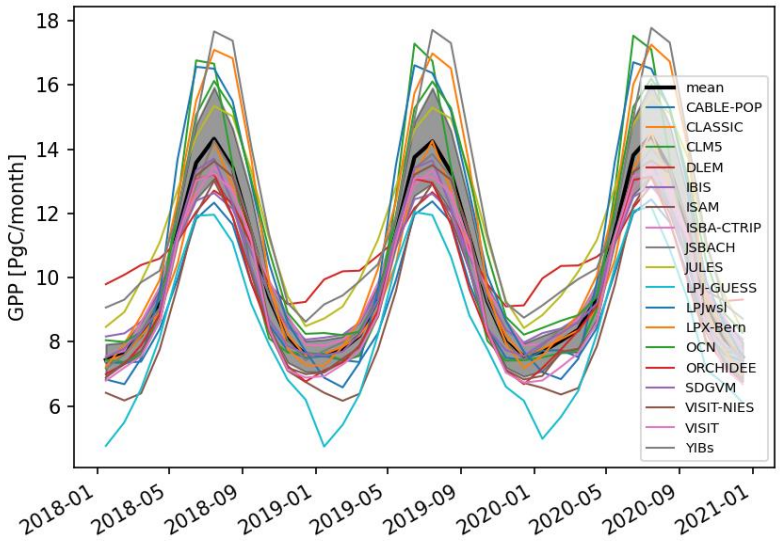

Figure 2: Seasonal cycles of total GPP estimates (PgC month^{-1}) from 18 LSMs of the TRENDY project between 2018 and 2020. The model ensemble mean is in black while each model is represented by a color.

In addition to the above methods, proxies for GPP and plant transpiration can be used as substitute measurements and provide estimates of these two fluxes that cannot be directly measured above the leaf scale. In particular, the emergence of new proxies for GPP and plant transpiration can offer additional constraints to inform processes representation and parameter calibrations in LSMs, thereby reducing uncertainty in the simulation and projection of GPP and LE.

1.3.2 Using proxies for large scale GPP and LE estimates

Remote sensing vegetation indices (VIs) have traditionally served as proxies for GPP or as input data for LUE models (Wang et al. 2004; Zhou et al. 2001; Running et al. 2004). VIs such as the normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and LAI reflect vegetation states related to greenness or leaf area, which are associated with photosynthetic capacity (Huete, 1997; Tucker & Sellers, 1986). However, these traditional VIs have limitations when estimating GPP. For instance, they tend to saturate over dense canopies (Turner et al., 2003; Running et al., 2004) and are influenced by the reflectance of soil, snow, or water bodies, as well

as atmospheric scattering (Huete et al., 1985, 2002). Additionally, VIs may not accurately capture rapid changes in GPP due to their temporal resolution or during periods of vegetation stress where greenness does not change. The spatial heterogeneity of vegetation types and conditions can also lead to variability in the relationship between VIs and GPP (D'Odorico et al., 2015; Zhou et al., 2014). VIs like NDVI (Goward & Huemmrich, 1992) and EVI (Xiao et al., 2004) have been used as proxies for LE by providing insights into surface conductance, which is closely tied to plant transpiration. Remote sensing of land surface temperature (LST) (Kalma et al., 2008), or a combination of VIs and LST data, has also been used to estimate global LE (Tang et al., 2010). VIs are generally used to estimate LE through empirical relationships based on flux tower LE measurements (Nishida et al., 2003) or to parameterize the conductance term in the Penman-Monteith equation to estimate evapotranspiration (Leuning et al., 2008). LST data can serve as input for a surface energy balance model (Kalma et al., 2008). However, Cleugh et al. (2007) found that inaccuracies in LSTs led to significant errors in LE estimates when using a surface energy balance model. Additionally, LE models that use VIs to represent transpiration as a function of surface conductance have been found to overestimate water fluxes in water-limited ecosystems, such as savannahs and arid and semi-arid regions (Barraza et al., 2017).

In recent years, our understanding of global GPP dynamics has improved through the use of newer satellite products more closely linked to the photosynthetic process. In particular, the near-infrared reflectance of vegetation (NIRv) and sun-induced fluorescence (SIF) have been investigated to provide information on vegetation physiology (Porcar-Castell et al., 2014; Wang et al., 2019). Compared with NDVI, NIRv and SIF are less affected by the soil background and atmospheric scattering (Badgley et al., 2017). NIRv, which is the product of NDVI and near-infrared reflectance, is demonstrated to be a good proxy of GPP at monthly to annual scales (Badgley et al., 2019; Wang et al., 2021). SIF is an electromagnetic signal emitted by chlorophyll *a* molecules when they are illuminated by sunlight, directly linking it to the light reactions of photosynthesis (Frankenberg & Berry, 2018). Several studies have demonstrated that satellite-based SIF retrievals exhibit a linear relationship with GPP at the ecosystem scale across various biomes (Guanter et al., 2014; Sun et al., 2017). Because SIF tracks plant activity and CO_2 and H_2O exchanges are closely linked through stomatal diffusion, SIF has also been used as a proxy for plant transpiration or LE (Alemohammad et al., 2017; Maes et al., 2021; Zhang et al., 2024). However, SIF represents only about 3% of the absorbed energy, making it a weak signal and leading to significant retrieval noise in passive measurements (Joiner et al., 2020). Additionally, the relationship between GPP and SIF is influenced by various factors such as canopy structure (van der Tol et al., 2019; Yang et al., 2019), cloud cover, sun-sensor geometry effects (Zhang et al., 2021), or environmental stressors (Wohlfahrt et al., 2018). Similarly, the empirical relationship between transpiration and SIF was found to be influenced by photosynthetically active radiation (PAR), VPD, and air temperature (Lu et al., 2018). Tracking GPP and LE variations at fine temporal and spatial scales is also constrained by the current coarse resolution of satellite measurements. However, the European Space Agency's (ESA) Fluorescence Explorer (FLEX), set to be launched in 2026, will be the first satellite mission specifically designed to measure SIF, offering data at a 300 m spatial resolution.

Tree rings and carbon and oxygen isotopes have also been demonstrated to be useful proxies for GPP. For example, ^{13}C in C_3 woody plants could be a valuable carbon isotope for the study of

photosynthesis (Lavergne et al., 2022). Plant tissues have a ^{13}C-depleted signature compared to atmospheric CO_2 due to slower diffusion of $^{13}CO_2$ through stomata and preferential fixation of $^{12}CO_2$ by Rubisco, a process known as carbon isotope discrimination ($\Delta^{13}C$) (Park & Epstein, 1960). Short-term $\Delta^{13}C$ variations in C_3 leaves have been found to be influenced by environmental factors like soil moisture and precipitation (Diefendorf et al., 2010; Kohn, 2010), VPD (Lloyd & Farquhar, 1994), temperature, atmospheric pressure (Cornwell et al., 2018). Tree ring studies also suggest that $\Delta^{13}C$ in C_3 woody plants increases with rising CO_2 concentration (McCarroll et al., 2009). Then, GPP estimates have been derived from oxygen isotopic analyses (^{16}O, ^{17}O, and ^{18}O) (Farquhar et al., 1993; Ciais et al., 1997; Peylin et al., 1999; Cuntz et al., 2003). During photosynthesis, plants incorporate oxygen from H_2O into organic compounds. The ratio of ^{18}O to ^{16}O of water in leaves, which is influenced by source water and evaporative enrichment, gets reflected in the organic matter (Liu et al., 2023).
Finally, in addition to the $\Delta^{13}C$ analysis in tree rings, tree ring widths have been found to have a consistent correlation with GPP, informing on trends in tree growth and terrestrial carbon stocks (Babst et al., 2014; Tei et al., 2019). Tree ring width and stable isotope data have also been used to evaluate and constrain vegetation carbon uptake and tree growth in LSMs (Barichivich et al., 2021; Panek & Waring, 1997). Worldwide tree ring databases, such as the International Tree-ring Data Bank (ITRDB), can be used for LSMs benchmarking or to constrain model parameters through data assimilation.

Despite the extensive use of the proxies mentioned above, significant uncertainties remain in GPP and plant transpiration estimates (Anav et al., 2015; Wei et al., 2017). Moreover, the relationship between each proxy and GPP or plant transpiration is affected by spatial and temporal scales, along with specific uncertainties tied to each proxy. Therefore, there is a need to investigate additional proxies that can address the limitations associated with traditional methods. Ideally, these proxies should offer a more direct measurement of GPP and plant transpiration, unlike indices that inform on vegetation greenness for example. These proxies should not require partitioning between GPP and respiration fluxes, and should be measured at fine temporal resolution to capture short-term variations in GPP and plant transpiration. Additionally, they should allow for integration over large-scale areas, such as through atmospheric concentration measurements.

1.3.3 Carbonyl sulfide: a proxy for GPP and plant transpiration

COS is a trace gas with a mean atmospheric concentration about 10^6 times smaller than the one of CO_2, with approximately 480 parts per trillions for COS (ppt), compared to 420 parts per millions (ppm) for CO_2 in 2023. However, COS shares similarities with CO_2 that have led to an increasing interest in investigating its potential to constrain the carbon cycle. Indeed, COS shows a seasonal cycle of atmospheric concentrations that follows the one of CO_2, as recorded by measurement stations from the National Oceanic and Atmospheric Administration (NOAA) network (Montzka et al., 2007). This close seasonal cycle with a drawdown in concentrations in spring and summer in the Northern hemisphere is due to vegetation uptake, which is the main sink of COS as it is of CO_2 (Whelan et al., 2018).

Therefore, COS uptake has been proposed as a proxy of GPP as both COS and CO_2 are absorbed by plant leaves following a similar diffusion pathway from the atmosphere to the leaf interior, which is represented by the boundary layer, the stomatal, and the internal conductance (Figure 3). Besides, COS has the advantage of not having a respiration flux analogous to the one of CO_2, as COS is irreversibly hydrolysed by the carbonic anhydrase enzyme (CA) inside the leaf following the reaction (Stimler et al., 2010),

$$COS + H_2O \rightarrow CO_2 + H_2S \qquad \text{(Equation 1)}$$

Then, COS is mainly not emitted back to the atmosphere by plants, and at the canopy scale, measurements of COS vegetation flux do not require a partitioning of the net flux as it is the case between respiration and photosynthesis for CO_2.

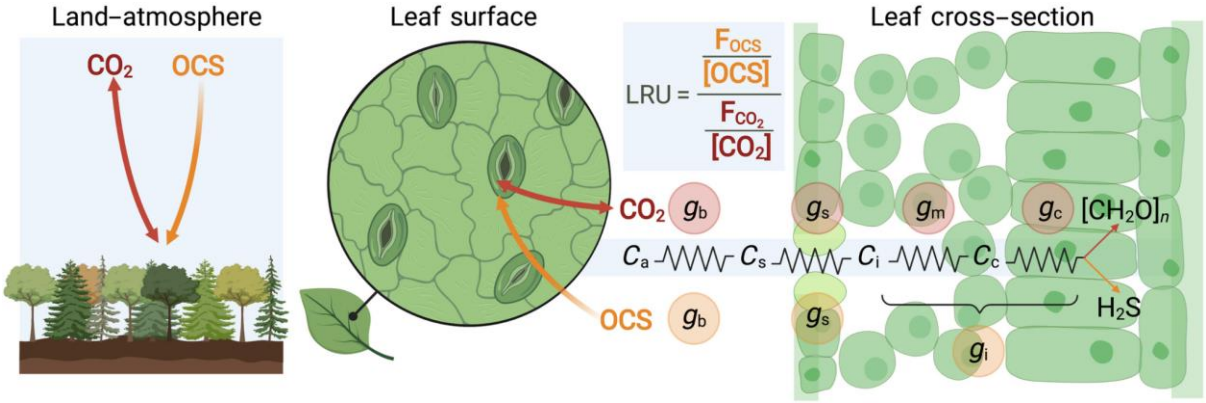

Figure 3: Relationship between CO_2 and COS (or OCS) vegetation uptake from the ecosystem to the leaf scale (Whelan et al., 2022). The diffusion pathway is characterized by a series of conductances with the boundary layer (g_b) and stomatal (g_s) conductances, followed by the internal conductance (g_i) for COS, and the mesophyll (g_m) and Rubisco reaction rate equivalent conductances (g_c) for CO_2. Along this diffusion pathway, the gradient of concentration is represented by the atmospheric (C_a), the leaf surface (C_s), the internal (C_i), and the chloroplast (C_c) concentrations.

In particular, COS leaf uptake has been more directly linked to stomatal conductance than GPP (Seibt et al., 2010; Wehr et al., 2017; Wohlfahrt et al., 2012; Berkelhammer et al., 2020). Indeed, COS internal conductance differs from the photosynthesis reaction in CO_2 internal conductance. Unlike CO_2 fixation by the Rubisco enzyme that depends on the energy from the light reaction part of photosynthesis, COS hydrolysis by CA is not light-dependent. Therefore, COS can provide additional information on GPP through its constraint on stomatal conductance (Asaf et al., 2013; Commane et al., 2015; Maseyk et al., 2014; Spielmann et al., 2019). Then, because stomatal conductance determines the coupling between GPP and plant transpiration, COS potential to constrain plant transpiration has also been investigated (Wehr et al., 2017; Wohlfahrt et al., 2012). A first relationship has been established to derive vegetation CO_2 uptake from COS, which is defined by the leaf relative uptake (LRU) that is the ratio of COS and CO_2 leaf uptakes normalized by their concentrations (Sandoval-Soto et al., 2005) (Figure 3). Once the LRU value has been

experimentally estimated, GPP can be inferred from measurements of COS flux, and COS and CO_2 concentrations. However, this extensively adopted approach has limited accuracy as the LRU values were found to vary with plant species (Sandoval-Soto et al., 2005; Stimler et al., 2012), atmospheric humidity (Sun et al., 2018; Kooijmans et al., 2019), and light (Stimler et al., 2011; Maseyk et al., 2014; Commane et al., 2015; Kooijmans et al., 2019).

The vegetation one-way COS flux also benefits from being spatially separated from the main COS sources (Whelan et al., 2018). Indeed, the ocean and anthropogenic activities are the main contributors to COS emissions. COS can be directly emitted by the ocean, or be indirectly produced from the oxidation of carbon disulfide (CS_2) or dimethyl sulfide (DMS), representing a total oceanic contribution estimated between +200 and +400 GgS y^{-1} (Lennartz et al., 2017, 2020). Anthropogenic emissions of the same order of magnitude have been evaluated between +220 and +580 GgS y^{-1}, with the main COS source originating from the viscose industry through the oxidation of CS_2 (Zumkehr et al., 2018). COS emissions of lower magnitude also come from biomass burning with +50 to +168 GgS y^{-1} (Campbell et al., 2015; Glatthor et al., 2017; Stinecipher et al., 2019), and volcanoes with +25 to +43 GgS y^{-1} (Whelan et al., 2018). Then, the second largest COS net sink is the soil with estimates ranging from -409 to -30 GgS y^{-1} (Kettle et al., 2002; Berry et al., 2013; Launois et al., 2015; Kooijmans et al., 2021; Abadie et al., 2022). Soils are mainly involved in COS absorption because the CA enzyme is present in soil microorganisms and drives COS consumption (Masaki et al., 2021), but soils can also emit COS under specific temperature and light conditions (Whelan & Rhew, 2015; Whelan et al., 2016, 2018; Kitz et al., 2017, 2020). Finally, the atmosphere is a small COS sink, through oxidation by the OH radical in the troposphere representing -130 to -82 GgS y^{-1}, and photolysis in the stratosphere with -80 to -30 GgS y^{-1}. A synthesis of COS sinks and sources is presented in Table 1.

Table 1: COS sinks and sources. Negative(/positive) estimates correspond to net uptake(/emission).

	Contribution	Estimated range (GgS y^{-1})	References
Net sinks	Vegetation	[-1335; -238]	Kettle et al. (2002); Montzka et al. (2007); Suntharalingam et al. (2008); Berry et al. (2013); Launois et al. (2015); Maignan et al. (2021); Kooijmans et al. (2021)
	Troposphere	[-130; -82]	Whelan et al. (2018)
	Stratosphere	[-80; -30]	Whelan et al. (2018)
	Total soil	[-409; -30]	Kettle et al. (2002); Berry et al. (2013); Launois et al. (2015); Kooijmans et al. (2021); Abadie et al. (2022)

		Oxic soils	[-510; -89]	Kettle et al. (2002); Berry et al. (2013); Launois et al. (2015); Kooijmans et al. (2021); Abadie et al. (2022)
Net sources		Anoxic soils	[+26; +101]	Kettle et al. (2002); Launois et al. (2015); Abadie et al. (2022)
	Total ocean		[+200; +400]	Lennartz et al. (2017, 2020)
		COS direct	[+50; +210]	Lennartz et al. (2017)
		COS from CS_2	[+5; +265]	Lennartz et al. (2017)
		COS from DMS	[+0; +80]	Lennartz et al. (2017)
	Anthropogenic		[+220; +580]	Zumkehr et al. (2018)
	Biomass burning		[+50; +168]	Campbell et al. (2015); Glatthor et al. (2017); Stinecipher et al. (2019)
	Volcanoes		[+25; +43]	Whelan et al. (2018)

The large range of estimates for the COS sink and source components leads to strong uncertainties in the global COS budget. Then, a better characterization of the different contributors to the global COS budget is required, including vegetation COS uptake to be able to fully exploit its potential as a proxy for GPP and plant transpiration. Currently, the global COS budget is not balanced, with a missing source likely located in the tropics and a missing biospheric sink in the high latitudes as suggested by atmospheric inversion studies (Ma et al., 2021; Remaud et al., 2022, 2023). Constraining the different COS sinks and sources is also of interest for the atmospheric sulfur cycle as COS is the longest-lived and most abundant sulfur-containing gas in the atmosphere.

1.4 Research questions and approach

Given the growing interest in COS measurements (Kooijmans et al., 2019; Kohonen et al., 2020; Vesala et al., 2022; Wehr et al., 2017; Berkelhammer et al., 2014; Commane et al., 2015; Kitz et al., 2019; Maseyk et al., 2014; Spielmann et al., 2019) and in the representation of land surface COS fluxes (Berry et al., 2013; Sun et al., 2015; Ogée et al., 2016), I decided to further explore this proxy to introduce new constraints on the simulated GPP and plant transpiration in the ORCHIDEE LSM. This work builds on the progress made during a first internship at LSCE, where I focused on integrating vegetation COS uptake in ORCHIDEE. However, the full potential of COS as a proxy for GPP and plant transpiration has yet to be quantified. Moreover, the additional constraints that COS can provide on these fluxes, beyond those captured by traditional eddy-covariance measurements of GPP and LE, need to be evaluated.

In this section, I present an overview of the manuscript structure, highlighting the main research questions and the approach taken to address them. Chapter 2 introduces the ORCHIDEE LSM, the key observation datasets utilized in this study, and the data assimilation tool used for parameter calibration. Then, chapters 3, 4, 5, and 6 each address the research questions outlined in the following. The connections between these chapters are illustrated in Figure 4.

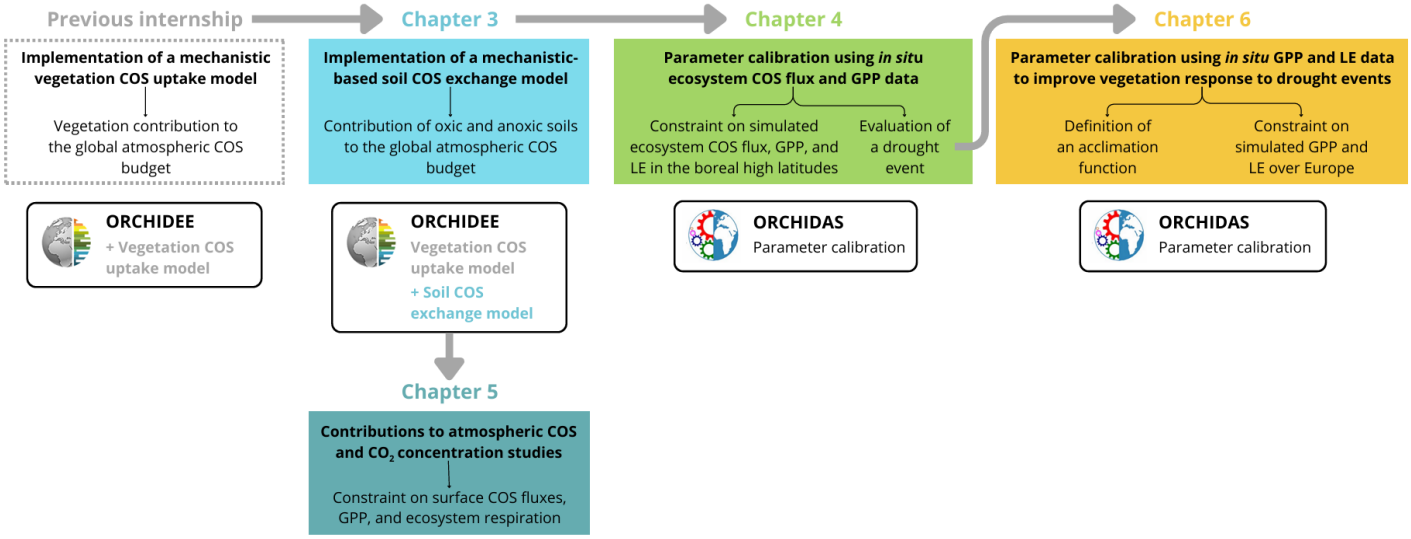

Figure 4: Overview of the main parts of the PhD, corresponding to the manuscript chapters. Large gray arrows indicate the connection between chapters.

Accurately characterizing the components of the global atmospheric COS budget is critical for using COS to constrain GPP and plant transpiration, and addressing the current imbalance in the COS budget. During a previous internship at LSCE before this PhD, I worked on implementing a vegetation COS uptake model in ORCHIDEE, aiming at exploiting the link between plant COS and CO_2 uptakes to constrain the simulated GPP using COS fluxes.
However, terrestrial biospheric COS fluxes also include the contribution from soils, which cannot be neglected as they have been estimated to be the second largest COS sink component (Section 1.3.3). Then, to complement the existing vegetation COS uptake model and simulate ecosystem COS fluxes at the global scale, a representation of soil COS exchanges needed to be implemented in ORCHIDEE. Unlike the first simple empirical soil COS flux models that estimated soil COS uptake while neglecting the soils' capacity to emit COS (Berry et al., 2013; Launois et al., 2015), two mechanistic models for soil COS exchanges (Ogée et al., 2015; Sun et al., 2015) were recently developed, based on fine scale soil measurements. These models account for the ability of soils to both take up and emit COS. Therefore, I implemented a mechanistic-based model of soil COS fluxes in ORCHIDEE, as presented in Chapter 3. This development enables simulating the spatial distribution and temporal variations of both oxic soil COS uptake and emission, as well as anoxic soil COS production. Moreover, this work allowed the evaluation of this mechanistic-based model, which involves parameters that depend on the vegetation type, across several sites representative of different biomes and climates. Indeed, this implementation of soil COS exchanges in ORCHIDEE, as well as the one made in parallel in another LSM (Simple Biosphere Model version 4, SiB4) based

on the same mechanistic soil COS flux model (Kooijmans et al., 2021), mark the first implementations of mechanistic-based representations of soil COS exchanges in LSMs.
Additionally, both vegetation and soil COS fluxes are driven by atmospheric COS concentration. However, previous works, including the first version of the vegetation COS uptake model in ORCHIDEE, considered a constant atmospheric COS concentration, neglecting the impact of its spatial and temporal variations on biospheric COS surface fluxes. Therefore, I included the consideration of a spatially and temporally varying atmospheric COS concentration in ORCHIDEE for computing vegetation and soil COS fluxes.
Consequently, in Chapter 3, I aim at answering the following main research questions:

- **What are the contributions of vegetation and soil COS fluxes to the global atmospheric COS budget based on mechanistic representations in a LSM?**
- **How does the contribution of anoxic soils compare to that of oxic soils in the global atmospheric COS budget?**
- **What is the impact of using spatially and temporally varying atmospheric COS concentrations, as opposed to a constant concentration, on the simulated surface vegetation and soil COS fluxes?**

In a second step, building on the developments presented in Chapter 3 to simulate ecosystem COS fluxes, I investigated how these simulated biospheric COS fluxes could be used to inform GPP and plant transpiration in ORCHIDEE. Although the process-based models represent an improvement over the LRU approach previously used to infer GPP from vegetation COS flux observations (Section 1.3.3), they include parameters that need to be calibrated. Therefore, I focused on using *in situ* biospheric COS flux observations to calibrate a set of ORCHIDEE parameters that influence GPP and plant transpiration in addition to vegetation COS fluxes, as presented in Chapter 4.
Ecosystem COS flux data, like GPP and LE estimates, can be obtained by eddy covariance measurements, which represent the sum of the vegetation and soil COS fluxes simulated in ORCHIDEE. However, multiple years of eddy covariance ecosystem COS flux data are available only at a few sites. Therefore, I chose to work on the Hyytiälä boreal forest in Finland, which provides the longest time series of *in situ* ecosystem COS flux measurements (Vesala et al., 2022). This site offers a large number of observation data under different environmental conditions, allowing for a more robust calibration of ORCHIDEE parameters.
Focusing on this site in the Northern high latitudes was also particularly interesting because previous studies have identified a missing COS sink in this region. Thus, I was able to assess the impact of ORCHIDEE parameter calibration using *in situ* ecosystem COS flux observations on the simulated biospheric COS sink estimate, in regard with the global atmospheric COS budget imbalance.
Therefore, the framework developed in Chapter 4 explores the additional constraint provided by ecosystem COS flux data to improve the representation of GPP and plant transpiration in ORCHIDEE through parameter calibration, and addresses these main research questions:

- **How can ecosystem COS flux observations constrain the simulated GPP and plant transpiration in ORCHIDEE?**
- **Can local ecosystem COS flux observations help refine regional biospheric COS flux estimates to address the global atmospheric COS budget imbalance?**

Following the COS model developments in ORCHIDEE and the use of simulated ecosystem COS fluxes described in Chapters 3 and 4, I was invited to contribute to several studies that relied on ORCHIDEE simulated biospheric COS fluxes and atmospheric COS concentrations. In Chapter 5, I expand on how I broadened my focus and expertise to use atmospheric COS and CO_2 concentrations for informing COS and CO_2 surface fluxes, with the aim of more accurately constraining the global atmospheric COS and CO_2 budgets.

The vegetation COS flux model implemented in ORCHIDEE assumes that vegetation can only act as a COS sink. The first study I contributed to tests this assumption by using local scale observations to investigate ecosystem COS flux processes. This work compares ORCHIDEE simulated ecosystem COS fluxes against field observations, including COS concentration measurements at two agroecosystem sites near LSCE. The identification of a potential missing crop COS source representation in ORCHIDEE led to a publication in which I am the second author.

While biospheric COS fluxes are the primary COS sinks, other components significantly contribute to terrestrial surface COS fluxes and need better constraints. Another study I participated in expanded the evaluation of surface COS fluxes at the regional scale, using a model that tracks atmospheric particle trajectories and COS concentration measurements at the Gif-sur-Yvette (GIF) site in France. This work helped identify discrepancies in anthropogenic COS emission estimates and in ORCHIDEE simulated biogenic COS fluxes.

In Chapter 3, I assessed the importance of considering spatially and temporally varying atmospheric COS concentrations to simulate vegetation and soil COS fluxes in ORCHIDEE. This was complemented by another study I contributed to, which investigated the impact of a recent decrease in atmospheric COS concentration, as observed at the GIF site, on simulated biogenic COS fluxes across the Northern hemisphere.

Evaluating surface COS fluxes can be achieved by transporting all surface COS flux components and comparing the resulting simulated concentrations with atmospheric COS concentration measurements. However, uncertainties related to the transport of these fluxes can limit this evaluation. To address this, I provided ORCHIDEE simulated biospheric COS fluxes for a study focusing on quantifying uncertainties in simulated atmospheric COS concentrations resulting from the transport of surface COS fluxes in seven atmospheric transport models (ATMs).

Following the improved characterization of surface COS fluxes from the studies presented above using atmospheric COS concentrations, we aimed to investigate how these concentrations can inform surface CO_2 fluxes, specifically GPP and ecosystem respiration, in addition to constraining surface COS fluxes. Expanding my focus beyond ORCHIDEE, I contributed to a joint assessment of atmospheric COS and CO_2 concentrations to evaluate potential seasonal amplitude and phase biases in simulated GPP and terrestrial ecosystem respiration from three model ensembles. This work also led to a publication in which I am the second author.

Finally, I participated in a study that not only evaluated surface COS and CO_2 fluxes using atmospheric COS and CO_2 concentrations but also optimized these fluxes using an atmospheric inversion modeling approach. This addressed the global COS budget imbalance issue.

Therefore, in Chapter 5, the studies I contributed to aim to answer the following main research questions:

- **How can atmospheric COS concentrations provide insights into the simulated ORCHIDEE biospheric COS fluxes and other surface COS component estimates to identify discrepancies and potential missing processes?**

- **How can atmospheric COS and CO_2 concentrations be used to constrain surface COS fluxes, GPP, and ecosystem respiration?**
- **Can a joint analysis of atmospheric COS and CO_2 concentrations help identify potential biases in GPP and terrestrial ecosystem respiration within model simulation ensembles?**

My work on COS, described in Chapters 3, 4, and 5, contributed to better constraining surface COS and CO_2 fluxes. Additionally, it allowed me to investigate specific GPP and plant transpiration processes implemented in ORCHIDEE. In Chapter 4, I focused on a drought event at the Hyytiälä site following ORCHIDEE parameter calibration using ecosystem COS flux data. This focus revealed structural deficiencies in the simulated GPP and LE under soil moisture stress conditions that could not be resolved solely through the calibration of existing model parameters within their physical range. This is critical as accurately modeling drought events in LSMs is crucial, given that their frequency and intensity are expected to increase due to climate change.

Therefore, in Chapter 6, I aimed to improve the representation of vegetation response to soil moisture stress in ORCHIDEE. For instance, the default physiological vegetation response to droughts in ORCHIDEE is not differentiated by vegetation type, although different species exhibit varied responses. Consequently, I investigated whether the default physiological vegetation response to soil moisture stress in ORCHIDEE could be refined to better account for the diversity of responses between biomes and the influence of environmental drivers. To achieve this, I calibrated a set of ORCHIDEE parameters, including those involved in vegetation response to soil moisture stress, using GPP and LE data at over 40 sites from the ICOS Warm Winter 2020 network. This dataset captured the impact of recent drought years across various biomes in Europe. This work led to the development of a new definition for the parameter determining the speed of stomatal closure under soil moisture stress conditions, incorporating vegetation's capacity to acclimate to long-term VPD conditions. Finally, I performed projections until 2100 to assess the impact of this refined parameter under future climate scenarios.

Consequently, in Chapter 6, I address the following main research questions:

- **How can the simulated physiological response of vegetation to drought events be improved in ORCHIDEE?**
- **How does a biome-dependent vegetation response to soil moisture stress compare to a response that also incorporates acclimation to long-term VPD conditions?**
- **How does incorporating vegetation acclimation to long-term VPD conditions in drought response simulations affect GPP and LE under a future climate scenario?**

2 DATA ASSIMILATION WITH ORCHIDEE

2.1 The ORCHIDEE land surface model

ORCHIDEE is the LSM component of the earth system model (ESM) developed at the Institut Pierre Simon Laplace (IPSL) (Boucher et al., 2020; Krinner et al., 2005). It simulates the carbon, water, and energy exchanges between the land surface and the atmosphere. ORCHIDEE can be run coupled with an atmospheric model, or forced with prescribed meteorological fields (offline mode), which was the configuration used during my PhD. Simulations can be performed from the site scale to the global scale.

Vegetation is characterized by plant functional types (PFTs), grouping plants with similar structure, phenology, photosynthetic pathway, and climate into 14 vegetated classes, with another PFT dedicated to bare soil. Each model grid cell is represented by fractions of PFTs that can be either imposed by ORCHIDEE users at the site scale, or prescribed using yearly varying vegetation maps at the global scale. Two global soil maps can be used to describe soil textures, which determine thermal and hydraulic soil properties like porosity and moisture thresholds, including the wilting point and field capacity for plant water uptake. The global map of soil textures can be derived from a simplification of the Zobler texture classification (Zobler, 1986) into three different soil texture classes. Alternatively, a more detailed global map is based on the Food and Agriculture Organization of the United Nations-United States Department of Agriculture (FAO-USDA) texture classification, which includes 12 texture classes (Reynolds et al., 2000). I have used the FAO-USDA map in ORCHIDEE as it offers a finer representation of soil textures.

Fast processes such as photosynthesis, plant transpiration and the energy budget are computed at a half-hourly time step, while slower processes related to phenology and carbon allocation are computed at a daily time step. The representation of GPP and plant transpiration in ORCHIDEE and their main drivers are illustrated in Figure 5.

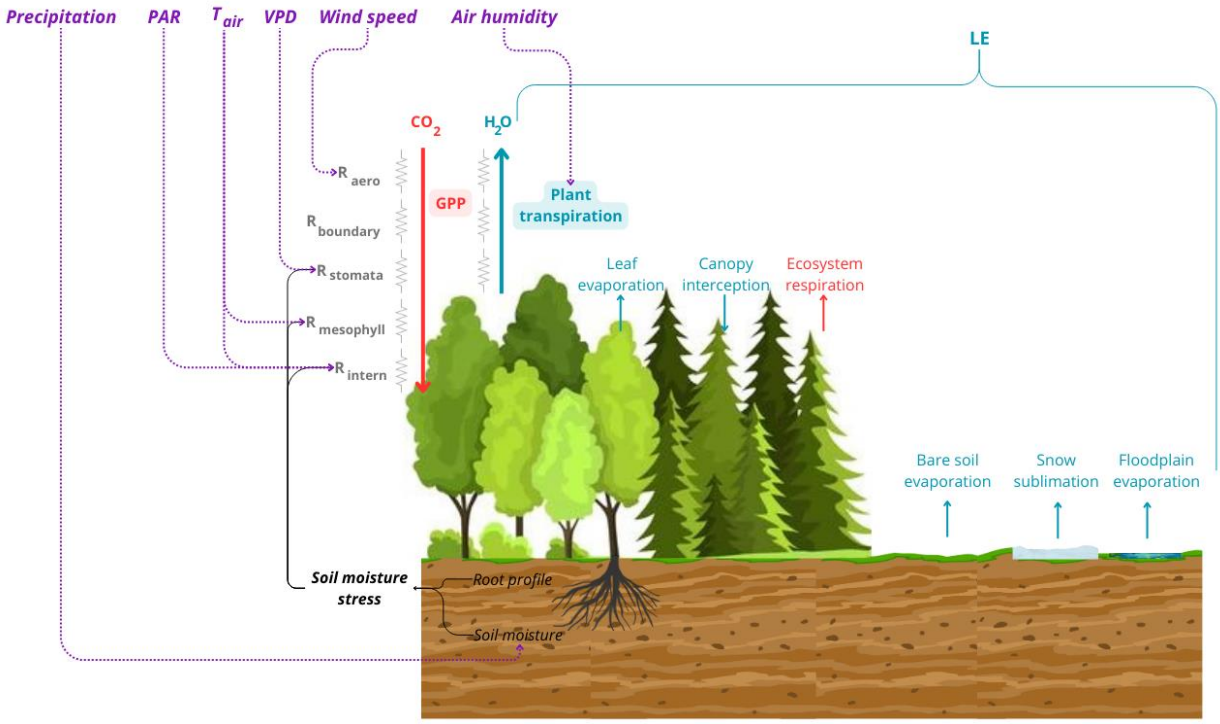

Figure 5: Simplified representations of GPP and plant transpiration in ORCHIDEE along with their main drivers. Blue arrows indicate water cycle fluxes, while red arrows represent carbon cycle fluxes. The resistance scheme for plant transpiration involves the aerodynamic resistance (R_{aero}), the leaf boundary layer resistance ($R_{boundary}$), the stomatal resistance ($R_{stomata}$). Additionally, the resistance scheme for GPP incorporates the mesophyll resistance ($R_{mesophyll}$) and the intern resistance (R_{intern}), which represents Rubisco activity. The main drivers influencing GPP and plant transpiration are depicted in purple. PAR: Photosynthetically Active Radiation; T_{air}: air temperature; VPD: Vapor Pressure Deficit.

Photosynthetic CO_2 assimilation is calculated for each PFT at the leaf level using the analytical algorithm described in Yin & Struik (2009), based on the photosynthesis model of Farquhar et al. (1980) for C_3 species and Collatz et al. (1992) for C_4 species. The Yin & Struik (2009) CO_2 assimilation model also integrates a description of stomatal and mesophyll conductances. Following Farquhar et al. (1980), the photosynthesis rate is expressed as the minimum between the Rubisco-limited and the electron transport-limited rates of CO_2 assimilation. The rates of maximum Rubisco carboxylation and RuBP regeneration are influenced by leaf age (Ishida et al., 1999; Krinner et al., 2005), and by an implicit nitrogen content, which diminishes deeper into the canopy (Johnson & Thornley, 1984). Air temperature also affects these rates through a modified Arrhenius function that incorporates acclimation to growth temperature, following Kattge & Knorr (2007). Furthermore, a constraint is applied to these maximum rates to reflect the downregulation of the productivity in response to elevated CO_2 concentrations (Sellers et al., 1996; Bounoua et al., 1999, 2010).

Stomatal conductance is determined by a residual stomatal conductance when the irradiance approaches zero, the intercellular CO_2 partial pressure and its compensation point in the absence of day respiration, and accounts for the effect of instantaneous changes in VPD, which reduces stomatal conductance when VPD increases. On the other hand, the mesophyll conductance is assumed to be only influenced by air temperature.

In ORCHIDEE, a soil moisture stress function has been integrated to the Yin & Struik (2009) CO_2 assimilation model to limit stomatal conductance, mesophyll conductance, and the Rubisco carboxylation and RuBP regeneration rates as soil moisture decreases (Section 6.1). Then, leaf boundary layer and aerodynamic conductances also limit CO_2 assimilation following Su et al. (2001), mainly influenced by wind speed, LAI, and canopy height.

LAI, which corresponds to the one-sided green leaf area per unit ground surface area (m^2 m^{-2}) is a prognostic variable tightly linked to the carbon allocation scheme. The calculated LAI is used to vertically discretize the canopy into a maximum of 20 LAI layers, with increasing thickness from top to bottom. Then, the gross carbon assimilation computed at the leaf level in a given LAI layer is scaled-up to the canopy level by summing over all LAI layers.

Plant transpiration is computed for each PFT directly at the canopy level, driven by the saturated moisture gradient between the surface and the air at 2 m, and regulated by the aerodynamic, leaf boundary layer, and total canopy stomatal conductances. The total canopy stomatal conductance corresponds to the integration of the stomatal conductance computed in each LAI layer over all LAI layers.
As for GPP computation, plant transpiration is also influenced by soil moisture stress through the soil moisture stress function limiting stomatal conductance when soil moisture decreases.

LE is the sum of all processes contributing to this flux at the grid cell level considering the fraction of the grid cell concerned by each process, which are plant transpiration, bare soil evaporation, canopy interception and evaporation, snow sublimation, and floodplain evaporation. Similar to plant transpiration, each other process contributing to LE is driven by the saturated moisture gradient between the surface and the air at 2 m, controlled by the aerodynamic and boundary layer conductances, and by a process-specific conductance (as total canopy stomatal conductance for transpiration).

2.2 COS, GPP, and LE observation datasets

In this PhD thesis, I have used various observational datasets, ranging from *in situ* measurements for data assimilation and model evaluation to global products for large-scale evaluation of the upscaled optimized simulated fluxes. In addition to traditional observation-based data such as GPP and LE eddy-covariance measurements, this work incorporates a wide range of COS observations. These include small-scale observations obtained through flux chamber measurements and atmospheric concentration flask measurements for regional-scale studies. The observational datasets are presented in Figure 6.

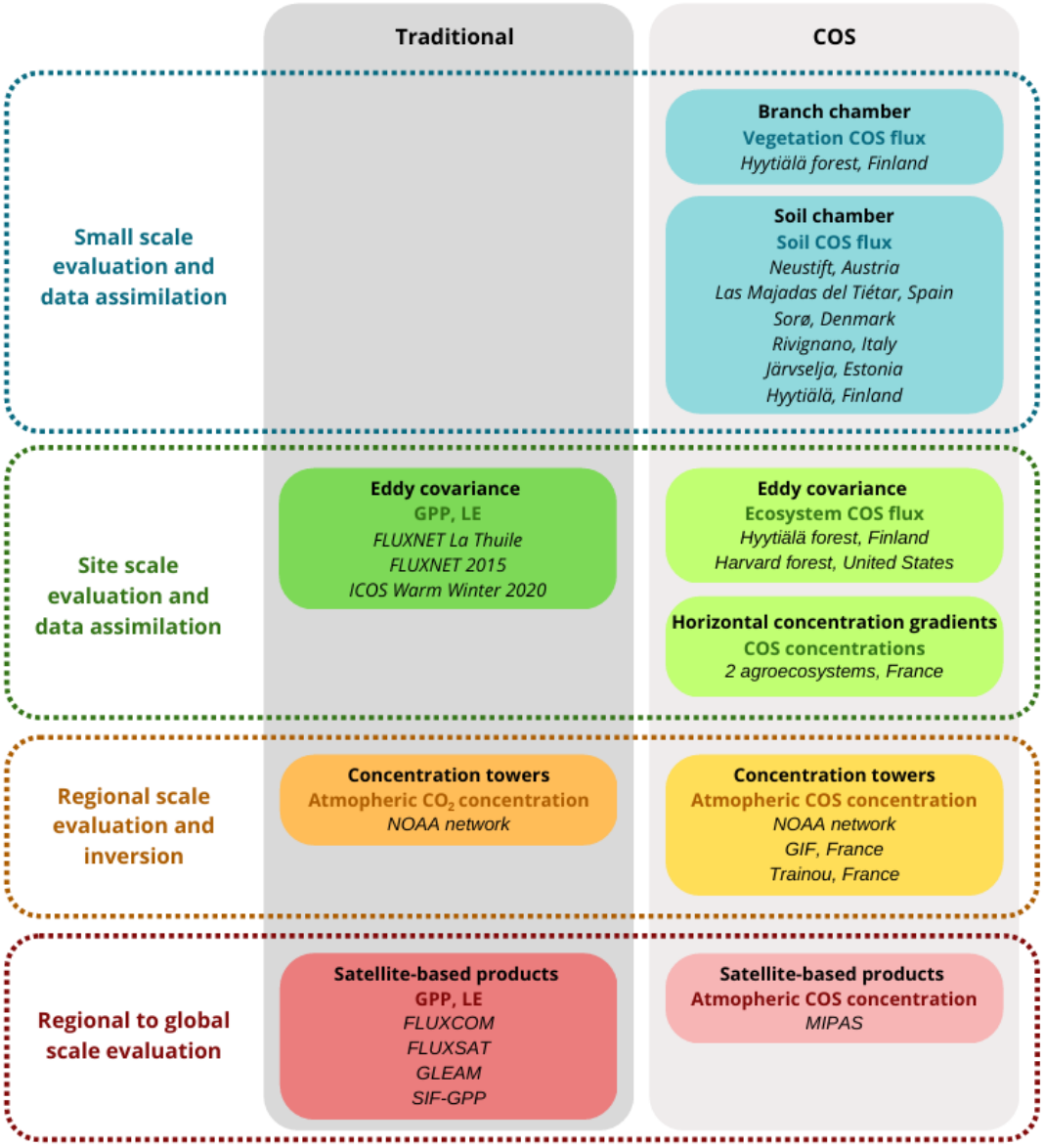

Figure 6: Overview of the main observation datasets used in this PhD work, presented at various spatial scales. The left column distinguishes datasets traditionally used for LSM evaluation and calibration, while the right column presents the COS observations.

GPP and LE data estimated using eddy covariance flux towers, which have been extensively used to benchmark LSMs (Section 1.3.1), are utilized in this study to calibrate ORCHIDEE parameters and evaluate model developments and calibrations. With the eddy covariance method, NEE is separated into GPP and ecosystem respiration (Re), for example by using night-time NEE data to estimate Re and subtracting it from NEE to isolate GPP (Reichstein et al., 2005). The FLUXNET La Thuile and 2015 networks have previously been used to calibrate and evaluate ORCHIDEE simulated carbon, water, and energy fluxes (Kuppel et al., 2012; Bastrikov et al., 2018, Peylin et al., 2016). Additionally, in this PhD work, I processed flux data from the ICOS Warm Winter 2020 database to provide new data for improving and evaluating ORCHIDEE fluxes over recent years in Europe.

Then, since direct GPP and LE observations are not available at regional and global scales, I have utilized several satellite-based global GPP and LE products for model evaluation at these broader

scales. The FLUXSAT GPP (Joiner et al., 2018) and FLUXCOM GPP and LE (Jung et al., 2019, 2020; Nelson et al., 2024) datasets are derived from machine learning algorithms that integrate data from eddy covariance flux towers with remote sensing observations. GLEAM also provides LE estimates by combining satellite observations with a model that integrates land surface and meteorological data (Miralles et al., 2011). Then, I used the SIF-GPP product from the Sen4GPP project (Duveiller et al., 2023), which derives GPP using linear empirical relationships with SIF data obtained from the ESA Sentinel-5P Tropospheric Monitoring Instrument (TROPOMI) (Guanter et al., 2021), with biome-specific calibrations. Using multiple global GPP and LE products helps account for uncertainties arising from different estimation methods and underlying assumptions when evaluating ORCHIDEE simulated fluxes.

To complement these commonly used GPP and LE data, I have incorporated various COS flux observation datasets. Unlike eddy covariance GPP estimates, eddy covariance COS flux measurements are typically not affected by the confounding influence of a vegetation emission flux. Using COS flux measurements also provides an independent constraint on GPP and LE, thereby reducing reliance on traditional GPP and LE eddy covariance data and offering complementary insights into these fluxes.
COS flux chamber measurements can be used to investigate COS flux processes at small scales. Leaf COS uptake can be measured using branch chambers (Seibt et al., 2010; Kooijmans et al., 2019). In particular, the Hyytiälä boreal evergreen needleleaf forest in Finland has been the subject of extensive COS flux measurements. Branch chambers measured COS fluxes on a Scots pine tree from March to July 2017 (Kooijmans et al., 2019). COS fluxes were derived from hourly changes in mole fractions when chambers were sealed, using a quantum cascade laser calibrated against reference standards (Kooijmans et al., 2016). Soil COS exchanges can also be inferred from chamber measurements. Soil COS flux chamber measurements have been conducted at several sites, covering a variety of biomes with forest, grassland, and agricultural sites (Kitz et al., 2020; Spielmann et al., 2019, 2020; Sun et al., 2018). While chambers offer the advantage of isolating individual components of the COS budget, they also have limitations beyond their small spatial scale. For example, some chamber components might emit COS, and using soil chambers necessitates removing the aboveground vegetation.
At the ecosystem scale, COS fluxes that include the contribution from both vegetation and soil are measured from eddy covariance flux towers. Several years of ecosystem COS flux measurements are available from 2013 to 2017 at the Hyytiälä forest (Vesala et al., 2022), and in 2012 and 2013 at the Harvard deciduous temperate forest in the United-States (Wehr et al., 2017).

Then, to address the lack of direct GPP observations and COS flux measurements at the regional scale, studies I contributed to used atmospheric CO_2 and COS concentration measurements to inform CO_2 and COS surface fluxes (Section 5). Analyzing atmospheric CO_2 and COS concentration data allows for the investigation of contributions from various components of the global CO_2 and COS budgets. The NOAA surface tower network has been providing long-term COS concentration measurements at 15 locations, with weekly to monthly frequencies, since 2000 (Montzka et al., 2007). These measurements, which also include CO_2 concentrations, are collected using paired flasks and analyzed with gas chromatography and mass spectrometry. Additionally, COS concentration measurements have been conducted at the GIF station in France from August 2014

to November 2021, and at the Trainou site, approximately 80 km south of GIF (Belviso et al., 2022a, 2022b). Finally, at the global scale, the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) spectrometer (Fisher et al., 2008), which is operated onboard of the ESA Environmental Satellite (ENVISAT), provides satellite retrievals of atmospheric COS concentration from 2002 to 2012 (Glatthor et al., 2015).

2.3 The ORCHIDEE data assimilation system

2.3.1 Data assimilation

Because the parameter values used in process representations of LSMs significantly contribute to the uncertainty of the simulated fluxes, a tool was developed at LSCE to optimize the parameters of the ORCHIDEE LSM. Data assimilation (DA) is an effective method for optimizing key parameters to enhance model accuracy and improve future climate predictions. The ORCHIDEE DA system (ORCHIDAS, https://orchidas.lsce.ipsl.fr/) integrates different DA techniques that rely on the minimization of a cost function (measuring the distance between model and observation) to determine a combination of optimized parameters (Bastrikov et al., 2018). Various observational data streams can be assimilated to constrain the model parameters, such as *in situ* or remote sensing data (Bacour et al., 2023; Kuppel et al., 2014; MacBean et al., 2022; Mahmud et al., 2021; Peylin et al., 2016). For each data stream, an observation operator is required in ORCHIDEE to simulate the corresponding variable, allowing the assimilation system to improve this simulated variable by reducing the misfit between the model and the observational data stream (Figure 7).

The cost function quantifies the discrepancy between observed and simulated variables and the deviation of optimized parameters from their prior values, considering the errors associated with all components. Assuming Gaussian distributions for model-data and parameter errors, the cost function is expressed as follows (Tarantola, 2005),

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

with y representing the observed variable and $H(x)$ the corresponding simulated variable. x_b is the prior, and x is the optimized parameter vector. R and B are the prior error covariance matrices for the observations and parameters, respectively. Note that R includes both the measurement and model structural errors. Only the diagonal elements are considered for R and B. The parameter error in B is defined as 15% of the parameter variation range, while the observation errors were set to the root mean squared deviation (RMSD) between observed data and the prior model simulations (Kuppel et al., 2012, 2013; Bacour et al., 2023).

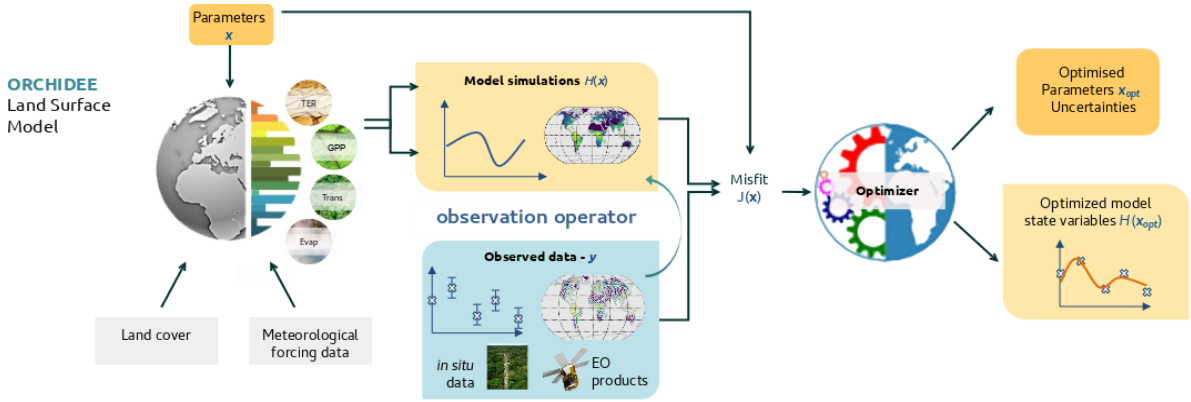

Figure 7: Overview of the data assimilation approach to optimize ORCHIDEE parameters using ORCHIDAS.

ORCHIDAS integrates two main minimization methods, a gradient-based approach (limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm with bound constraints, L-BFGS-B), and a global random search with a Genetic Algorithm (GA) approach (Bastrikov et al., 2018). In this work, I have used GA to optimize the model parameters (Goldberg, 1989; Haupt & Haupt, 2004; Santaren et al., 2014).

GA is inspired by genetics and natural selection (Goldberg, 1989; Haupt & Haupt, 2004). It treats the model parameter vector as a chromosome, with each gene representing a parameter. Iteratively, the algorithm generates a pool of a given number of chromosomes. The initial pool is created by randomly perturbing parameters, and subsequent iterations generate new chromosomes via two processes. New chromosomes can be generated through crossover, corresponding to an exchange of gene sequences between two parent chromosomes, or through mutation, randomly perturbing selected genes of a parent chromosome. The pool is then updated with the best chromosomes from both the parent and offspring pools, based on the lowest cost function values. Chromosomes are ranked by their cost function values, and the selection process ensures that the best chromosomes are more likely to produce offspring, adhering to the selection principle.

While GA is computationally more demanding than gradient-based approaches due to its slower convergence, it offers a significant advantage by considering a population of solutions and performing a global search. This thorough exploration of the search space helps avoid premature convergence. In contrast, gradient-based methods rely on a single solution trajectory, limiting the exploration of the search space and increasing the risk of converging to local optima. For example, in ORCHIDEE, Santaren et al. (2014) demonstrated that GA outperformed the gradient-based method in minimizing the cost function to the correct minimum at one FLUXNET site, constrained by water and carbon fluxes, as the gradient method frequently became trapped in local minima.

2.3.2 Sensitivity analysis

ORCHIDEE incorporates over 130 parameters to simulate land-surface interactions. To manage the computational cost of parameter optimization and to exclude parameters that do not significantly impact the variables of interest, sensitivity analyses (SAs) are conducted as a

preliminary step to DA. This process helps identify and select a subset of key parameters for calibration. Moreover, as ORCHIDEE is a LSM that can be used for global scale simulations and projections of land carbon, water, and energy fluxes, we aim to achieve genericity in the optimized parameter values. For instance, in ORCHIDEE, plant species are categorized into 14 vegetated PFTs, each associated with processes that depend on PFT-specific parameters. Ensuring parameter value genericity allows each PFT-dependent parameter to be applicable across different environmental conditions within a PFT. Therefore, by selecting key parameters for optimization, we reduce the risk of overfitting, where the optimized parameters fit the assimilated observational data too closely but fail to generalize well to varying conditions or independent validation data.

ORCHIDAS integrates two SA methods to test the impact of each parameter on a simulated variable, the Morris method (Morris, 1991) and the Sobol method (Sobol et al., 1993). The Sobol method provides a quantitative measure of the impact of each parameter and their interactions on the variable variance, but it is computationally expensive. In this PhD work, I have used the Morris method due to its time efficiency. The Morris method enables a qualitative identification of key parameters by ranking them according to their importance for the simulated variable.

3 REPRESENTING ECOSYSTEM COS FLUXES IN ORCHIDEE

3.1 Vegetation COS uptake

Vegetation COS uptake is estimated to be the largest sink for COS, but it also exhibits the widest range of estimates among all COS budget components, spanning from -1335 to -238 GgS y^{-1} (Table 1). This significant uncertainty in the vegetation contribution to the global COS budget highlights the need for a deeper understanding and tighter constraints on the spatial and temporal dynamics of vegetation COS fluxes to effectively use COS uptake as a proxy for GPP and plant transpiration. Berry et al. (2013) initially developed a mechanistic model of vegetation COS flux, and its implementation in LSMs can yield new global estimates of vegetation COS uptake.
GPP can be inferred from observed vegetation COS flux using the LRU relationship (Section 1.3.3). A previous study by Launois et al. (2015), which considered three LSMs including ORCHIDEE, linked vegetation COS fluxes to simulated global GPP using the LRU approach. However, the integration of a mechanistic vegetation COS model into LSMs offers the potential to provide new insights into the simulated GPP and LE fluxes by constraining the model parameters that govern these three fluxes through the use of COS flux measurements and data assimilation techniques.

Therefore, during a 3-month internship at LSCE prior to my PhD thesis, I worked on implementing the mechanistic model of vegetation COS uptake initially developed by Berry et al. (2013) in ORCHIDEE. This work, detailed in Maignan et al. (2021), where I am the second author and presented in Appendix 9.1, involved my contributions to the vegetation COS model implementation, performing simulations from site to global scales, and participating in the analysis of results and the writing of the article.
In this model, vegetation COS flux is represented by a one-way diffusion equation from the atmosphere to the leaf interior, limited by a series of conductances,

$$Fcos_{veg} = [COS]_{atm} \cdot \left[\frac{1}{gb_{COS}} + \frac{1}{gs_{COS}} + \frac{1}{gi_{COS}}\right]^{-1} \quad \text{(Equation 2)}$$

with $Fcos_{veg}$ the vegetation COS uptake (pmol COS m^{-2} s^{-1}), $[COS]_{atm}$ the atmospheric COS concentration (ppt), and gb_{COS}, gs_{COS}, gi_{COS} respectively the boundary layer, stomatal, and internal COS conductance (mol COS m^{-2} s^{-1}).
Contrary to the empirically based LRU approach relying on a constant LRU value per PFT as previously used in ORCHIDEE (Launois et al., 2015), this mechanistic model accounts for the influence of environmental factors, such as light, on the relationship between GPP and vegetation COS uptake.

The simulated vegetation COS fluxes in ORCHIDEE were evaluated against *in situ* measurements at two forest sites: Harvard in the United-States and Hyytiälä in Finland. This evaluation yielded relative RMSDs between 21% and 41% at both diel and seasonal scales across the two sites. Global simulations (Figure 8) estimated vegetation COS uptake at −756 GgS y^{-1} over the 2000-2009 period, in the range of previous estimates (Table 1). Additionally, new LRU values were derived for each

PFT, based on GPP and plant COS flux simulated in ORCHIDEE. The mechanistic and LRU approaches were compared through the transport of all COS components in the LMDZ atmospheric transport model and evaluation of the simulated COS concentrations at NOAA sites. Results showed that while the mechanistic approach was more accurate to represent high-temporal-resolution measurements, both methods produced similar results after transport of monthly mean COS fluxes. This study highlighted that uncertainties in the global COS budget are a larger limiting factor for using COS concentrations to constrain GPP in LSMs than the differences between the two modeling approaches.

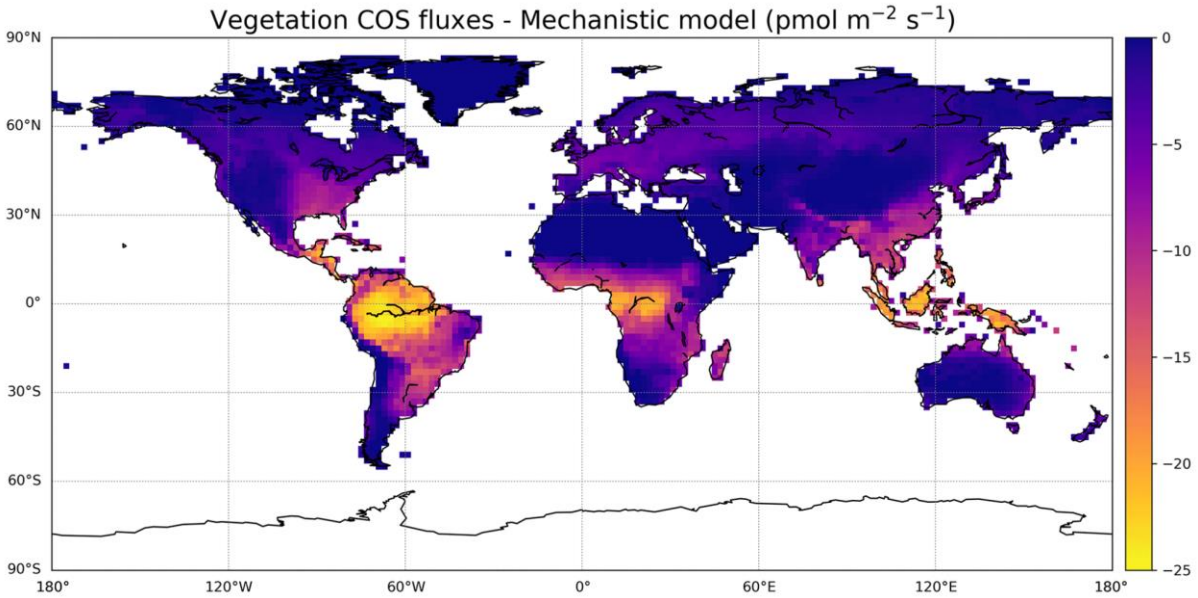

Figure 8: Map of average ORCHIDEE simulated vegetation COS fluxes (pmol m^{-2} s^{-1}) over the 2000-2009 period.

3.2 Soil COS exchanges

To further constrain the biospheric COS flux contribution to the global atmospheric COS budget and to accurately represent COS flux at the ecosystem scale, a representation of soil COS exchanges needed to be implemented in ORCHIDEE in addition to the vegetation COS model. Consequently, the implementation of a mechanistic-based model of soil COS uptake and production in ORCHIDEE has followed the implementation of the vegetation COS model. This work is presented in my first publication as the lead author in the next section.

Biogeosciences, 19, 2427–2463, 2022
https://doi.org/10.5194/bg-19-2427-2022

Biogeosciences Open Access EGU

Global modelling of soil carbonyl sulfide exchanges

Camille Abadie[1], Fabienne Maignan[1], Marine Remaud[1], Jérôme Ogée[2], J. Elliott Campbell[3], Mary E. Whelan[4], Florian Kitz[5], Felix M. Spielmann[5], Georg Wohlfahrt[5], Richard Wehr[6], Wu Sun[7], Nina Raoult[1], Ulli Seibt[8], Didier Hauglustaine[1], Sinikka T. Lennartz[9,10], Sauveur Belviso[1], David Montagne[11], and Philippe Peylin[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
[2]INRA, UMR 1391 ISPA, 33140 Villenave d'Ornon, France
[3]Sierra Nevada Research Institute, University of California, Merced, California 95343, USA
[4]Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey 08901, USA
[5]Department of Ecology, University of Innsbruck, Innsbruck, 6020, Austria
[6]Center for Atmospheric and Environmental Chemistry, Aerodyne Research, Inc., Billerica, Massachusetts 01821, USA
[7]Department of Global Ecology, Carnegie Institution for Science, Stanford, California 94305, USA
[8]Department of Atmospheric & Oceanic Sciences, University of California Los Angeles, Los Angeles, California 90095, USA
[9]Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, 26129 Oldenburg, Germany
[10]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[11]AgroParisTech, INRAE, Université Paris-Saclay, UMR ECOSYS, 78850 Thiverval-Grignon, France

Correspondence: Camille Abadie (camille.abadie.research@gmail.com)

Received: 22 October 2021 – Discussion started: 12 November 2021
Revised: 1 April 2022 – Accepted: 1 April 2022 – Published: 11 May 2022

Abstract. Carbonyl sulfide (COS) is an atmospheric trace gas of interest for C cycle research because COS uptake by continental vegetation is strongly related to terrestrial gross primary productivity (GPP), the largest and most uncertain flux in atmospheric CO_2 budgets. However, to use atmospheric COS as an additional tracer of GPP, an accurate quantification of COS exchange by soils is also needed. At present, the atmospheric COS budget is unbalanced globally, with total COS flux estimates from oxic and anoxic soils that vary between -409 and $-89\,\mathrm{GgS\,yr^{-1}}$. This uncertainty hampers the use of atmospheric COS concentrations to constrain GPP estimates through atmospheric transport inversions. In this study we implemented a mechanistic soil COS model in the ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) land surface model to simulate COS fluxes in oxic and anoxic soils. Evaluation of the model against flux measurements at seven sites yields a mean root mean square deviation of $1.6\,\mathrm{pmol\,m^{-2}\,s^{-1}}$, instead of $2\,\mathrm{pmol\,m^{-2}\,s^{-1}}$ when using a previous empirical approach that links soil COS uptake to soil heterotrophic respiration. However, soil COS model evaluation is still limited by the scarcity of observation sites and long-term measurement periods, with all sites located in a latitudinal band between 39 and 62° N and no observations during wintertime in this study. The new model predicts that, globally and over the 2009–2016 period, oxic soils act as a net uptake of $-126\,\mathrm{GgS\,yr^{-1}}$ and anoxic soils are a source of $+96\,\mathrm{GgS\,yr^{-1}}$, leading to a global net soil sink of only $-30\,\mathrm{GgS\,yr^{-1}}$, i.e. much smaller than previous estimates. The small magnitude of the soil fluxes suggests that the error in the COS budget is dominated by the much larger fluxes from plants, oceans, and industrial activities. The predicted spatial distribution of soil COS fluxes, with large emissions from oxic (up to $68.2\,\mathrm{pmol\,COS\,m^{-2}\,s^{-1}}$) and anoxic (up to $36.8\,\mathrm{pmol\,COS\,m^{-2}\,s^{-1}}$) soils in the tropics, especially in India and in the Sahel region, marginally improves the latitudinal gradient of atmospheric COS concentrations, after transport by the LMDZ (Laboratoire de Météorologie Dynamique) atmospheric transport model. The impact of different soil COS flux representations on the latitudinal gra-

Published by Copernicus Publications on behalf of the European Geosciences Union.

dient of the atmospheric COS concentrations is strongest in the Northern Hemisphere. We also implemented spatiotemporal variations in near-ground atmospheric COS concentrations in the modelling of biospheric COS fluxes, which helped reduce the imbalance of the atmospheric COS budget by lowering soil COS uptake by 10 % and plant COS uptake by 8 % globally (with a revised mean vegetation budget of $-576\,\mathrm{GgS\,yr^{-1}}$ over 2009–2016). Sensitivity analyses highlighted the different parameters to which each soil COS flux model is the most responsive, selected in a parameter optimization framework. Having both vegetation and soil COS fluxes modelled within ORCHIDEE opens the way for using observed ecosystem COS fluxes and larger-scale atmospheric COS mixing ratios to improve the simulated GPP, through data assimilation techniques.

1 Introduction

Carbonyl sulfide (COS) has been proposed as a tracer for constraining the simulated gross primary productivity (GPP) in land surface models (LSMs) (Launois et al., 2015; Remaud et al., 2022; Campbell et al., 2008). COS is an atmospheric trace gas that is scavenged by plants at the leaf level through stomatal uptake and irreversibly hydrolysed in a reaction catalysed by the enzyme carbonic anhydrase (CA) (Protoschill-Krebs et al., 1996). This enzyme also interacts with CO_2 inside leaves. COS and CO_2 follow a similar pathway from the atmosphere to the leaf interior. However, while CO_2 is also released during respiration, plants generally do not emit COS (Montzka et al., 2007; Sandoval-Soto et al., 2005; Wohlfahrt et al., 2012). To infer GPP at the regional scale using COS observations, modellers can use measurements of ecosystem COS fluxes directly or measurements of atmospheric COS concentrations combined with an atmospheric transport inversion model, provided all COS flux components are taken into account. In both cases, net soil COS flux estimates are needed, as well as a functional relationship between GPP and COS uptake by foliage.

One important limitation for using COS as a tracer for GPP is the uncertainty that remains on the COS budget components. Several atmospheric transport inversion studies have suggested that an unidentified COS source located over the tropics, of the order of 400–600 $\mathrm{GgS\,yr^{-1}}$, was needed to close the contemporary COS budget (Berry et al., 2013; Glatthor et al., 2015; Kuai et al., 2015; Ma et al., 2021; Remaud et al., 2022). It was recently estimated to account for 432 $\mathrm{GgS\,yr^{-1}}$ by Ma et al. (2021). The hypothesis of a strong tropical oceanic source has not been substantiated by in situ COS and CS_2 measurements in sea waters (Lennartz et al., 2017, 2020, 2021), except by Davidson et al. (2021), that invoke an oceanic source of $600 \pm 400\,\mathrm{GgS\,yr^{-1}}$ based on direct measurements of sulfur isotopes. Clearly, an accurate characterization of all flux components of the atmospheric COS budget is still needed. In particular, the contribution of soils to the COS budget is poorly constrained, and improved estimates of their contribution may therefore provide clues to the attribution of the missing source.

A distinction is usually made between oxic soils that mainly absorb COS and anoxic soils that emit COS (Whelan et al., 2018). Regarding COS uptake, COS diffuses into the soil, where it is hydrolysed by CA contained in soil microorganisms such as fungi and bacteria (Smith et al., 1999). It is to be noted that COS can also be consumed by other enzymes, like nitrogenase, CO dehydrogenase, or CS_2 hydrolase (Smith and Ferry, 2000; Masaki et al., 2021), but these enzymes are less ubiquitous than CA. The rate of uptake varies with soil type, temperature, and soil moisture (Kesselmeier et al., 1999; van Diest and Kesselmeier, 2008; Whelan et al., 2016). With high temperature or radiation, soils were also found to emit COS through thermal or photo degradation processes (Kitz et al., 2017, 2020; Whelan and Rhew, 2015; Whelan et al., 2016, 2018). Although such COS emissions can be large in some conditions, they have usually not been considered in atmospheric COS budgets.

Using the empirical relationship between soil COS uptake and soil respiration by Yi et al. (2007), Berry et al. (2013) provided new global estimates of COS uptake by oxic soils. Launois et al. (2015) proposed another empirical model, linking oxic soil COS uptake to H_2 deposition based on the correlation between these two processes observed at Gif-sur-Yvette (Belviso et al., 2013). Models with a physical representation of the involved processes are also available. Sun et al. (2015) proposed such a mechanistic model including COS diffusion and reactions within layered soil. Ogée et al. (2016) also developed a mechanistic model including both COS uptake and production, with steady-state analytical solutions in homogeneous soils. When including such models in an LSM, the challenge is to spatialize them, which requires new variables or parameters not readily available at the global scale but inferred from field or lab experiments.

In this study, our goal is to provide and evaluate new global estimates of net soil COS exchange. To this end, we did the following:

i. We implemented an empirical-based and a mechanistic-based soil COS model in the ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) LSM.

ii. We evaluated the soil COS models at seven sites against in situ flux measurements.

iii. We estimated soil contributions to the COS budget at the global scale.

iv. We transported all COS sources and sinks using an atmospheric model and evaluated the concentrations against measurements of the National Oceanic and Atmospheric Administration (NOAA) air sampling network.

2 Methods

2.1 Description of the models

2.1.1 The ORCHIDEE Land Surface Model

The ORCHIDEE Land Surface Model is developed at the Institut Pierre-Simon Laplace (IPSL). The model version used here is the one involved in the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Boucher et al., 2020; Cheruy et al., 2020). ORCHIDEE computes the carbon, water, and energy balances over land surfaces. It can be run at the site level or at the global scale. Fast processes such as soil hydrology, photosynthesis, and respiration are computed at a half-hourly time step. Other processes such as carbon allocation, leaf phenology, and soil carbon turnover are evaluated at a daily time step. Plant species are classified into 14 plant functional types (PFTs), according to their structure (trees, grasslands, or croplands), bioclimatic range (boreal, temperate, or tropical), leaf phenology (broadleaf or evergreen), and photosynthetic pathway (C_3 or C_4). The vegetation distribution in each grid cell is prescribed using yearly varying PFT maps, derived from the ESA Climate Change Initiative (CCI) land cover products (Poulter et al., 2015).

Soil parameters such as soil porosity, wilting point, and field capacity are derived from a global map of soil textures based on the FAO–USDA (Food and Agriculture Organization of the United Nations–United States Department of Agriculture) texture classification with 12 texture classes (Reynolds et al., 2000). The different textures for the USDA classification are presented in Table S1 in the Supplement. To better represent the observed soil conditions at the different sites that will be used for evaluation in this study, we substituted the soil textures initially assigned in ORCHIDEE from the USDA texture global map with the field soil textures translated into USDA texture classes (Table S2). In a previous study of vegetation COS fluxes in ORCHIDEE, Maignan et al. (2021) used the global soil map based on the Zobler texture classification (Zobler, 1986), which is reduced to three different textures in ORCHIDEE. However, the USDA soil classification gives a finer description of the different soil textures than the Zobler soil classification, considering 12 soil textures instead of 3. The move from the coarse Zobler classes to the finer USDA classes is found to be more important to the mechanistic model than to the empirical model. Since the USDA texture classes are more accurate with its finer discretization of soil textures, in the rest of this study, we only illustrate the results based on the USDA texture classification.

For site level simulations, the ORCHIDEE LSM was forced by local micro-meteorological measurements obtained from the FLUXNET network at the FLUXNET sites following the Creative Commons (CC-BY 4.0) license (Pastorello et al., 2020) and at the remaining sites by other local meteorological measurements performed together with the COS fluxes measurements when available, eventually gap-filled using the $0.25° \times 0.25°$ hourly reanalysis from the fifth generation of meteorological analyses of the European Centre for Medium-Range Weather Forecasts (ECMWF) (ERA5) (Hersbach et al., 2020). Global simulations were forced by the 0.5° and 6-hourly CRU JRA reanalysis (University of East Anglia Climatic Research Unit–Japanese Reanalysis; Friedlingstein et al., 2020). Near-surface COS concentrations (denoted C_a below) were prescribed using monthly mean atmospheric COS concentrations at the first vertical level of the LMDZ (Laboratoire de Météorologie Dynamique) atmospheric transport model (GCM, general circulation model; see description below in Sect. 2.1.3), forced with optimized COS surfaces fluxes. The latter have been inferred by atmospheric inverse modelling from the COS surface measurements of the NOAA network (Remaud et al., 2022). Simulations with constant atmospheric COS concentrations at a mean global value of 500 ppt were also run to evaluate the impact of spatiotemporal variations in near-surface COS concentrations versus a constant value. Near-surface CO_2 concentrations were estimated using global yearly mean values provided by the TRENDY (Trends in the land carbon cycle) project (Sitch et al., 2015).

2.1.2 COS soil models

The empirical soil COS flux model

We implemented in the ORCHIDEE LSM the soil COS flux model from Berry et al. (2013), which assumes that COS uptake is proportional to CO_2 production by soil respiration, following Yi et al. (2007). Although Yi et al. (2007) reported a relationship between soil COS uptake and total soil respiration, including root respiration, Berry et al. (2013) assumed that COS flux was proportional to soil heterotrophic respiration only. The rationale behind this assumption is that soil CA concentration is related to soil organic matter content and thus ecosystem productivity (Berry et al., 2013). As heterotrophic respiration is also linked to productivity, Berry et al. (2013) considered soil COS uptake to be proportional to soil heterotrophic respiration. However, soil respiration alone did not correlate well in incubation studies (Whelan et al., 2016). As the proportionality between COS fluxes and soil respiration has only been demonstrated for the total (heterotrophic and autotrophic) soil respiration (Yi et al., 2007), we used in this study total soil respiration as a scaling factor for soil COS uptake. This model will be referred to as the empirical model.

The influence of soil temperature and moisture are included in the calculation of soil respiration. Thus, we computed soil COS flux $F_{\text{soil, empirical}}$ (pmol COS $\text{m}^{-2}\,\text{s}^{-1}$) as follows:

$$F_{\text{soil, empirical}} = -k_{\text{soil}} \cdot \text{Resp}_{\text{tot}}, \quad (1)$$

where $\mathrm{Resp_{tot}}$ is total soil respiration (µmol $CO_2\,m^{-2}\,s^{-1}$) and k_{soil} is a constant equal to 1.2 pmol COS per µmol CO_2 that converts CO_2 production from respiration to COS uptake. The value of 1.2 pmol COS per µmol CO_2 was estimated from field chamber measurements in a pine and broadleaf mixed forest (Dinghushan Biosphere Reserve, southern China) from Yi et al. (2007). In ORCHIDEE, we calculated the total soil respiration as the sum of soil heterotrophic respiration within the soil column, including that of the litter, and root autotrophic respiration.

The mechanistic soil COS flux model

The mechanistic COS soil model of Ogée et al. (2016) describes both soil COS uptake and production. This model includes COS diffusion in the soil matrix, COS dissolution, and hydrolysis in the water-filled pore space and COS production under low redox conditions. The soil is assumed to be horizontally homogeneous so that the soil COS concentration C (mol m^{-3}) is only a function of time t (s) and soil depth z (m). The mass balance equation for COS can then be written as (Ogée et al., 2016)

$$\frac{\partial \varepsilon_{tot} C}{\partial t} = -\frac{\partial F_{diff}}{\partial z} + P - S, \tag{2}$$

where ε_{tot} is the soil total porosity (m^3 air per cubic metre soil), F_{diff} is the diffusional flux of COS (mol $m^{-2}\,s^{-1}$), S is the COS consumption rate (mol $m^{-3}\,s^{-1}$), and P the COS production rate under low redox conditions (mol $m^{-3}\,s^{-1}$).

Under steady-state conditions and uniform soil temperature, moisture, and porosity profiles, an analytical solution of Eq. (2) can be found (Ogée et al., 2016). We assume that the environmental conditions, such as soil temperature and moisture, are constant in ORCHIDEE over the 30 min model time step. We also assume chemical equilibrium between the gaseous and the dissolved COS, neglecting advection as suggested by Ogée et al. (2016). In these conditions, the typical timescale for COS diffusion in the upper active soil layer is much shorter than the 30 min model time step. Although Eq. (2) could also be solved numerically using the soil discretization in ORCHIDEE, we preferred to use the analytical solution, using the mean soil moisture and temperature averaged over the first few soil layers (down to about 9 cm deep), weighted by the thickness of each soil layer. Assuming fully mixed atmospheric conditions within and below the vegetated canopy, we also assumed that the COS concentration at the soil surface $C(z = 0)$ is equal to the near-surface COS concentration C_a. With these boundaries' conditions, the steady-state COS flux at the soil surface $F_{soil,\,mechanistic}$ (mol $m^{-2}\,s^{-1}$) is (Ogée et al., 2016)

$$F_{soil,\,mechanistic} = \sqrt{kB\theta D}\left(C_a - \frac{z_1^2 P}{D}\left(1 - \exp(-z_{max}/z_1)\right)\right), \tag{3}$$

where k is the first-order COS consumption rate constant within the soil (s^{-1}), B is the solubility of COS in water (m^3 water per cubic metre air), θ is the soil volumetric water content (m^3 water per cubic metre soil), D is the total effective COS diffusivity ($m^2\,s^{-1}$), $z_1 = \sqrt{D/kB\theta}$ (m), and z_{max} is the soil depth below which the COS production rate and the soil COS gradient are assumed negligible (Ogée et al., 2016). In the following, z_{max} is set at 0.09 m.

COS diffusion

The total effective COS diffusivity in soil D includes the effective diffusivity of gaseous COS $D_{eff,a}$ (m^3 air per metre soil per second) and dissolved COS $D_{eff,l}$ (m^3 water per metre soil per second) through the soil matrix:

$$D = D_{eff,a} + D_{eff,l}\,B. \tag{4}$$

The solubility of COS in water B is calculated using Henry's law constant K_H (mol m^{-3} Pa^{-1}):

$$B = K_H\,R\,T, \tag{5}$$

where $R = 8314\,J\,mol^{-1}\,K^{-1}$ is the ideal gas constant, T is the soil temperature (K), and (Wilhelm et al., 1977)

$$K_H = 0.00021\,\exp[24\,900/R(1/T - 1/298.15)]. \tag{6}$$

The effective diffusivity of gaseous COS $D_{eff,a}$ is expressed as (Ogée et al., 2016)

$$D_{eff,a} = D_{0,a}\,\tau_a\,\varepsilon_a, \tag{7}$$

where $D_{0,a}$ is the binary diffusivity of COS in the air (m^2 air s^{-1}), τ_a is the air tortuosity factor representing the tortuosity of the air-filled pores, and ε_a is the air-filled porosity (m^3 air per cubic metre soil). The binary diffusivity of COS in the air $D_{0,a}$ is expressed following the Chapman–Enskog theory for ideal gases (Bird et al., 2002) and depends on temperature and pressure:

$$D_{0,a}(T, p) = D_{0,a}(T_0 p_0)\left(\frac{T}{T_0}\right)^{1.5}\left(\frac{p}{p_0}\right), \tag{8}$$

where $D_{0,a}(T_0, p_0) = D_{0,a}(25\,°C, 1\,atm) = 1.27 \times 10^{-5}\,m^2\,s^{-1}$ (Massman, 1998).

The expression of the air tortuosity factor τ_a depends on whether the soil is repacked or undisturbed. In ORCHIDEE, repacked soils correspond to the agricultural soils represented by the C_3 and C_4 crops. Soils not covered by crops are considered undisturbed soils. The expression of τ_a for repacked soils $\tau_{a,r}$ is given by Moldrup et al. (2003):

$$\tau_{a,r} = \varepsilon_a^{3/2}/\varphi, \tag{9}$$

where φ is the soil porosity ($m^3\,m^{-3}$) that includes the air-filled and water-filled pores. Soil porosity is assumed constant through the soil column in ORCHIDEE and is determined by the USDA texture global map. The air-filled porosity ε_a is calculated as $\varepsilon_a = \varphi - \theta$.

The expression of τ_a for undisturbed soils $\tau_{a,u}$ is given in Deepagoda et al. (2011). We chose this expression rather than the expression proposed by Moldrup et al. (2003) for undisturbed soils because it appears to be more accurate and does not require information on the pore-size distribution (Ogée et al., 2016):

$$\tau_{a,u} = [0.2(\varepsilon_a/\varphi)^2 + 0.004]/\varphi. \quad (10)$$

In a similar way to COS diffusion in the gas phase, the effective diffusivity of dissolved COS $D_{eff,l}$ is described by Ogée et al. (2016):

$$D_{eff,l} = D_{0,l}\,\tau_l\,\theta, \quad (11)$$

where $D_{0,l}$ is the binary diffusivity of COS in the free water (m^2 water s^{-1}) and τ_l is the tortuosity factor for solute diffusion. The binary diffusivity of COS in the free water $D_{0,l}$ is described using an empirical formulation proposed by Zeebe (2011) for CO_2, which only depends on temperature:

$$D_{0,l}(T) = D_{0,l}(T_0)\left(\frac{T}{T_0} - 1\right)^2, \quad (12)$$

where $T_0 = 216\,K$ (Ogée et al., 2016) and $D_{0,l}(25\,°C) = 1.94 \times 10^{-9}\,m^2\,s^{-1}$ (Ulshöfer et al., 1996).

The expression of τ_l is the same for repacked and undisturbed soils. We used the expression given by Millington and Quirk (1961) as a good compromise between simplicity and accuracy (Moldrup et al., 2003):

$$\tau_l = \theta^{7/3}/\varphi^2. \quad (13)$$

COS consumption

COS can be destroyed by biotic and abiotic processes. The abiotic process corresponds to COS hydrolysis in soil water at an uncatalysed rate k_{uncat} (s^{-1}), which depends on soil temperature T (K) and pH (Elliott et al., 1989):

$$k_{uncat} = 2.15 \times 10^{-5} \exp\left(-10\,450\left(\frac{1}{T} - \frac{1}{298.15}\right)\right) + 12.7 \times 10^{-pK_w+pH} \exp\left(-6040\left(\frac{1}{T} - \frac{1}{298.15}\right)\right), \quad (14)$$

where pK_w is the dissociation constant of water.

This uncatalysed hydrolysis is quite low compared to the COS hydrolysis catalysed by soil microorganisms, which is the main contribution of COS uptake by soils (Kesselmeier et al., 1999; Sauze et al., 2017; Meredith et al., 2018). The enzymatic reaction catalysed by CA follows Michaelis–Menten kinetics. The turnover rate k_{cat} (s^{-1}) and the Michaelis–Menten constant K_m (mol m^{-3}) of this reaction depend on temperature. The temperature dependence of the ratio $\frac{k_{cat}}{K_m}$ is expressed as (Ogée et al., 2016)

$$x_{CA}(T) = \frac{\exp\left(-\frac{\Delta H_a}{RT}\right)}{1 + \exp\left(-\frac{\Delta H_d}{RT} + \frac{\Delta S_d}{R}\right)}, \quad (15)$$

where ΔH_a, ΔH_d, and ΔS_d are thermodynamic parameters, such as $\Delta H_a = 40\,kJ\,mol^{-1}$, $\Delta H_d = 200\,kJ\,mol^{-1}$, and $\Delta S_d = 660\,J\,mol^{-1}\,K^{-1}$.

The total COS consumption rate by soil k (s^{-1}) is described with respect to the uncatalysed rate at $T = 298.15\,K$ and pH = 4.5 (Ogée et al., 2016):

$$k = f_{CA} k_{uncat}\,(298.15, 4.5)\,\frac{x_{CA}\,(T)}{x_{CA}\,(298.15)}, \quad (16)$$

where f_{CA} is the CA enhancement factor, which characterizes the soil microbial community that can consume COS. The CA enhancement factor depends on soil CA concentration, temperature, and pH. Ogée et al. (2016) reported that its values range between 21 600 and 336 000, with a median value at 66 000. We adapted the values of f_{CA} found in Meredith et al. (2019) to have a CA enhancement factor that depends on ORCHIDEE biomes (Table A1 in Appendix A).

Oxic soil COS production

Abiotic oxic soil COS production has been observed at high soil temperature (Maseyk et al., 2014; Whelan and Rhew, 2015; Kitz et al., 2017, 2020; Spielmann et al., 2019a, 2020). However, photodegradation has also been proposed as an abiotic production mechanism in oxic soils (Whelan and Rhew, 2015; Kitz et al., 2017, 2020). Abiotic COS production is still not well understood but was assumed to originate from biotic precursors (Meredith et al., 2018).

In Ogée et al. (2016), the production rate P is described as independent of soil pH but depends on soil temperature and redox potential. This dependence on soil redox potential enables us to consider the transition between oxic and anoxic soils. However, because little information is available on soil redox potential at the global scale, its influence cannot yet be represented in a spatially and temporally dynamic way in a land surface model such as ORCHIDEE. Thus, we decided to use the production rate described in Whelan et al. (2016) that only depends on soil temperature and land use type:

$$P_{oxic} = e^{\alpha + \beta T}, \quad (17)$$

where P_{oxic} is expressed in pmol g^{-1} min^{-1}, T is soil temperature (°C), and α and β are parameters determined by Whelan et al. (2016) for each land use type using the least-squares fitting approach. We adapted the values of α and β given for four land use types to ORCHIDEE biomes (Table A2 in Appendix A). Values of α and β for deserts

could not be estimated by Whelan et al. (2016) because COS emission for this biome was not found to increase with temperature. Figure 11 in Whelan et al. (2016) shows that COS emission from a desert soil is always near zero for temperatures ranging from 10 to 40 °C. Moreover, COS emission from a desert soil is also found to be near zero in Fig. 1 of Meredith et al. (2018). This could be explained by a lack of organic precursors to produce COS (Whelan et al., 2016). Therefore, we considered that desert soils, which correspond to a specific non-vegetated PFT in ORCHIDEE, do not emit COS. For other ORCHIDEE biomes, COS production was estimated using α and β for each PFT and the mean soil temperature over the top 9 cm. The unit of P_{oxic} was converted from pmol g^{-1} min^{-1} to mol m^{-3} s^{-1} (in Eq. 3) using soil bulk density information from the Harmonized World Soil Database (HWSD; FAO/IIASA/ISRIC/ISSCAS/JRC, 2012).

Anoxic soil COS emission

Several studies have shown direct COS emissions by anoxic soils (Devai and DeLaune, 1997; de Mello and Hines, 1994; Whelan et al., 2013; Yi et al., 2007). This has been linked to a strong activity of sulfate reduction metabolisms in highly reduced environments such as wetlands (Aneja et al., 1981; Kanda et al., 1992; Whelan et al., 2013; Yi et al., 2007). A previous approach developed by Launois et al. (2015) was based on the representation of seasonal methane emissions by Wania et al. (2010) in the LPJ–WHyME (Lund–Potsdam–Jena–Wetland Hydrology and Methane) model to represent anoxic soils in ORCHIDEE. The mean values of soil COS emissions from Whelan et al. (2013) were used to attribute to each grid point a value of soil COS emission. In this approach by Launois et al. (2015), salt marshes were not represented despite their strong COS emissions found in Whelan et al. (2013). Emissions from rice paddies were also neglected. Thus, COS emissions from anoxic soils peaked in summer over the high latitudes, following methane production.

Because of the scarce knowledge on anoxic soil COS exchange, here we propose another approach to represent the contribution of anoxic soils, which could be compared to the previous approach developed by Launois et al. (2015). To represent the distribution of anoxic soils, we selected the regularly flooded wetlands from the map developed by Tootchi et al. (2019), as represented in Fig. 1. The regularly flooded wetlands cover 9.7 % of the global land area, which is among the average values found in the literature ranging from 3 % to 21 % (Tootchi et al., 2019). Then, in ORCHIDEE each pixel is considered either anoxic following the wetland map distribution from Tootchi et al. (2019) or oxic for the rest of the land surfaces. The pixels defined as anoxic soils are considered flooded through the entire year: the seasonal variations of the flooding, as happen during the monsoon seasons, are consequently neglected.

On anoxic pixels, we represent anoxic soil COS flux with a production rate based on the expression developed by Ogée et al. (2016):

$$P_{\text{anoxic}} = P_{\text{ref}} z_{\max} Q_{10}^{\frac{(T - T_{\text{ref}})}{10}}, \quad (18)$$

where P_{ref} (mol m^{-2} s^{-2}) is the reference production term, 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). As anoxic soil production ranges from 10 to 300 pmol m^2 s^{-1} for salt marshes and is usually below 10 pmol m^{-2} s^{-1} for freshwater wetlands (Whelan et al., 2018), the reference production term was set to 10 pmol m^{-2} s^{-1}.

All the variables and constants of the empirical and mechanistic models are presented in Tables A3 and A4 in Appendix A.

2.1.3 The atmospheric chemistry transport model LMDZ

To simulate the COS atmospheric distribution, we use an "offline" version of the Laboratoire de Météorologie Dynamique general circulation model (GCM), LMDZ 6 (Hourdin et al., 2020), which has been used as the atmospheric component in the IPSL coupled model for CMIP6. The LMDZ GCM has a spatial resolution of 3.75° long. × 1.9° lat. with 39 sigma-pressure layers extending from the surface to about 75 km, corresponding to a vertical resolution of about 200–300 m in the planetary boundary layer, and a first level at 33 m above sea or ground level. The model u and v wind components were nudged towards winds from the ERA5 reanalysis with a relaxation time of 2.5 h to ensure realistic wind advection (Hourdin and Issartel, 2000; Hauglustaine et al., 2004). The ECMWF fields are provided every 6 h and interpolated onto the LMDZ grid. This version has been shown to reasonably represent the transport of passive tracers (Remaud et al., 2018). The offline model uses pre-computed mass fluxes provided by this full LMDZ GCM version and only solves the continuity equation for the tracers, which significantly reduces the computation time. In the following, we refer to this offline version as LMDZ. The model time step is 30 min, and the output concentrations are 3-hourly averages.

The atmospheric COS oxidation is computed from pre-calculated OH monthly concentration fields produced from a simulation of the INCA (Interaction with Chemistry and Aerosols) model (Folberth et al., 2006; Hauglustaine et al., 2004, 2014) coupled to LMDZ. The atmospheric OH oxidation of COS amounts to 100 GgS yr^{-1} in the model. Similarly, the COS photolysis rates are also pre-calculated with the INCA model, which uses the Troposphere Ultraviolet and Visible (TUV) radiation model adapted for the stratosphere (Terrenoire et al., 2022). The temperature-dependent carbonyl sulfide absorption cross-sections from 186.1 to

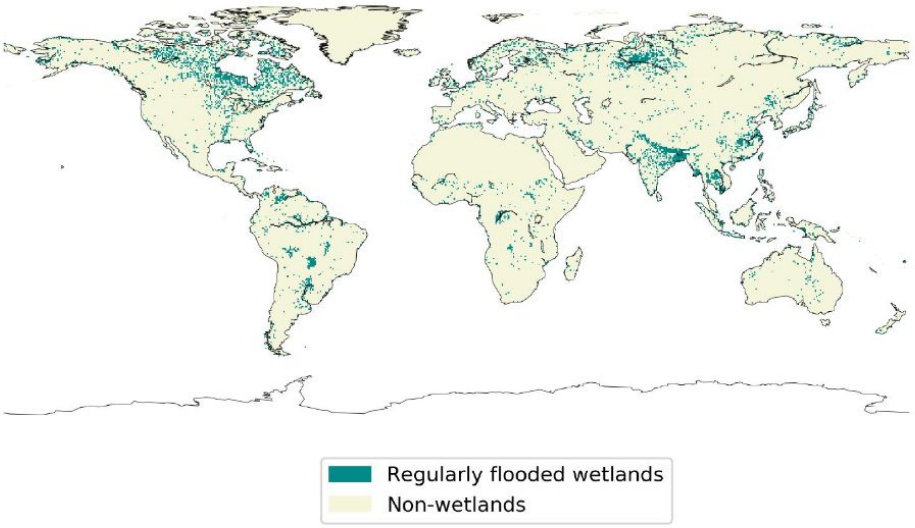

Figure 1. Map of wetlands distribution used to represent anoxic soils in ORCHIDEE. The map resolution is $0.5° \times 0.5°$ (adapted from Tootchi et al., 2019).

296.3 nm are taken from Burkholder et al. (2019). The calculated photolysis rates are averaged over the period 2008–2018 and prescribed to LMDZ. Implemented in LMDZ, the COS photolysis in the stratosphere amounts to about 30 GgS yr^{-1}, which is of the same order of magnitude as previous estimates: 21 GgS yr^{-1} (71 % of 30 GgS yr^{-1}) by Chin and Davis (1995), between 11 and 21 GgS yr^{-1} by Kettle et al. (2002), and between 16 and 40 GgS yr^{-1} by Ma et al. (2021).

2.2 Observation data sets

2.2.1 Description of the sites

The description of the studied sites is given in Table 1.

2.2.2 Soil COS flux determination at selected sites

Soil COS flux chamber measurements were conducted in 2015 at AT-NEU; in 2016 at DK-SOR, ES-LMA, and ET-JA; and in 2017 at IT-CRO (abbreviations as in Table 1). The aboveground vegetation was removed 1 d before the measurements if needed, and the fluxes were derived from concentration measurements using a quantum cascade laser (see Kitz et al., 2020, and Spielmann et al., 2020, 2019a). At AT-NEU, DK-SOR, ES-LMA, and IT-CRO, a random forest model was calibrated against the manual chamber measurements and then used to simulate half-hourly soil COS fluxes in Spielmann et al. (2019a). We compared the ORCHIDEE half-hourly simulated fluxes to half-hourly outputs of the random forest model. This enabled studying the diel cycle and computing daily observations with no sampling bias for the study of the seasonal cycle. Soil COS fluxes for ET-JA were derived by using the same training method as the one used in Spielmann et al. (2019a).

At FI-HYY, soil COS fluxes were measured using two automated soil chambers in 2015. These chambers were connected to a quantum cascade laser spectrometer to calculate soil COS fluxes from concentration measurements (see Sun et al., 2018, for more information on the experimental setup). Any vegetation was removed from the chambers before the measurements.

At US-HA, soil COS fluxes in 2012 and 2013 were not directly measured but derived from flux-profile measurements, connected to CO_2 soil chamber measurements and profiles. A sub-canopy flux gradient approach was used to partition canopy uptake from soil COS fluxes. For more information on this approach and its limitations, see Wehr et al. (2017).

In the study of soil COS fluxes, the difficulty of performing soil COS flux measurements must be acknowledged, as well as the differences between experimental setups and methods to retrieve soil COS fluxes. These limitations are illustrated in the set of observations selected here. Aboveground vegetation had to be removed at some sites to not measure the plant contribution in addition to soil COS fluxes (Sun et al., 2018; Spielmann et al., 2019a; Kitz et al., 2020). Vegetation removal prior to the measurements might lead to artefacts in the observations. Some components of the measuring system can also emit COS. In this case, a blank system is needed to apply a post-correction to the measured fluxes (Sun et al., 2018; Kitz et al., 2020). Litter was left in place at the measurement sites.

Table 1. Lists the sites' characteristics including their identification name, location, climate, soil type, dominant vegetation and species, corresponding PFT fractions we used for the ORCHIDEE simulations, and reference studies for more details. The spatial distribution of the sites is represented in Fig. B1 in Appendix B.

	Grassland	Savannah-like grassland	Deciduous broadleaf forest	Agricultural soybean field	Evergreen needleleaf forest	Boreal evergreen needleleaf forest	Temperate deciduous broadleaf forest
Country	Austria	Spain	Denmark	Italy	Estonia	Finland	United States
Sampling site	Neustift	Las Majadas del Tiétar	Sorø	Rivignano	Järvselja	Hyytiälä	Harvard
ID	AT-NEU	ES-LMA	DK-SOR	IT-CRO	ET-JA	FI-HYY	US-HA
Coordinates	47.12° N, 11.32° E	39.94° N, 5.77° W	55.49° N, 11.64° E	45.87° N, 13.08° E	58.22° N, 27.28° E	61.85° N, 24.30° E	42.54° N, 72.17° W
Climate	Humid continental	Mediterranean	Temperate maritime	Humid subtropical	Temperate	Boreal	Cool, moist temperate
Soil type	Fluvisol	Abruptic Luvisol	Luvisols or Chernozems	Silt loam	Haplic Gleysol	Haplic Podzol	Podzol and Regosol
Dominant vegetation	Graminoids: *Dactylis glomerata*, *Festuca pratensis* Forbs: *Ranunculus acris*, *Taraxacum officinale*	Tree: *Quercus ilex* Grass: *Vulpia bromoides*	European beech (*Fagus sylvatica*)	Soybean	Norway spruce (*Picea abies*)	Scots pine (*Pinus sylvestris*)	Red oak (*Quercus rubra*), red maple (*Acer rubrum*), hemlock (*Tsuga canadensis*)
ORCHIDEE PFT representation	100 % temperate natural grassland (C_3) (PFT 10)	20 % temperate broadleaf evergreen (PFT 5), 80 % temperate natural grassland (C_3) (PFT 10)	80 % boreal broadleaf summergreen (PFT 8), 20 % boreal natural grassland (C_3) (PFT 15)	100 % C_3 crops (PFT 12)	50 % boreal needleleaf evergreen (PFT 7), 40 % boreal broadleaf summergreen (PFT 8), 10 % boreal natural grassland (C_3) (PFT 15)	80 % boreal needleleaf evergreen (PFT 7), 20 % boreal natural grassland (C_3) (PFT 15)	80 % temperate broadleaf summergreen (PFT 6), 20 % of temperate natural grassland (C_3) (PFT 10)
References	Hörtnagl et al. (2011), Hörtnagl and Wohlfahrt (2014), Spielmann et al. (2019a), Kitz et al. (2020)	Lopez-Sangil et al. (2011), El-Madany et al. (2018), Weiner et al. (2018), Spielmann et al. (2019a), Kitz et al. (2020)	Pilegaard et al. (2011), Wu et al. (2013), Brændholt et al. (2018), Spielmann et al. (2019a), Kitz et al. (2020)	Spielmann et al. (2019a)	Noe et al. (2011, 2015), Kitz et al. (2020)	Kolari et al. (2009), Sun et al. (2018)	Urbanski et al. (2007), Wehr et al. (2017)

2.2.3 COS concentrations at the NOAA Earth System Research Laboratories (ESRL) sites

The NOAA surface flask network provides long-term measurements of the COS mole fraction at 14 locations at weekly to monthly frequencies from the year 2000 onwards. We use an extension of the data initially published in Montzka et al. (2007). The data were collected as paired flasks analysed using gas chromatography and mass spectrometry. The stations located in the Northern Hemisphere had sample air masses coming from the entire Northern Hemisphere domain above 30°. Among them, the sites LEF, NWR, HFM, and WIS have mostly continental footprints (Remaud et al., 2022), while the sites SPO, CGO, and PSA sample mainly oceanic air masses of the Southern Hemisphere (Montzka et al., 2007). The locations of these sites are depicted in Fig. B1 in Appendix B.

2.3 Simulations

2.3.1 Spin-up phase

A "spin-up" phase was performed before each simulation, which enabled all carbon pools to stabilize and the net biome production to oscillate around zero. Reaching the equilibrium state is accelerated in the ORCHIDEE LSM thanks to a pseudo-analytical iterative estimation of the carbon pools, as described in Lardy et al. (2011). For site simulations, the spin-up was performed by cycling the years available in the forcing files of each site, for a total of about 340 years. For global simulations, the spin-up phase of 340 years was performed by cycling over 10 years of meteorological forcing files in the absence of any disturbances.

2.3.2 Transient phase

Following the spin-up phase we ran a transient simulation of about 40 years that introduced disturbances such as climate change, land use change, and increasing CO_2 atmospheric concentrations.

This transient phase was performed by cycling over the available years for site simulations. For global simulations, the transient phase was run where we introduced disturbances from 1860 to 1900. After this transient phase, COS fluxes were simulated from 1901 to 2019.

2.3.3 Atmospheric simulations: sampling and data processing

We ran the LMDZ6 version of the atmospheric transport model described above for the years 2009 to 2016. We started from a uniform initial condition, and we removed the first year, as it is considered to be part of the spin-up period. The COS fluxes used as model inputs are presented in Table 2. The fluxes are given as a lower boundary condition, called the surface, of the atmospheric transport model (LMDZ), which then simulates the transport of COS by large-scale advection and sub-grid scale processes such as convection and boundary layer turbulence. In this study, we only evaluate the sensitivity of the latitudinal gradient and seasonal cycle of COS concentrations to the soil COS fluxes. The horizontal gradient aims at validating the latitudinal repartition of the surface fluxes, while the seasonal cycle partly reflects the seasonal exchange with the terrestrial sink, which peaks in spring/-summer. This study does not aim at reproducing the mean value, as the top-down COS budget is currently unbalanced, with a source component missing (Whelan et al., 2018; Remaud et al., 2022; see Table 3).

For each COS observation, the 3D simulated concentration fields were sampled at the nearest grid point to the station and at the closest hour of the measurements. For each station, the curve fitting procedure developed by the NOAA Climate Monitoring and Diagnostic Laboratory (NOAA CMDL) (Thoning et al., 1989) was applied to modelled and observed COS time series to extract a smooth detrended seasonal cycle. We first fitted a function including a first-order polynomial term for the growth rate and two harmonic terms for seasonal variations. The residuals (raw time series minus the smooth curve) were fitted using a low-pass filter with either 80 or 667 d as short-term and long-term cut-off values. The detrended seasonal cycle is defined as the smooth curve (full function plus short-term residuals) minus the trend curve (polynomial plus long-term residuals). Regarding vegetation COS fluxes (Maignan et al., 2021), we added the possibility of using spatially and temporally varying atmospheric COS concentrations, as for soil.

2.4 Numerical methods for model evaluation and parameter optimization

2.4.1 Statistical scores

We evaluated modelled soil COS fluxes against field measurements using the root mean square deviation (RMSD) as

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{n=1}^{N}\left(F_{\mathrm{COS}}^{\mathrm{Obs}}(n) - F_{\mathrm{COS}}^{\mathrm{Mod}}(n)\right)^2}{N}}, \tag{19}$$

where N is the number of considered observations, $F_{\mathrm{COS}}^{\mathrm{Obs}}(n)$ is the nth observed COS flux, and $F_{\mathrm{COS}}^{\mathrm{Mod}}(n)$ is the nth modelled COS flux, and the relative RMSD (rRMSD) as

$$\mathrm{rRMSD} = \frac{\mathrm{RMSD}}{\frac{\sum_{n=1}^{N} F_{\mathrm{COS}}^{\mathrm{Obs}}(n)}{N}}, \tag{20}$$

which is the RMSD divided by the mean value of observations.

Simulated atmospheric COS concentrations were evaluated by computing the normalized standard deviation (NSD),

Table 2. Sink and source components of COS budget used in this study. Mean magnitudes and standard deviations of different types of fluxes are given for the period 2009–2016.

Type of COS flux	Temporal resolution	Total ($GgS\,yr^{-1}$)	Standard deviation ($GgS\,yr^{-1}$)	Data source
Anthropogenic	Monthly, interannual	+394	21	Zumkehr et al. (2018) for which the fluxes for the year 2012 were repeated after 2012
Ocean	Monthly, interannual	+313	14	Lennartz et al. (2021) and Masotti et al. (2016) for indirect oceanic emissions (via CS2 – carbon disulfide – and DMS – dimethyl sulfide – respectively) and Lennartz et al. (2017) for direct oceanic emissions
Biomass burning	Monthly, interannual	+48	9	Stinecipher et al. (2019)
Soil	Monthly, interannual	See Table 3	5 (oxic) 2 (anoxic)	This work, including mechanistic and empirical approaches (Berry et al., 2013; Launois et al., 2015)
Vegetation uptake	Monthly, interannual	−576	7	Maignan et al. (2021)
Atmospheric OH oxidation	Monthly, interannual	−100	(–)	Hauglustaine et al. (2004)
Photolysis in the stratosphere	Monthly, interannual	−30	(–)	Remaud et al. (2022)

Table 3. Comparison of soil COS budget per year ($GgS\,yr^{-1}$). The net total COS budget is computed by adding all sources and sinks of COS (anthropogenic, ocean, biomass burning, soils, vegetation, atmospheric OH oxidation, and photolysis in the atmosphere) used to transport COS fluxes (Table 2). CLM: Community Land Model. SiB: Simple Biosphere Model.

	Kettle et al. (2002)	Berry et al. (2013)	Launois et al. (2015)			Kooijmans et al. (2021)	This study	
			ORCHIDEE	LPJ	CLM4	SiB4 (modified)	Empirical soil model	Mechanistic soil model
Period	2002	2002–2005	2006–2009			2000–2020	2009–2016	
Plants	−238	−738	−1335	−1069	−930	−664	−576	
Soil oxic	−130	−355	−510			−89	−214	−126
Soil anoxic	+26	Neglected	+101			Neglected	Neglected	+96
Soil total	−104	−355	−409			−89	−214	−30
Net total	+64	+1	−566	−300	−161	(–)	−165	+19

which is the standard deviation of the simulated concentrations divided by the mean of the observed concentrations, and the Pearson correlation coefficients (r) between simulated and observed COS concentrations. The closer NSD and r values are to 1, the better the model accuracy is.

2.4.2 Data assimilation

One of the main difficulties with the implementation of a model is to define the parameter values that lead to the most accurate representation of the processes in ORCHIDEE. Calibrating the model parameters is of interest as Ogée et al. (2016) indicate that some of the model parameters such as f_{CA} and the production term parameters have to be constrained by observations. Moreover, the default values for the soil COS model parameters used in this study (Tables A1 and A2 in Appendix A) are determined by laboratory experiments (Ogée et al., 2016; Whelan et al., 2016), which is why it is interesting to study how the values obtained by calibration against field observations differ from these default values. Data assimilation (DA) aims at producing an optimal estimate by combining observations and model outputs. In this study, we used DA to find the model parameter values that improve the fit between simulated and observed soil COS fluxes from the empirical and the mechanistic models. We

used the ORCHIDEE Data Assimilation System (ORCHIDAS), which is based on a Bayesian framework. ORCHIDAS has been described in detail in previous studies (Bastrikov et al., 2018; Kuppel et al., 2014; MacBean et al., 2018; Peylin et al., 2016; Raoult et al., 2021), so below we only briefly present the method. Assuming that the observations and model outputs follow a Gaussian distribution, we aim at minimizing the following cost function $J(\boldsymbol{x})$ by optimizing the model parameters (Tarantola, 2005):

$$J(\boldsymbol{x}) = \frac{1}{2}\Big[(M(\boldsymbol{x}) - y)^T \cdot E^{-1} \cdot (M(\boldsymbol{x}) - y) + (\boldsymbol{x} + \boldsymbol{x}^b)^T \cdot B^{-1} \cdot (\boldsymbol{x} + x^b)\Big], \quad (21)$$

where $\boldsymbol{x}$ is the vector of parameters to optimize and y is the observations. The first part of the cost function measures the mismatch between the observations and the model, and the second part represents the mismatch between the prior parameter values $\boldsymbol{x}^b$ and the considered set of parameters $\boldsymbol{x}$. Both terms of the cost function are weighted by the prior covariance matrices for the observation errors E^{-1} and parameter errors B^{-1}. The minimization of the cost function follows the genetic algorithm (GA) method, which is derived from the principles of genetics and natural selection (Goldberg, 1989; Haupt and Haupt, 2004) and is described for ORCHIDAS in Bastrikov et al. (2018).

For each soil COS model, we selected the eight most important parameters to which soil COS fluxes are sensitive following sensitivity analyses (Sect. 2.4.3). The observation sites selected for sensitivity analyses and DA are the ones with the largest number of observations for model parameter calibration, which are FI-HYY and US-HA.

2.4.3 Sensitivity analyses

We conducted sensitivity analyses at two contrasting sites (FI-HYY and US-HA) to determine which model parameters have the most influence on the simulated soil COS fluxes from the empirical and the mechanistic models. Sensitivity analyses can help to identify the key parameters before aiming at calibrating these parameters. Indeed, focusing on the key model parameters for calibration limits both the computational cost of optimization that increases with the number of parameters and the risk of overfitting.

The Morris method (Morris, 1991; Campolongo et al., 2007) was used for the sensitivity analysis, as it is relatively time efficient and enables ranking the parameters by importance. This qualitative method requires only a small number of simulations, $(p+1)n$, where p is the number of parameters and n is the number of random trajectories generated (here, $n = 10$).

We selected a set of parameters for the Morris sensitivity analyses based on previous sensitivity analyses conducted on soil parameters in ORCHIDEE (Dantec-Nédélec et al., 2017; Raoult et al., 2021; Mahmud et al., 2021). A distinction is made between the soil COS model parameters called first-order parameters (f_{CA}, α, and β for the mechanistic model and k_{soil} for the empirical model) and parameters called second-order parameters related to soil hydrology, carbon uptake and allocation, phenology, conductance, or photosynthesis (18 parameters; see Tables S3 and S4). The range of variation in the second-order parameters is described in previous studies using ORCHIDEE (Dantec-Nédélec et al., 2017; Raoult et al., 2021; Mahmud et al., 2021). For the first-order parameters, the range of variation is described in Yi et al. (2007) for k_{soil} (± 1.08 pmol COS per µmol CO_2) and in Table 1 in Meredith et al. (2019) for f_{CA}. The ranges of variation for α and β parameters are not directly given in the literature and were calculated based on information from the production parameters defined in Meredith et al. (2018) (Text S1 and Table S5).

3 Results

3.1 Site-scale COS fluxes

3.1.1 Soil COS flux seasonal cycles

Figure 2 shows the seasonal cycles of soil COS fluxes at the different sites where measurements were conducted. The empirical model mainly differs from the mechanistic model with a stronger seasonal amplitude of soil COS fluxes (34 % higher), except at the sites where a net COS production is found with the mechanistic model in summer (ES-LMA and IT-CRO). At all sites, the empirical model shows that the simulated uptake increases in spring, reaching a maximum in summer, and decreases in autumn with a minimal uptake during winter. The strong COS uptake in summer from the empirical model can be explained by the proportionality of soil COS uptake to simulated soil respiration, which increases with the high temperatures in summer. In contrast, the mechanistic model depicts almost no seasonality at all the sites where no net COS production is found over the year. As the mechanistic model represents both soil COS uptake and production, the increase in COS production due to higher temperature in summer compensates part of the COS uptake (Fig. C1 in Appendix C). While the uptake from the empirical model is often higher than the one computed with the mechanistic model in summer, soil COS uptake in winter is stronger with the mechanistic representation.

The scarcity of field measurements at AT-NEU, ES-LMA, IT-CRO, DK-SOR, and ET-JA does not allow for an evaluation of the simulated seasonality of COS fluxes. However, at US-HA, the absence of seasonality from May to October in the observations is also found in the mechanistic model, while a maximum net soil COS uptake is reached with the empirical model.

We found that the mechanistic model is in better agreement with the observations for four (IT-CRO, ET-JA, FI-HYY, and US-HA) out of the seven sites, with a mean

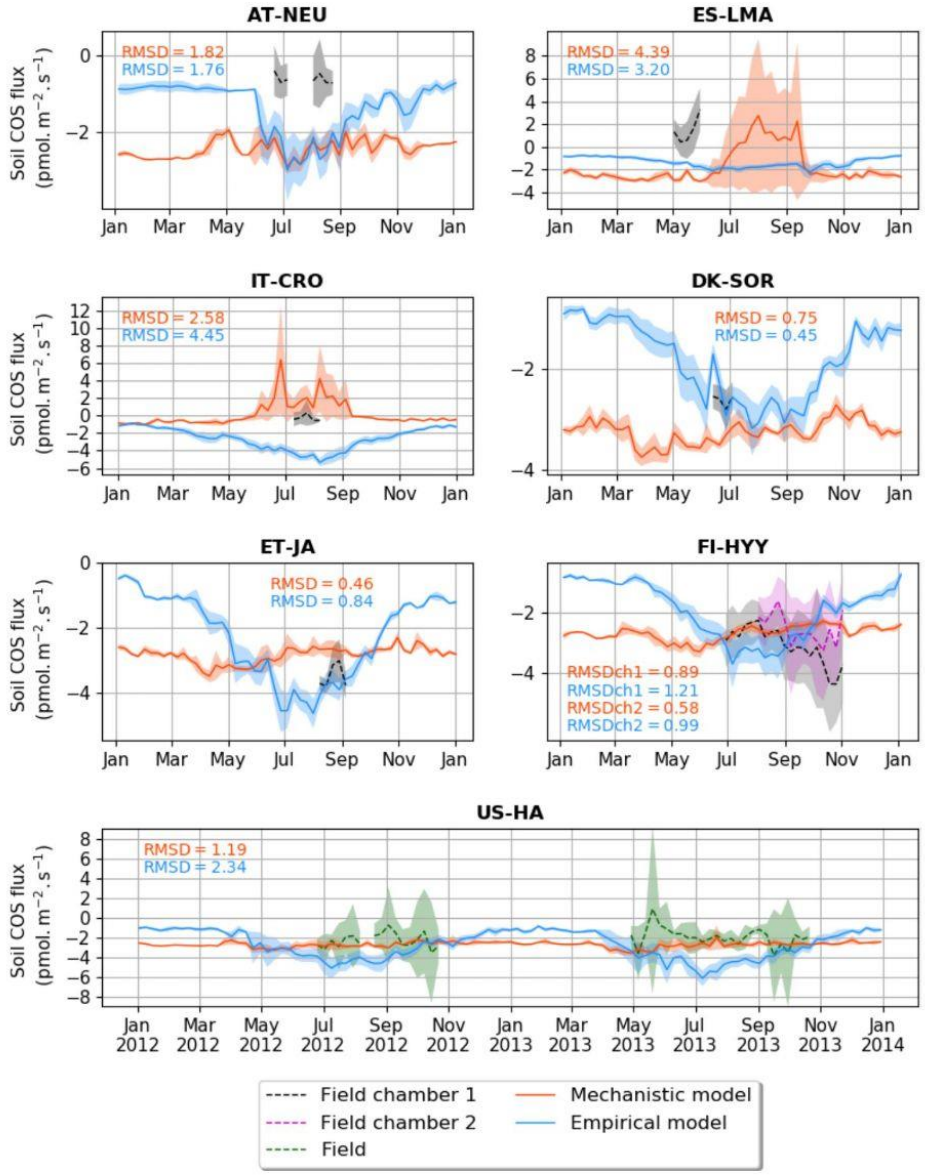

Figure 2. Seasonal cycle of weekly average net soil COS fluxes (pmol m^{-2} s^{-1}) at AT-NEU, ES-LMA, IT-CRO, DK-SOR, ET-JA, FI-HYY, and US-HA. The shaded areas around the observation and simulation curves represent the standard deviation over a week for each site. Soil COS fluxes are computed with a variable atmospheric COS concentration. RMSD values between the simulated and observed fluxes are given with the respective model colour at each site and for both soil chambers at FI-HYY (ch1 and ch2).

of 1.58 pmol m^{-2} s^{-1} and 2.03 m^{-2} s^{-1} for the mechanistic and empirical model, respectively. However, the mechanistic model struggles to reproduce soil COS fluxes at AT-NEU and ES-LMA, with an overestimation of soil COS uptake or an underestimation of soil COS production at AT-NEU and a delay in the simulated net COS production at ES-LMA. We might suspect that the removal of vegetation at these sites prior to the measurements could have artificially enhanced COS production in the observations. Indeed, the removal of vegetation could change soil structure and increase the availability of soil organic matter to degradation (Whelan et al., 2016). AT-NEU and ES-LMA are grassland sites for which soils are expected to receive higher light intensity than forest soils. These sites also show a high mean soil temperature of about 20 °C during the measurement periods. Therefore, high soil temperature and light intensity on soil surface could have enhanced soil COS production, as it was related to thermal or photo degradation of soil organic matter (Kitz et al., 2017, 2020; Whelan and Rhew, 2015; Whelan et al., 2016, 2018). This is not the case at FI-HYY, ET-JA, or DK-SOR, where soil temperature is much lower (mean value about 10 °C at FI-HYY and 15 °C at ET-JA and DK-SOR during the measurement periods) and the forested cover decreases the radiation level reaching the soil. Note that herba-

ceous biomass is also likely to be higher in grasslands than in forests. Besides, AT-NEU and ES-LMA are managed grassland sites with nitrogen inputs. Then, soil COS production could also be enhanced by a high nitrogen content as suggested by several studies (Kaisermann et al., 2018; Kitz et al., 2020; Spielmann et al., 2020), which is not represented in our models. The mechanistic model is able to represent a net COS production at IT-CRO but overestimates it. This might highlight the importance of adapting the production parameters (α and β) in this model to adequately represent net COS production. In this model, the net soil COS production is related to an increase in soil temperature. However, it is to be noted that IT-CRO is an agricultural site with nitrogen fertilization. Therefore, soil COS production in the observations could also be enhanced by nitrogen inputs. As expected, the empirical model is unable to correctly simulate the direction of the observed positive soil COS exchange rates at ES-LMA and IT-CRO.

3.1.2 Soil COS flux diel cycles

Figure 3 shows the comparison between the simulated and observed mean diel cycles over a month. The observations show a minimum net soil COS uptake or a maximum net soil COS production reached between 11:00 and 13:00 at AT-NEU (UTC+2), ES-LMA (UTC+2), IT-CRO (UTC+1), and DK-SOR (UTC+2). At AT-NEU and ES-LMA, neither model is able to represent the observed diel cycle. At these grassland sites, Spielmann et al. (2020) and Kitz et al. (2020) found that the daytime net COS emissions were mainly related to high radiations reaching the soil surface, the impact of which is not represented in the soil COS models. At IT-CRO and DK-SOR, the diel cycles simulated by the mechanistic model show patterns similar to the observations with a peak in the middle of the day but with an overestimation of the net soil COS production and a delay in the peak at IT-CRO and an overestimation of the net soil COS uptake at DK-SOR. The mechanistic model reproduces the absence of a diel cycle observed at FI-HYY and ET-JA but with an underestimation of the net soil COS uptake at ET-JA. AT US-HA, the observed soil COS flux does not exhibit diel variations, while the mechanistic model shows a peak with a decrease in the net soil COS uptake around 15:00. Wehr et al. (2017) explain this absence of the diel cycle in the observations by a range of variations for soil temperature and soil water content that is too low to influence soil COS flux. In ORCHIDEE, the simulated range of temperature at US-HA is larger than the one measured on site, and temperature is the main driver of the decrease in net soil COS uptake at this site (not shown). Therefore, the enhancement of soil COS production by soil temperature could be only found in the simulated flux. Another possibility is that it could be totally compensated by soil COS uptake in the observations. The mismatch between the model and the observations could be due to several factors including (i) an insufficient representation of the vegetation complexity by the division in PFTs; (ii) a poor calibration of the PFT-specific parameters (f_{CA}, α, and β); or (iii) missing processes in the model, such as considering the effect of nitrogen content on soil COS fluxes.

The empirical model shows a maximum soil COS uptake around 15:00 at ET-JA, FI-HYY, US-HA, and IT-CRO, which is not found in the observations at FI-HYY and is in contradiction with the observed diel variations at IT-CRO and ES-LMA. Considering all sites, the mechanistic model leads to a smaller error between the simulations and the observations, with a mean RMSD of 1.38 pmol m^2 s^{-1} against 1.87 pmol m^2 s^{-1} for the empirical model.

3.1.3 Dependency on environmental variables

Figure 4 represents simulated net soil COS fluxes versus soil temperature and soil water content at the different sites. At the sites where only a net soil COS uptake is simulated by the mechanistic model (all sites except IT-CRO and ES-LMA), soil COS uptake generally decreases with increasing soil water content, which appears to be the main driver of soil COS fluxes. This behaviour can be explained by a decrease in COS diffusivity through the soil matrix with increasing soil moisture, reducing soil COS availability for microorganism consumption. Furthermore, an optimum soil water content for net soil COS uptake is found between 10 % and 15 %, which was also observed in Ogée et al. (2016) and in several field studies to be around 12 % (Kesselmeier et al., 1999; Liu et al., 2010; van Diest and Kesselmeier, 2008). This optimum soil water content for soil COS uptake is related to a site-specific temperature optimum, which is found between 13 and 15 °C at US-HA for example. Indeed, Ogée et al. (2016) also describe a temperature optimum with a value that depends on the studied site (Kesselmeier et al., 1999; Liu et al., 2010; van Diest and Kesselmeier, 2008). At IT-CRO and ES-LMA, where a strong net soil COS production is simulated by the mechanistic model, the main driver of soil COS fluxes becomes soil temperature. At these sites, the net soil COS production increases with soil temperature, due to the exponential response of soil COS production term to soil temperature. The increase in soil COS production with soil temperature at IT-CRO and ES-LMA is supported by the observations (Fig. S1 in the Supplement).

Contrary to the mechanistic model, soil COS uptake computed with the empirical model is mainly driven by soil temperature, with a soil COS uptake that increases with increasing soil temperature. This response of the empirical model to soil temperature is due to its relation to soil respiration, which is enhanced by strong soil temperature. However, this net increase in soil COS uptake with soil temperature at all sites is not found in the observations (Fig. S1). It can be noted that low soil moisture values were found to limit soil COS uptake for the empirical model, as seen at ES-LMA for a soil water content below 8 %.

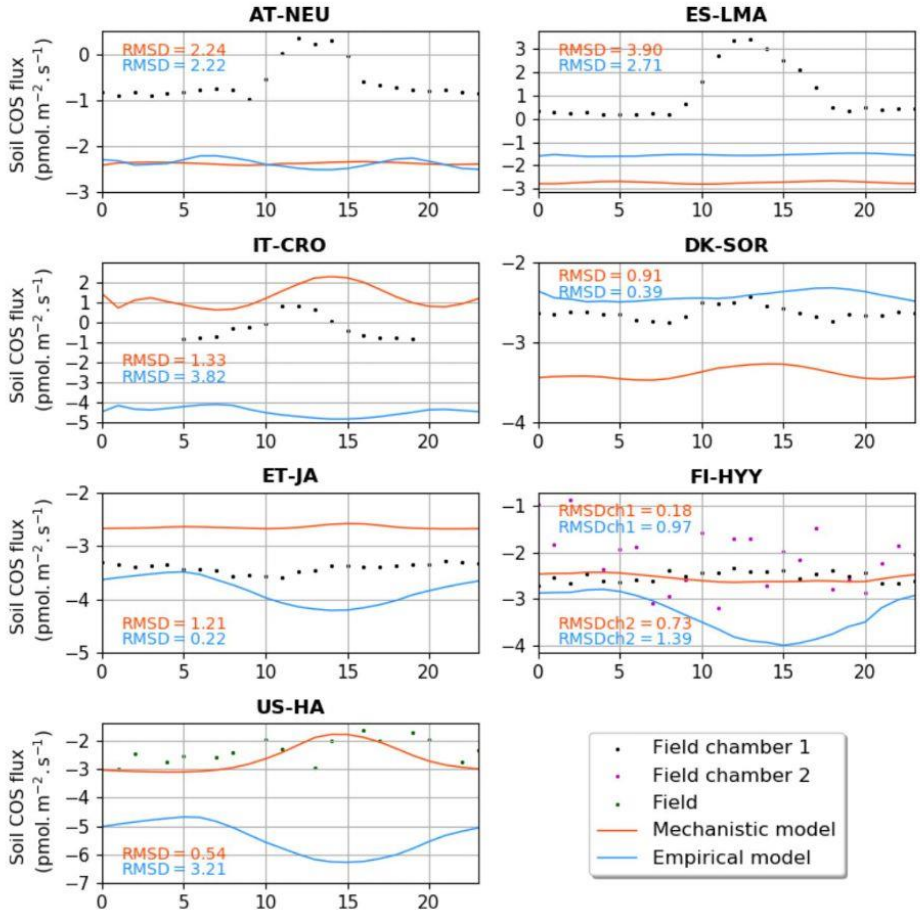

Figure 3. Mean diel cycle of net soil COS fluxes (pmol m^{-2} s^{-1}) over a month at AT-NEU (August 2015), ES-LMA (May 2016), IT-CRO (July 2017), DK-SOR (June 2016), ET-JA (August 2016), FI-HYY (August 2015), and US-HA (July 2012). Soil COS fluxes are computed with a variable atmospheric COS concentration. The observation-based diel cycles (dots) are computed using random forest models at AT-NEU, ES-LMA, IT-CRO, DK-SOR and ET-JA. At AT-NEU and ES-LMA, RMSD values between the simulated and observed fluxes are given with the respective model colour at each site and for both soil chambers at FI-HYY (ch1 and ch2).

3.1.4 Sensitivity analyses of soil COS fluxes to parameterization

Sensitivity analyses including a set of parameters (19 for the empirical model and 21 for the mechanistic model) were performed to evaluate the sensitivity of soil COS fluxes to each of the selected parameter. The Morris scores were normalized by the highest values to help rank the parameters by their relative influence on soil COS fluxes, where a score of 1 represents the most important parameter and that of 0 represents the parameters that have no influence on soil COS fluxes. For reasons of clarity, in the following we present the results only for the parameters that were found to have an impact on soil COS fluxes (Morris scores not equal to 0).

Figure 5 shows the results of the Morris sensitivity experiments highlighting the key parameters influencing soil COS fluxes from the empirical and the mechanistic models at FI-HYY and US-HA. For the empirical model at both sites, the first-order parameter (k_{soil}) is the most important parameter in the computation of soil COS fluxes, as it directly scales soil respiration to soil COS fluxes. The following parameters to which soil COS fluxes are the most sensitive are the scalar on the active soil C pool content (soilC) and the temperature-dependency factor for heterotrophic respiration (soil_Q_{10}). Indeed, the soilC parameter determines the soil carbon active pool content, which can be consumed by soil microorganisms during respiration, therefore impacting soil COS fluxes from the empirical model. The soil_Q_{10} parameter impacts soil COS fluxes at both sites, as it determines the response of soil heterotrophic respiration to temperature, which is included in the proportionality of soil COS fluxes to the total soil respiration in the empirical model. Similarly, one of the second-order parameters, the minimum soil wetness to limit the heterotrophic respiration (min_SWC_resp), has an impact on soil COS fluxes from the empirical model only. The importance of min_SWC_resp for soil COS fluxes is found at US-HA but not at FI-HYY. This can be explained by the difference in soil moisture between the two sites, with an annual

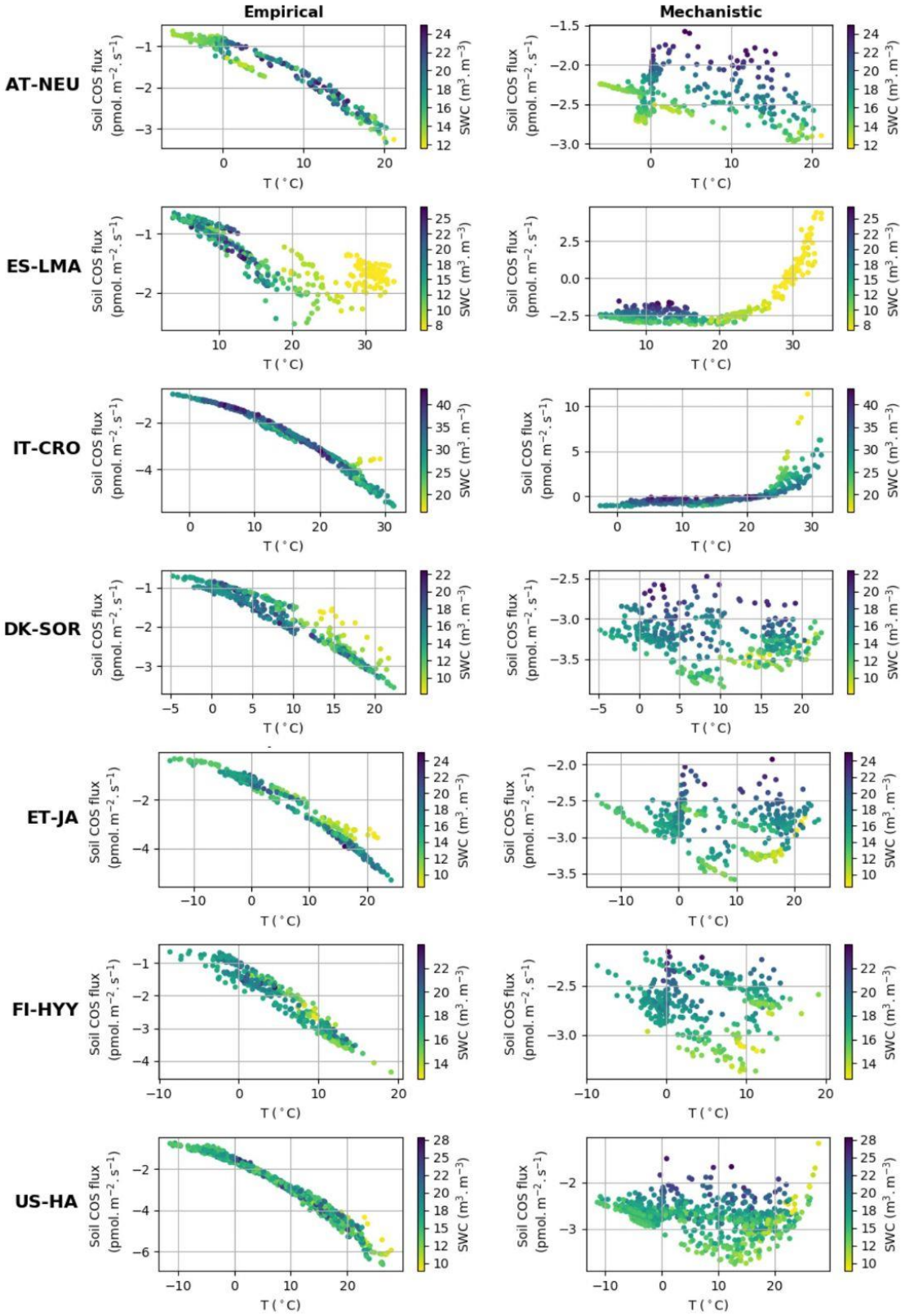

Figure 4. Simulated daily average net soil COS flux (pmol m^2 s^{-1}) versus soil temperature (°C) and soil water content (SWC) (m^3 m^{-3}) at AT-NEU, ES-LMA, IT-CRO, DK-SOR, ET-JA, US-HA, and FI-HYY, for the empirical and the mechanistic model.

mean of 16.2 % at US-HA and reaching a minimum of only 8.8 % against an annual mean of 17.5 % with a minimum of 12.4 % at FI-HYY.

Contrary to the empirical model, soil COS fluxes computed with the mechanistic model are more sensitive to two second-order parameters, the van Genuchten water retention curve coefficient n and the saturated volumetric water content (θSAT). These two second-order parameters are strongly linked to soil hydrology and determine the soil water content, which affects COS diffusion through the soil matrix and its uptake. The van Genuchten coefficients occur in the relationships linking hydraulic conductivity and diffusivity to soil water content (van Genuchten, 1980). At both sites, the strong impact of the van Genuchten water retention curve coefficient n on soil COS fluxes simulated with the mechanistic model highlights the critical importance of soil architecture. Thus, soil COS fluxes computed with the mechanistic model are expected to strongly vary according to the different soil types. Then, the first-order parameters (f_{CA}, α, and β) also influence soil COS fluxes from the mechanistic model. However, the uptake parameter (f_{CA} of PFT 15, boreal C_3 grass) has the most influence on soil COS fluxes at FI-HYY, while it is the production-related parameter (α of PFT 6, temperate broadleaved summergreen forest) that has the largest impact at US-HA. The stronger influence of the production parameter involved in the temperature response at US-HA might be explained by the difference in temperature between the two sites, which ranges from -10 to 25 °C at US-HA with an annual mean of 7.5 °C in 2013, while only ranging from -5 to 15 °C with an annual mean of 4.3 °C at FI-HYY in 2015. Similar to the difference in the main driver of soil COS fluxes found in Fig. 4, the most important first-order parameters to which soil COS fluxes are sensitive seem to differ between uptake and production parameters depending on the site conditions. It is to be noted that at US-HA, the most important production parameters are the ones of the dominant PFT at this site (PFT 6), which also correspond to a stronger response of the production term to temperature than for PFT 10 (temperate C_3 grass). However, at FI-HYY the most influential uptake parameter is for PFT 15 (boreal C_3 grass) that only represents 20 % of the PFTs at this site, while PFT 7 (boreal needleleaf evergreen forest) is the dominant PFT. This can be explained by the range of variation that is assigned to f_{CA} of PFT 7 by Meredith et al. (2019), which is larger than the one of f_{CA} for PFT 15 (9000 against 3100).

Finally, a set of parameters related to photosynthesis, conductance, phenology, hydrology, and carbon uptake has an impact on soil COS fluxes computed with both the empirical and the mechanistic models at the two sites. The specific leaf area (SLA), maximum rate of Rubisco activity-limited carboxylation at 25 °C (V_{cmax25}), residual stomatal conductance (g_0), and minimum photosynthesis temperature (T_{min}) have an impact on soil COS fluxes, as they also indirectly affect soil moisture through their influence on transpiration and stomatal opening. The second-order parameters related to soil hydrology (a, K_s, Z_{root}, θ_{WP}, θ_{FC}, θ_R, and θ_{Transp_max}) impact the soil water availability, which affects soil respiration for the empirical model and soil COS diffusion and uptake in the mechanistic model. For example, the parameter for the root profile (Z_{root}) determines the density and depth of the roots and therefore how much water can be taken up by roots.

3.1.5 Soil COS flux optimization

Figure 6 presents soil COS fluxes before and after optimization of the model parameters to better fit the observations at FI-HYY and US-HA. For the mechanistic model, the optimization at the two sites mainly changes the mean value of soil COS fluxes, by reducing the net uptake at US-HA and increasing it at FI-HYY. Similar to the mechanistic model optimization, the posterior soil COS uptake computed with the empirical model is enhanced at FI-HYY and reduced at US-HA. However, at US-HA, the increase in soil COS uptake is only found between April and October, while the winter soil COS fluxes are not impacted by the optimization. Using the optimized parameterization improves the RMSD by 7 % and 5 % at US-HA and by 23 % and 25 % at FI-HYY for the mechanistic and the empirical model, respectively. While it leads to similar posterior RMSD values between the two models at US-HA, the optimization of the mechanistic model gives a lower RMSD than the empirical model at FI-HYY, with 0.54 pmol m^{-2} s^{-1} against 0.95 pmol m^{-2} s^{-1}.

At FI-HYY, the difference between prior and posterior soil COS fluxes from the empirical model seems to mainly come from the change in the soil_Q_{10} value (Fig. E1 in Appendix E). The soil_Q_{10} value drops from 0.83 to 0.53, which corresponds to a prior Q_{10} value of 2.29 versus a posterior value of 1.70, decreasing the heterotrophic respiration response to soil temperature. Soil COS fluxes computed with the empirical model were found to be strongly sensitive to soil_Q_{10} (Fig. 5). The posterior value of this parameter has nearly attained the lower bound of its variation range. Since the range of variation represents the realistic values this parameter can take, we need to be careful about the fact that this parameter is trying to take values close to, or potentially beyond, these meaningful values. Furthermore, the optimization deviates the Q_{10} value at FI-HYY from the ones calculated in the observations over the measurement period (3.0 for soil chamber 1 and 2.5 for soil chamber 2). We could assume that k_{soil} should be defined as temperature dependent for linking soil COS flux to soil respiration (Berkelhammer et al., 2014; Sun et al., 2018), instead of being considered a constant. Thus, the optimization of the empirical model could in fact be aliasing the error of k_{soil} onto soil_Q_{10} because of the impossibility to account for the temperature dependence of soil COS to the CO_2 uptake ratio (Sun et al., 2018). At US-HA, the optimization also leads to a decrease in soil_Q_{10} but to a lesser extent, with the parameter remaining comfortably within its range of variation.

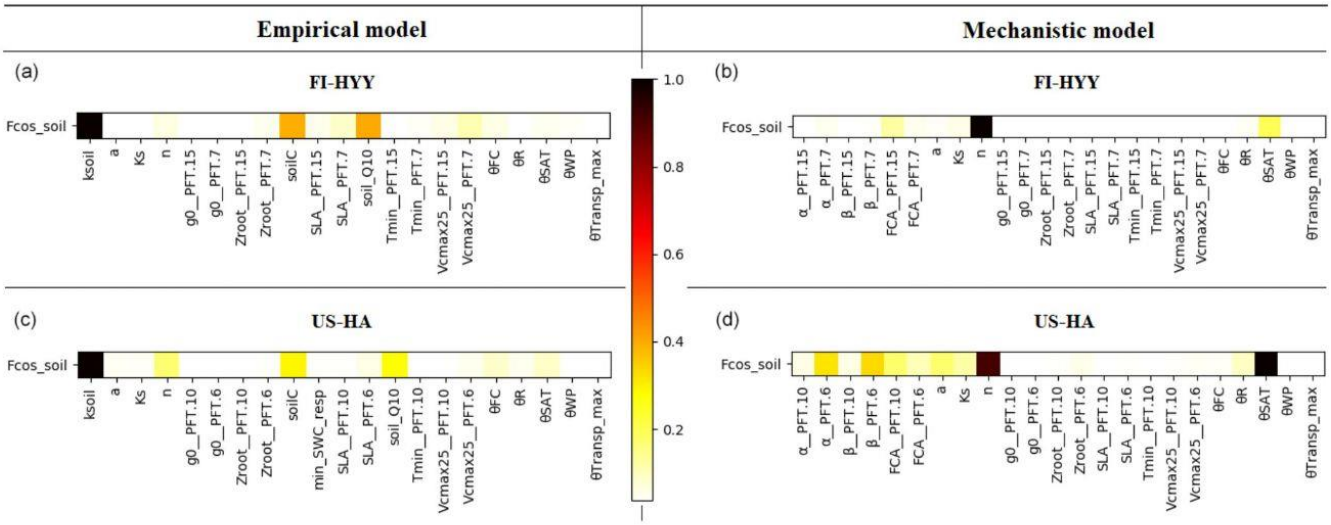

Figure 5. Morris sensitivity scores of the key parameters to which soil COS fluxes are sensitive, for the empirical **(a, c)** and the mechanistic **(b, d)** models. The two studied sites are FI-HYY **(a, b)** and US-HA **(c, d)**. Full descriptions of each tested parameter can be found in Tables S3 and S4 in the Supplement. The PFT is indicated at the end of the parameter names for the PFT-dependent parameters (at FI-HYY, PFT 7 is boreal needleleaf evergreen, and PFT 15 is boreal natural C_3 grassland; at US-HA, PFT 6 is temperate broadleaf summergreen, and PFT 10 is temperate natural C_3 grassland).

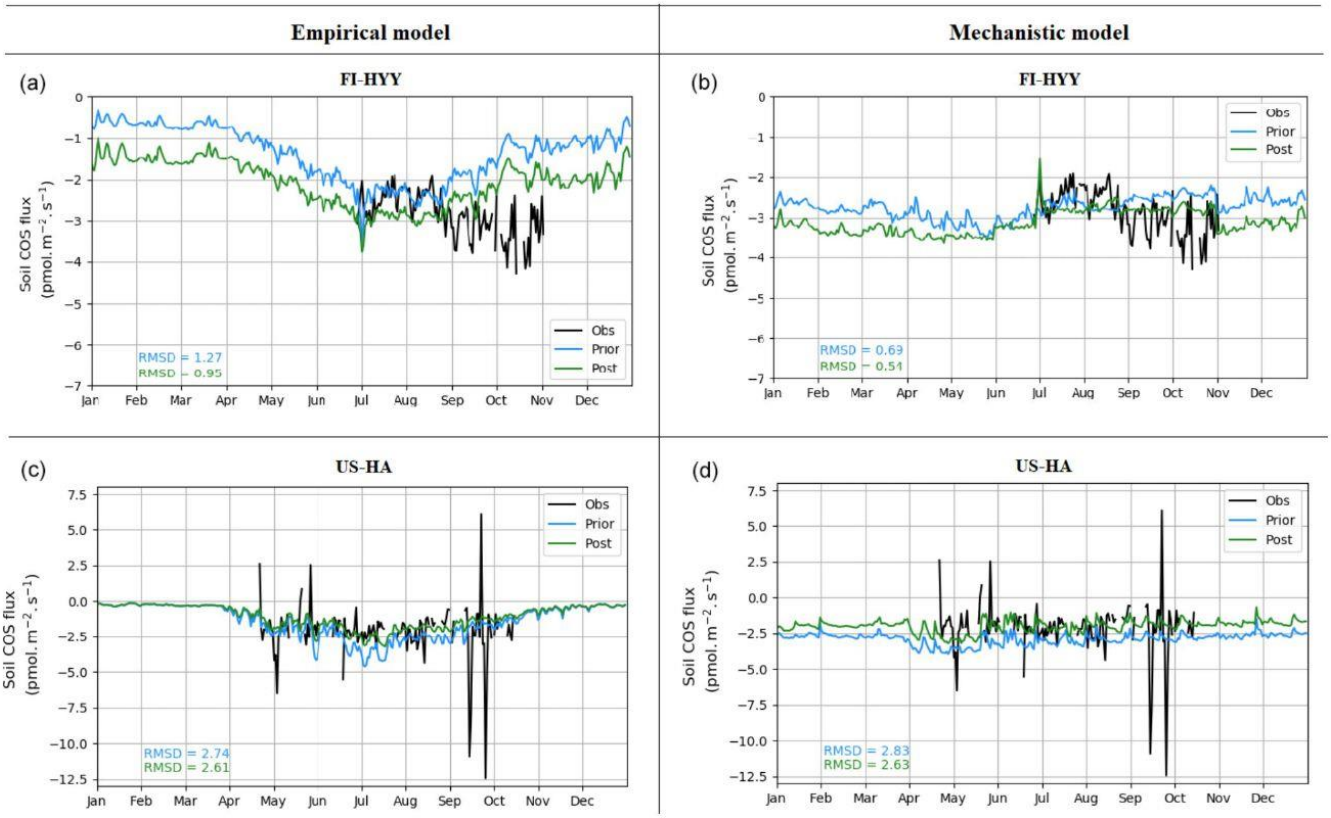

Figure 6. Prior and post-optimization net soil COS fluxes (pmol m^{-2} s^{-1}) for the empirical **(a, c)** and the mechanistic **(b, d)** models. The two studied sites are FI-HYY **(a, b)** in 2015 and US-HA **(c, d)** in 2013.

For the mechanistic model, the optimization reduces the enhancement factor value (f_{CA}) for PFT 10 at US-HA and increases the value of the production parameter α for the dominant PFT (PFT 6). This enhances the reduction in net soil COS uptake, which was slightly overestimated with the prior model parametrization. At FI-HYY, the optimized parameters show higher values of f_{CA} and of α for PFT 15 and of both production parameters (α and β) for the dominant PFT (PFT 7). This increase in both soil COS uptake and production after optimization could correspond to an attempt to better simulate the larger range of variation found in the observations compared to the modelled fluxes.

Finally, the optimization also affects hydrology-related parameters for both models. However, while it improves the simulated water content compared to the observations for the mechanistic model at the two sites (RMSD decreases by 28 % at FI-HYY and 22 % at US-HA), it leads to a degradation at FI-HYY for the empirical model (RMSD increases by more than 3 times). Since the empirical model is quite a simplistic model with few parameters, it relies on parameters from different processes to help better fit the observations – sometimes degrading the fit to the other processes. The mechanistic model is able to both improve the fit to the COS observations and soil moisture values, implying its parameterization is more consistent.

This optimization experiment has been promising, highlighting how observations can be used to improve the models. However, since we only optimized over two sites due to the scarcity of soil COS flux observations, for the global-scale simulations in the rest of this study, we will rely on the default parameter values of each parameterization.

3.2 Global-scale COS fluxes

3.2.1 Soil COS fluxes

The spatial distribution of oxic soil COS fluxes shows a net soil COS uptake everywhere except in India, in the Sahel region, and in some areas in the tropical zone, where net soil COS production is simulated (Fig. 7a). The strongest uptake rates are found in western North America and South America, as well as in China, with a mean maximum uptake of -4.4 pmol COS m^{-2} s^{-1} over 2010–2019. The difference in magnitude between the maximum uptake value and the maximum of production can be noticed, with a net production reaching 67.2 pmol COS m^{-2} s^{-1} in the Sahel region. India and the Sahel region, where oxic soil COS production is concentrated, are represented in ORCHIDEE by a high fraction of C_3 and C_4 crops (Fig. S4). In the mechanistic model, crops are associated with the lowest f_{CA} value due to overall lower fungal diversity and abundance in agricultural fields (Meredith et al., 2019) and the strongest response of oxic soil COS production to temperature as observed by Whelan et al. (2016). Thus, these PFT-specific parameters combined with high temperature in the tropical region can explain the net oxic soil COS production found in these regions. C_3 crops are also dominant in China near the Yellow Sea (Fig. S4). However, the mean soil temperature in this region is about 15 °C lower than the mean soil temperature in India, leading to a lower enhancement of soil COS production. The highest atmospheric COS concentration is also found in this region with about 800 ppt (Fig. S3). Indeed, recent inventories have shown that China was related to strong anthropogenic COS emissions due to industry, biomass burning, coal combustion, agriculture, or vehicle exhaust (Yan et al., 2019; Zumkehr et al., 2018). High atmospheric COS concentrations increase soil COS diffusion and uptake that can compensate part of soil COS production. The highest values of soil COS fluxes for anoxic soils are located in northern India, with a mean maximum value reaching 36.8 pmol COS m^{-2} s^{-1} (Fig. 7b). This region is characterized by rice paddies, which were also associated with strong COS production in previous studies (Zhang et al., 2004).

The total soil COS fluxes (oxic and anoxic) computed with the mechanistic model (Fig. 7c) show a very different spatial distribution than the one obtained with the empirical model (Fig. 7d). Soil COS fluxes from the empirical model are on the same order of magnitude for net COS uptake than the mechanistic model, with a mean maximum uptake of -6.41 pmol COS m^{-2} s^{-1}. However, most soil COS uptakes simulated by the empirical model is located in the tropical region, where soil respiration is strong due to high temperature. The distribution and magnitude of soil COS flux from the empirical approach is similar to the one presented in Kooijmans et al. (2021) (see Fig. S15 in the Supplement of Kooijmans et al., 2021), when implemented in SiB4. For the mechanistic model, the comparison of oxic soil COS flux distribution with the one in SiB4 shows a net soil COS emission in India in both SiB4 and ORCHIDEE. However, the maximum oxic soil COS flux is about 60 pmol m^{-2} s^{-1} higher in ORCHIDEE than in SiB4. The regions with the strongest net oxic soil COS uptake also differ between SiB4 and ORCHIDEE, as it is concentrated in the tropics in SiB4, as well as in western North America and South America, and in China for ORCHIDEE.

The difference in soil COS fluxes between the mechanistic model and the empirical model ranges from -4.1 pmol COS m^{-2} s^{-1} to $+68.0$ pmol COS m^{-2} s^{-1} (Fig. D1 in Appendix D). Over western North America and South America; northern and southern Africa; western Asia; and eastern, northern, and central Asia, the net COS uptake from the mechanistic model exceeds the uptake from the empirical model. On the contrary, soil COS uptake from the empirical approach is higher than the net COS uptake simulated with the mechanistic model over eastern North America and South America; western, central, and eastern Africa; and Indonesia. The absence of soil COS production representation in the empirical approach leads to the strongest differences in India and in the Sahel region, reaching $+68.0$ pmol COS m^{-2} s^{-1}.

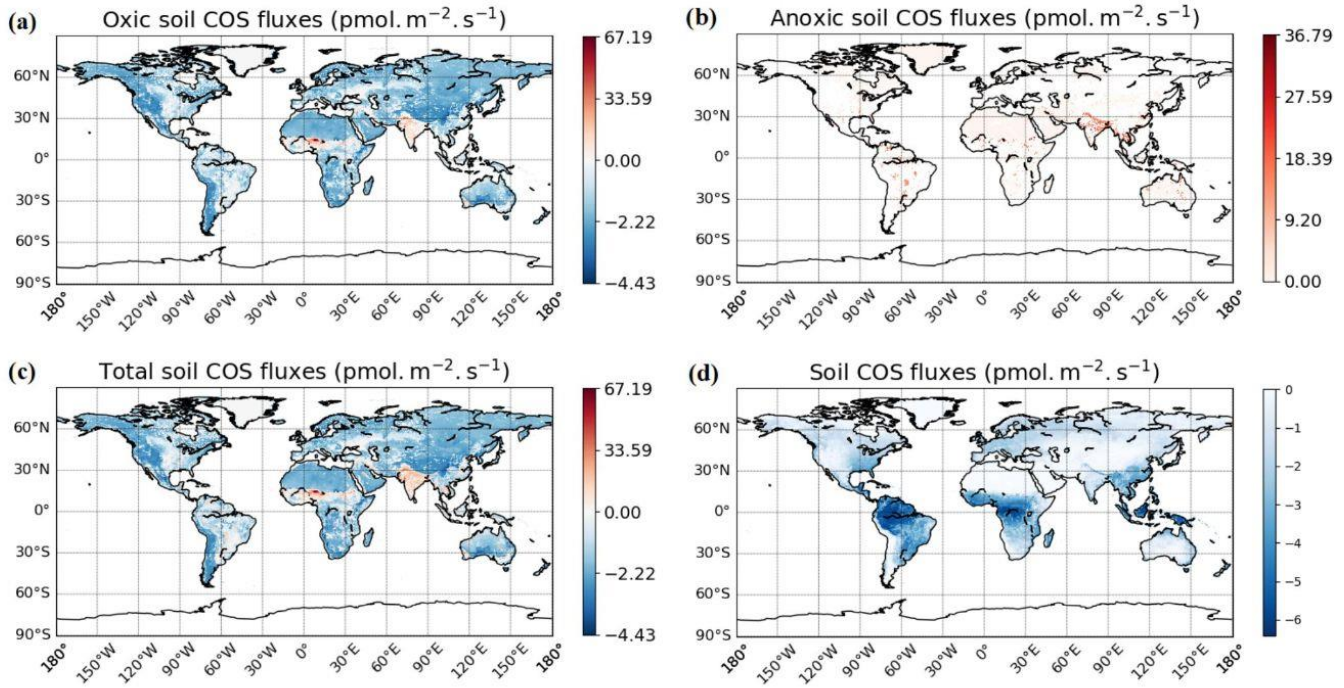

Figure 7. Maps of mean soil COS fluxes for the mechanistic **(a, b, c)** and the empirical model **(d)**, computed over 2010–2019 with a variable atmospheric COS concentration. Colour scales were normalized between the minimum and maximum soil COS flux values and centred on zero for oxic and total soil COS fluxes computed with the mechanistic model. The map resolution is $0.5° \times 0.5°$.

3.2.2 Temporal evolution of the soil COS budget

We computed the mean annual soil COS budget over the period 2010–2019 using the monthly variable atmospheric COS concentration, and we compared its evolution to the variations in the mean annual atmospheric COS concentration.

The evolution of the mean annual soil COS budget (Fig. 8) shows small variations in the budget for oxic soils computed with the mechanistic model between 2010 and 2015, with a net sink ranging from -133 to $-124\,\mathrm{GgS\,yr^{-1}}$. Then, from 2016 we see a sharp decrease in this budget, which reaches $-98\,\mathrm{GgS\,yr^{-1}}$ in 2019. This decrease also corresponds to the decrease in atmospheric COS concentration observed between 2016 and 2019 with a loss of 25 ppt in 3 years. Several monitoring stations recorded a drop in atmospheric COS concentration over Europe, as for the Gif-sur-Yvette station with -42 ppt between 2015 and 2021 (updated after Belviso et al., 2020). Note that the decrease in the oxic soil COS budget computed with the mechanistic model is sharper than the drop in atmospheric COS concentration because changes in oxic soil COS budget result from the combined effect of decreasing atmospheric COS concentration and changes in the drivers of soil COS fluxes (i.e. changes in soil temperature and water content during the 10-year period which are not homogenously distributed around the globe; not shown). On the contrary, the soil COS net uptake computed with the empirical model slightly increases from $-212\,\mathrm{GgS\,yr^{-1}}$ in 2010 to $-219\,\mathrm{GgS\,yr^{-1}}$ in 2019. As the empirical model defines soil COS flux as proportional to the total soil respiration independently of atmospheric COS concentration, the budget obtained with this model is not impacted by the variations observed in atmospheric COS concentration. The anoxic soil COS budget follows soil temperature variations (not shown), with an increasing trend of about $0.17\,\mathrm{GgS\,yr^{-1}}$ over the studied period.

3.3 Transport and site-scale concentrations

3.3.1 Interhemispheric gradient

We transported total COS fluxes for the different configurations (i.e. including not only the soil fluxes but also other components of the COS atmospheric budget, listed in Table 2) with the LMDZ6 atmospheric transport model as described in Sect. 2.1.3. We analysed COS concentrations derived from simulated COS fluxes obtained with the mechanistic and two empirical approaches with regards to the COS concentrations observed at 14 NOAA sites depicted in Fig. B1 in Appendix B. Note that atmospheric mixing ratios of COS result from the transport of all COS sources and sinks and that, due to other sources of errors (transport and errors in the other COS fluxes), the comparison presented in the following should be taken as a sensitivity study of COS seasonal cycle and interhemispheric gradient to the soil exchange fluxes rather than a complete val-

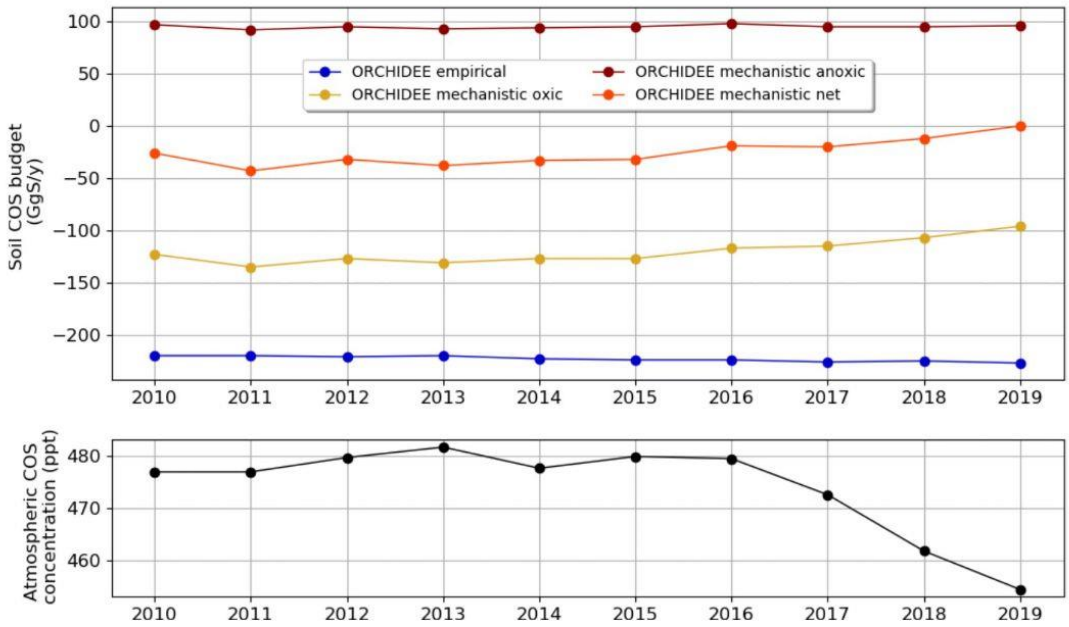

Figure 8. Evolution of mean annual soil COS budget and mean annual atmospheric COS concentration between 2010 and 2019, computed with a variable atmospheric COS concentration.

idation of one approach or the other. Figure 9 shows the COS atmospheric concentrations at NOAA sites as a function of latitude for each simulated soil flux and for the observations. Here as we want to focus on the latitudinal variations in atmospheric COS mixing ratios; the atmospheric COS concentrations have been vertically shifted to have the same mean as the observations. This means that the concentrations values cannot be compared at each site; we can only compare the interhemispheric gradients of simulated and observed concentrations. The RMSD for the mechanistic model with oxic soils only, the mechanistic model with oxic and anoxic soils, the empirical Berry model (with oxic soils only), and the empirical Launois model (with oxic and anoxic soils) are 36.5, 39.4, 43.0, and 51.0 ppt, respectively. While the different approaches show similar gradient patterns in the southern latitudes, they lead to strong differences in the simulated concentrations in the Northern Hemisphere. Compared to empirical approaches, the mechanistic approach marginally improves the latitudinal distribution of the atmospheric mixing ratios by decreasing the concentrations in the high latitudes. The lower atmospheric mixing ratios above 60° N reflect the stronger soil absorption in the mechanistic model (see Fig. 9), where soil COS uptake is dominant and the compensation by COS production is small (Fig. D2 in Appendix D). Despite this slight improvement, there are persistent biases as overestimated concentrations at the high-latitude sites ALT, BRW, and SUM and underestimated concentrations at most tropical sites, i.e. WIS, MLO, and SMO. These model–observation mismatches have led top-down studies to identify vegetation as an underestimated sink in the high latitudes (Ma et al., 2021; Remaud et al., 2022) and the tropical oceanic emissions as the missing source (Berry et al., 2013; Launois et al., 2015; Kuai et al., 2015; Ma et al., 2021; Remaud et al., 2022; Davidson et al., 2021). The present anoxic soil fluxes have little impact on the surface latitudinal distributions and therefore are unlikely to shed new light on the tropical missing source. An explanation for the small impact is that they are located outside areas experiencing deep convection events (e.g. the Indian monsoon domain), and thus the surface concentrations are less sensitive to these fluxes.

3.3.2 Seasonal cycle at NOAA sites

Figure 10 shows the detrended temporal evolution of COS concentrations for the mechanistic and empirical approaches at Alert (ALT, Canada) and Harvard Forest (HFM, USA). Because of the mean westerly flow, the HFM site is influenced by continental regions to the west (Sweeney et al., 2015) and is more sensitive to the soil fluxes than other mid-latitude sites located to the west of the ocean as MHD; see Fig. 1 in Remaud et al. (2022). The ALT site samples air masses come not only from high-latitude ecosystems (Peylin et al., 1999) but also from regions further south due to atmospheric transport (Parazoo et al., 2011). The reader is referred to Table B2 in Appendix B for the other sites. At both sites, the mechanistic approach tends to weaken the total seasonal amplitude and increase the model–data mismatch. At HFM, since the mechanistic soil model shows overall good agreement with the observed soil fluxes (e.g. Fig. 2), the model–observation mismatch likely arises from errors in other components of

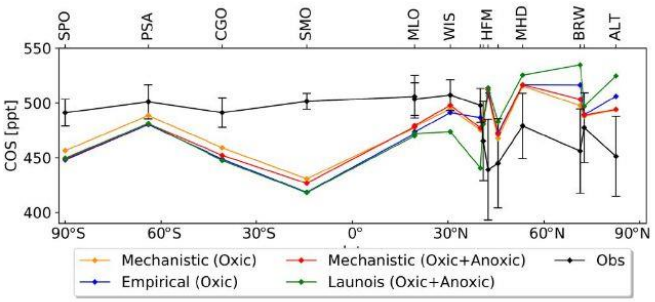

Figure 9. Comparison of the latitudinal variations in the COS abundances simulated by LMDZ at NOAA sites with the observations (black). The LMDZ COS abundances have been vertically shifted such that the means of the simulated concentrations are the same as the mean of the observations. The error bars around the black curve represent the standard deviation over the whole studied period at each NOAA site. The orange curve is obtained using the oxic soil fluxes of the mechanistic model. The red curve is obtained using the oxic and anoxic soil fluxes of the mechanistic model. The blue curve is given by LMDZ using the oxic soil fluxes from the Berry empirical model. The green curve is obtained using the soil fluxes from the empirical approach of Launois et al. (2015). For more clarity, the names of the stations KUM (19.74° N, 155.01° W), NWR (40.04° N, 105.54° W), LEF (45.95° N, 90.28° W), and SUM (72.6° N, 38.42° W) are not shown in this figure due to their proximity to other stations (Fig. B1 and Table B1 in Appendix B).

the COS budget (in particular oceanic and vegetation fluxes). Therefore, empirical approaches give a more realistic seasonality of atmospheric concentrations for the wrong reasons, which likely hides an underestimated vegetation uptake. Indeed, as Maignan et al. (2021) showed that the vegetation uptake magnitude in ORCHIDEE was consistent with measurements, the introduction of variable atmospheric COS concentrations decreased the vegetation uptake, which, as a result, is very likely underestimated now. Moreover, the comparison between simulated and observed concentrations shows a degradation of the simulated concentrations in this study compared to Maignan et al. (2021). It is to be noted that in addition to using a variable atmospheric COS concentration in this study, the transported ocean COS fluxes from Masotti et al. (2016) and Lennartz et al. (2017, 2021) differ from the ones used in Maignan et al. (2021), Kettle et al. (2002), and Launois et al. (2015). These results illustrate the necessity of well constraining both the soil and vegetation fluxes in order to optimize the GPP with the help of atmospheric inverse modelling.

4 Discussion

4.1 Soil budget

According to the mechanistic approach of this study, the COS budget for oxic soil is a net sink of $-126\,\mathrm{GgS\,yr^{-1}}$ over 2009–2016, which is close to the value of $-130\,\mathrm{GgS\,yr^{-1}}$ found by Kettle et al. (2002) (Table 3). This net COS uptake by oxic soils is higher than the one found in SiB4 by Kooijmans et al. (2021) with $-89\,\mathrm{GgS\,yr^{-1}}$, also based on the mechanistic model described in Ogée et al. (2016). In SiB4 and in ORCHIDEE, the mechanistic model gives the lowest oxic soil COS net uptake compared to all previous studies, which were using empirical approaches. This budget is also 41 % lower than the one found with the Berry empirical approach in this study, with an uptake of $-214\,\mathrm{GgS\,yr^{-1}}$. The anoxic soil COS budget computed with the mechanistic approach is $+96\,\mathrm{GgS\,yr^{-1}}$, which is close to the budget found by Launois et al. (2015) of $+101\,\mathrm{GgS\,yr^{-1}}$ based on methane emissions. However, while COS emissions from anoxic soils were only located in the northern latitudes in Launois et al. (2015), the COS production in this study is also distributed in the tropical region. Thus, we can expect that despite similar budget values for anoxic soils, the difference in flux distribution will impact the latitudinal gradient of COS fluxes. Finally, adding the anoxic soil COS budget to oxic soil COS budget results in a total soil COS budget of only $-30\,\mathrm{GgS\,yr^{-1}}$ for the mechanistic approach.

When computing the net total COS budget considering all sources and sinks of COS (Table 2), we found that neglecting the potential COS production of oxic soils and COS emissions from anoxic soils leads to an overestimation of COS sink or an underestimation of COS source to close the budget ($-165\,\mathrm{GgS\,yr^{-1}}$). On the contrary, the total COS budget computed with the mechanistic soil model is closed given the uncertainties on each component (Table 2). However, despite a closed budget, the mismatch between the observed and simulated latitudinal gradients of atmospheric COS concentration highlights errors in COS flux component distributions (Fig. 9).

It is also to be noted that the mechanistic model better simulates the lack of seasonality in the soil COS flux at US-HA compared to the empirical model (Fig. 2). US-HA is represented by 80 % of PFT 6 (temperate broadleaved summergreen forest), and the absence of seasonality by this PFT has also been reported at a mid-latitude site at Gif-sur-Yvette (Belviso et al., 2020). This PFT is largely found in the temperate region such as in Europe and in the southern United States. Moreover, NWR, HFM, and LEF stations are mainly influenced by COS exchanges from PFT 6. Therefore, the use of the mechanistic model would be recommended to carry out new comparisons at these mid-latitude sites.

4.2 Variable atmospheric COS concentration versus constant atmospheric COS concentration

We studied the impacts of using a constant versus a variable atmospheric COS concentration on soil COS fluxes. At the site scale we found a distinction between the sites where soil COS production is strong (IT-CRO and ES-LMA) and the sites mainly showing a net soil COS uptake. The impact of using a constant atmospheric COS concentration is lower at

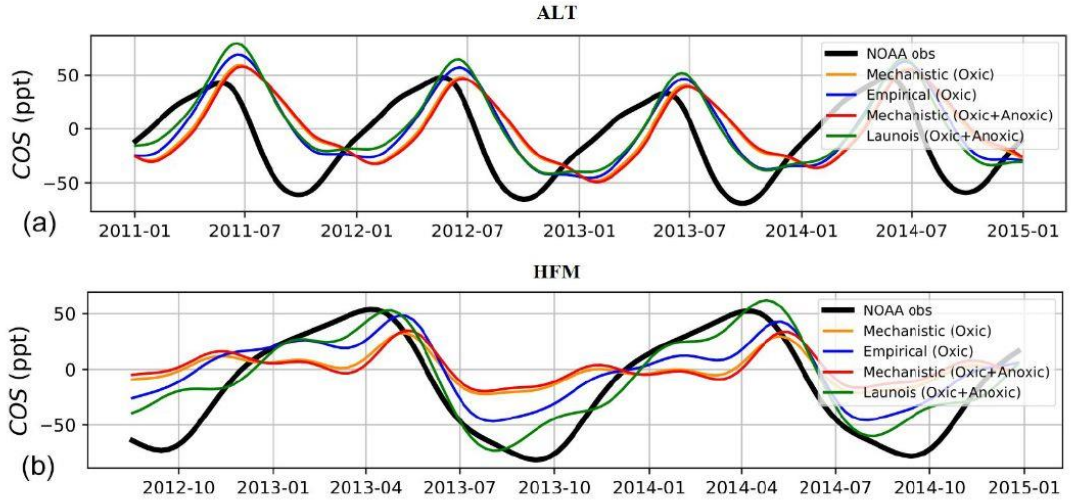

Figure 10. Detrended temporal evolution of simulated and observed COS concentrations at two selected sites, simulated with LMDZ6 transport between 2011 and 2015. The simulated concentrations are obtained by transporting the surface fluxes described in Table 2 and changing only the contribution from soils, with mechanistic (oxic soils alone and oxic + anoxic soils) and empirical approaches (Berry et al., 2013; Launois et al., 2015). **(a)** Alert station (ALT, Canada) and **(b)** Harvard Forest station (HFM, USA). The curves have been detrended beforehand and filtered to remove the synoptic variability (see Sect. 2.3.3). Please note that the date format in this figure is year-month.

IT-CRO and ES-LMA because the atmospheric COS concentration does not directly impact the soil COS production term but participates in the net soil COS flux. Our study shows that at the sites where a net soil COS uptake is dominant, using a constant atmospheric COS concentration leads to a lower soil COS flux in winter and a higher soil COS flux from spring to autumn (not shown). Indeed, during the growing season, plant uptake decreases atmospheric COS concentration (Fig. S2), which reduces COS availability for soil COS diffusion, whereas during winter, a higher atmospheric COS concentration enhances COS diffusion into the soil.

At the global scale, as the variable atmospheric COS concentration used in this study shows a decrease of about 25 ppt in the recent years (Fig. 8), considering a constant atmospheric COS concentration would not enable the representation of the impact of this strong variation on soil COS fluxes. When computing the soil COS budget over 2016 to 2019, we found a net uptake of $-126\,\mathrm{GgS\,yr^{-1}}$ with the mechanistic model using a constant atmospheric COS concentration against the $-110\,\mathrm{GgS\,yr^{-1}}$ computed with a monthly spatially variable concentration. Using a constant atmospheric COS concentration would then lead to a 13 % higher net soil COS uptake over the past 4 years.

We also studied the impact of considering a constant versus a variable atmospheric COS concentration on the seasonal variations in mean monthly soil COS fluxes over 2010–2019, simulated with the mechanistic model (not shown). We found that using a constant atmospheric COS concentration leads to an increase in net soil COS uptake over the whole year in the southern latitudes and from June to February in the northern latitudes (reaching $1.62\,\mathrm{pmol\,m^{-2}\,s^{-1}}$). This increase is higher over the regions with the lowest atmospheric COS concentrations, which limits COS diffusion through the soil matrix. On the contrary when atmospheric COS concentration is high in the northern latitudes between April and May, considering a constant atmospheric COS concentration decreases the net soil COS uptake. We notice that this lower net soil COS uptake with a constant atmospheric COS concentration can be found as early as March over Europe, where atmospheric COS concentration is higher in this region. In eastern Asia, where atmospheric COS concentration is higher than 800 ppt, the decrease in the net soil COS uptake can reach $-2.34\,\mathrm{pmol\,m^{-2}\,s^{-1}}$ when considering a constant atmospheric COS concentration.

It is to be noted that the modelled COS concentrations we used have their own uncertainty, which is however smaller than their difference with the fixed value (Remaud et al., 2022).

4.3 Foreseen improvements

The mechanistic representation of soil COS fluxes was found to be in better agreement with the observations at field sites. However, there can be strong differences between the simulated fluxes and the observations at some sites, especially at AT-NEU and ES-LMA. In the mechanistic approach, the influence of light on soil COS fluxes is not considered. Several field studies have reported light-induced emissions in oxic soils (Kitz et al., 2017; Meredith et al., 2018; Spielmann et al., 2019a; Whelan and Rhew, 2015), assumed to be related to the effect of light on soil organic matter. Spielmann et al. (2019a) related strong soil COS emissions during daytime to light at the sites where direct solar radiations reached the surface, such as ES-LMA and AT-NEU. At these

sites, the mechanistic model was unable to represent the soil COS emission peak during daytime. The optimization we performed showed that, as expected, adjusting the parameters to site observations improves the fit between the simulated and observed fluxed. However, it is necessary to represent all important processes in the mechanistic approach before calibrating the parameters. Thus, a next step in our modelling approach could be to include the light influence on soil COS fluxes, which can be of major importance for the sites where radiations strongly affect soil COS fluxes. Several studies also found that soil COS production could be related to nitrogen content, which increases with nitrogen fertilizer application (Kaisermann et al., 2018; Meredith et al., 2018, 2019). At the sites where soil is enriched with nitrogen inputs, such as agricultural fields or managed and fertilized grasslands and forests, the fertilization practices would also need to be included when representing the dynamics of soil COS fluxes. However, the soil nitrogen content and soil microbial nitrogen biomass vary not only with fertilization but also with location. Then, in addition to indications on land use, information on the total soil nitrogen content should be included in the model to consider nitrogen impact on soil COS flux. In the soil COS models, the impact of snow cover is also not represented. Indeed, due to the scarcity of soil COS flux observations in winter and with snow cover, its effect on soil COS flux could not be implemented in soil COS models yet. However, Helmig et al. (2009) found that COS uptake was not zero when soil is covered by snow at Niwot Ridge, Colorado.

Moreover, one difficulty with the study of soil COS fluxes arises from the scarcity of field measurements that could be used for model validation and calibration. Besides, the observation sites considered here are all located in a small latitudinal range between 39 and 62° N. Measurements in the tropics and in the Southern Hemisphere are needed. Especially, soil COS flux observations in northern India could help to validate the net soil COS production simulated in both SiB4 and ORCHIDEE. In the tropical rainforest, soil COS flux measurements were performed at La Selva Biological Station in Costa Rica (Sun et al., 2014). When available, these measurements could allow for a first comparison between the observed and simulated soil COS flux in a tropical region.

Then, the characterization of the soil microbial community should also be addressed to improve the scaling of CA content and activity, represented by the f_{CA} parameter (Meredith et al., 2019).

The implementation of the soil COS flux mechanistic model from Ogée et al. (2016) in SiB4 (Kooijmans et al., 2021) shows a seasonal cycle with a maximum net soil COS uptake in summer for the sites without crops, while the fluxes computed in ORCHIDEE show almost no seasonality. The expression of the production term P differs between the two models, which is based on Meredith et al. (2018) in SiB4 and on Whelan et al. (2016) in ORCHIDEE. The observation sites that are common to the two studies (FI-HYY, US-HA, AT-NEU, and DK-SOR) are also represented by different fractions of biomes between SiB4 and ORCHIDEE, which changes the parameterization to compute soil COS fluxes. Finally, the parameter values for the enhancement factor f_{CA} for grass differ as the value for tropical grass is also assigned to C_3 and C_4 grass in SiB4. Soil COS flux field data are mainly available in summer; therefore having field measurements over a whole year could better inform the seasonality of observed soil COS fluxes to compare to the simulations.

The optimization does not modify the respective seasonality of both soil COS models, with a seasonal cycle that agrees with the one of soil respiration for the empirical model and a lack of seasonality for the mechanistic model. The lack of observations in winter does not enable validating or constraining soil COS fluxes in winter. Therefore, having field observations over a whole year could help to determine if both models could be calibrated with a constraint over the whole year instead of only during summer and autumn. Moreover, the optimized set of parameters for the empirical models leads to a degradation of the simulated soil water content compared to the observations at FI-HYY, while the optimized parameters of the mechanistic model improve the representation of soil water content at US-HA and FI-HYY. Thus, the mechanistic approach is to be preferred over the empirical model and should be selected for future COS studies in ORCHIDEE.

The sensitivity analyses showed the importance of the hydrology-related parameters in the computation of soil COS fluxes with the mechanistic model. Thus, assuming an accurate representation of soil COS fluxes, soil COS fluxes could have the potential to add a new constraint on hydrology-related parameters.

In this work, soil COS fluxes are computed in the top 9 cm, which assumes that soil COS uptake and production depend on the conditions in the first soil layers. Indeed, soil COS uptake depends on diffusive supply of COS from the atmosphere. However, since soil COS production does not depend on COS supply, deeper soil layers could also contribute to soil COS production. A study by Yang et al. (2019) presents COS profile measurements in an orchard, which shows a non-zero COS concentration in deeper soil layers but no direct evidence for attributing it to soil COS production. Thus, we could consider deeper soil layers in the future to study the impact on soil COS fluxes compared to considering only the top soil layers.

The anoxic soil map of regularly flooded wetlands from Tootchi et al. (2019) enables approximating the spatial distribution of anoxic soil. However, in our approach, seasonality is only represented through soil temperature seasonality. Anoxic soil temporal dynamics were initially included in the model described by Ogée et al. (2016) with the soil redox potential but is not implemented in land surface models such as ORCHIDEE yet. We could also refine our approach by distinguishing between the different types of wetlands and define a P_{ref} value for each wetland type instead of a global

value of 10 pmol COS m^{-2} s^{-1}. Then, a distinction could also be made for anoxic soil COS fluxes from boreal peatlands, as Meredith et al. (2019) give a value of f_{CA} specific to this biome. Moreover, indirect COS emissions from DMS oxidation in anoxic soils have been reported (Kettle et al., 2002; Watts, 2000) but are not represented in this study. Finally, the anoxic map used here represents 9.7 % of the global land area, but the distribution of anoxic soils can greatly vary depending on the study (between 3 % and 21 %, Tootchi et al., 2019). Therefore, it would also be interesting to investigate the impact of anoxic soil coverage on soil COS flux uncertainty.

5 Conclusions and outlooks

We have implemented in the ORCHIDEE LSM a mechanistic and an empirical model for simulating soil COS fluxes. The mechanistic model, which performs a spatialization of the Ogée et al. (2016) model, enables us to consider that oxic soils can be net COS producers, as illustrated at some of the observation sites. The interhemispheric gradient of the COS surface atmospheric mixing ratio is marginally improved when all known COS sources and sinks are transported with the LMDZ model. This study also highlights the sensitivity of simulated atmospheric COS concentrations to soil COS flux representation in the northern latitudes. Thus, the uncertainty in soil COS fluxes could complicate GPP estimation using COS in the Northern Hemisphere.

The soil COS budget at the global scale over the 2009–2016 period is -30 GgS yr^{-1}, resulting from the contribution of oxic soils that represent a net sink of -126 GgS yr^{-1} and of anoxic soils that represent a source of $+96$ GgS yr^{-1}. It is to be noted that the contribution from anoxic soils, while leading to a global budget similar to Launois et al. (2015), has a different spatial distribution based on the repartition of regularly flooded wetlands from Tootchi et al. (2019). This repartition seems more accurate, as it also includes anoxic soil COS flux in the tropical region and considers a larger variety of anoxic soils, such as salt marshes and rice paddies.

During this work, we have also shown the importance of considering spatially and temporally variable atmospheric COS concentrations on soil COS fluxes, with an especially large impact at the global scale. This result evidences the impact of the recently decreasing atmospheric COS concentrations on the estimated soil COS fluxes.

Regarding the ORCHIDEE model, we performed a sensitivity study highlighting the key parameters to optimize for the soil models. The impact of soil model parameter optimization was studied at two sites. This study exhibited strong arguments in favour of the mechanistic model, as performing an optimization of the empirical model parameters can lead to aliasing errors and a degradation of the simulated soil water content. A larger database of COS flux measurements at the site scale and especially full year time series would greatly help for the next step, which would be to optimize the parameters of ecosystem COS fluxes.

Appendix A: Parameters, variables, and constants for soil COS models

Table A1. Carbonic anhydrase enhancement factor adapted to ORCHIDEE biomes.

ORCHIDEE biomes	Biomes from Meredith et al. (2019)	f_{CA} value from Meredith et al. (2019) (unitless)
1 – Bare soil	Desert	$13\,000 \pm 5400$
2 – Tropical broadleaved evergreen	Temperate broadleaf forest	$32\,000 \pm 1800$
3 – Tropical broadleaved raingreen	Temperate broadleaf forest	$32\,000 \pm 1800$
4 – Temperate needleleaf evergreen	Temperate coniferous forest	$32\,000 \pm 3100$
5 – Temperate broadleaved evergreen	Temperate broadleaf forest	$32\,000 \pm 1800$
6 – Temperate broadleaved summergreen	Temperate broadleaf forest	$32\,000 \pm 1800$
7 – Boreal needleleaf evergreen	Temperate coniferous forest	$32\,000 \pm 3100$
8 – Boreal broadleaved summergreen	Temperate broadleaf forest	$32\,000 \pm 1800$
9 – Boreal needleleaf summergreen	Temperate coniferous forest	$32\,000 \pm 3100$
10 – C_3 grass	Mediterranean grassland	$17\,000 \pm 9000$
11 – C_4 grass	Mediterranean grassland	$17\,000 \pm 9000$
12 – C_3 agriculture	Agricultural	6500 ± 6900
13 – C_4 agriculture	Agricultural	6500 ± 6900
14 – Tropical C_3 grass	Tropical grassland	45 000
15 – Boreal C_3 grass	Mediterranean grassland	$17\,000 \pm 9000$

Table A2. α and β parameters for COS production term adapted to ORCHIDEE biomes.

ORCHIDEE biomes	Biomes from Whelan et al. (2016)	α parameter from Whelan et al. (2016) (unitless)	β parameter from Whelan et al. (2016) ($°C^{-1}$)
1 – Bare soil	Desert	n/a	n/a
2 – Tropical broadleaved evergreen	Rainforest	−8.2	0.101
3 – Tropical broadleaved raingreen	Rainforest	−8.2	0.101
4 – Temperate needleleaf evergreen	Temperate forest	−7.77	0.119
5 – Temperate broadleaved evergreen	Temperate forest	−7.77	0.119
6 – Temperate broadleaved summergreen	Temperate forest	−7.77	0.119
7 – Boreal needleleaf evergreen	Temperate forest	−7.77	0.119
8 – Boreal broadleaved summergreen	Temperate forest	−7.77	0.119
9 – Boreal needleleaf summergreen	Temperate forest	−7.77	0.119
10 – C_3 grass	Savannah	−9.54	0.108
11 – C_4 grass	Savannah	−9.54	0.108
12 – C_3 agriculture	Soy field	−6.12	0.096
13 – C_4 agriculture	Soy field	−6.12	0.096
14 – Tropical C_3 grass	Savannah	−9.54	0.108
15 – Boreal C_3 grass	Savannah	−9.54	0.108

n/a – not applicable.

Table A3. Variables for the empirical and mechanistic COS soil models.

Variable name	Description	Unit	Reference
Empirical COS soil model			
$F_{\text{soil, empirical}}$	Empirical model soil COS flux	pmol COS m^{-2} s^{-1}	Berry et al. (2013), Yi et al. (2007)
Resp_{tot}	Total (heterotrophic and autotrophic) soil respiration	µmol CO_2 m^{-2} s^{-1}	Yi et al. (2007)
Mechanistic COS soil model			
ε_{tot}	Total soil COS porosity	m^3 air per cubic metre soil	Ogée et al. (2016)
C	Soil COS concentration	mol m^{-3}	Ogée et al. (2016)
F_{diff}	Soil COS diffusional flux	mol m^{-2} s^{-1}	Ogée et al. (2016)
S	Soil COS consumption rate	mol m^{-3} s^{-1}	Ogée et al. (2016)
P	Soil COS production rate	mol m^{-3} s^{-1}	Whelan et al. (2016)
$F_{\text{soil, mechanistic}}$	Soil COS flux in the mechanistic model	mol m^{-2} s^{-1}	Ogée et al. (2016)
k	Total COS consumption rate by soil	s^{-1}	Ogée et al. (2016)
B	Solubility of COS in soil water	m^3 water per cubic metre air	Ogée et al. (2016)
θ	Soil volumetric water content	m^3 water per cubic metre soil	Ogée et al. (2016)
D	Total effective COS diffusivity in soil	m^2 s^{-1}	Ogée et al. (2016)
z_1	Characteristic deep for soil COS flux	m	Ogée et al. (2016)
k_{uncat}	Uncatalysed rate of COS hydrolysis in the soil water	s^{-1}	Elliott et al. (1989)
k_{cat}	Turnover rate of COS enzymatic reaction catalysed by CA	s^{-1}	Ogée et al. (2016)
K_{m}	Michaelis–Menten constant of CA catalysis	mol m^{-3}	Ogée et al. (2016)
x_{CA}	Temperature dependence of the ratio $k_{\text{cat}}/K_{\text{m}}$	–	Ogée et al. (2016)
k	Soil total COS consumption rate	s^{-1}	Ogée et al. (2016)
f_{CA}	CA enhancement factor	–	Meredith et al. (2019)
$D_{\text{eff,a}}$	Effective diffusivity of gaseous COS in soil	m^3 air per metre soil per second	Ogée et al. (2016)
$D_{\text{eff,l}}$	Effective diffusivity of dissolved COS in soil	m^3 water per metre soil per second	Ogée et al. (2016)
K_{H}	Henry's law constant	mol m^{-3} Pa^{-1}	Bird et al. (2002)
$D_{0,\text{a}}$	Binary diffusivity of COS in the free air	m^2 air s^{-1}	Bird et al. (2002)
τ_{a}	Tortuosity factor for gaseous diffusion	–	Ogée et al. (2016)
$\tau_{\text{a,r}}$	Tortuosity factor for gaseous diffusion in repacked soils	–	Moldrup et al. (2003)
$\tau_{\text{a,u}}$	Tortuosity factor for gaseous diffusion in undisturbed soils	–	Deepagoda et al. (2011)
$D_{0,\text{l}}$	Binary diffusivity of COS in the free water	m^2 water s^{-1}	Zeebe (2011)
τ_{l}	Tortuosity factor for solute diffusion	–	Millington and Quirk (1961)
α	COS production parameter	–	Whelan et al. (2016)
β	COS production parameter	–	Whelan et al. (2016)
ORCHIDEE LSM			
p	Pressure		ORCHIDEE LSM
ε_{a}	Air-filled porosity	m^3 air per cubic metre soil	ORCHIDEE LSM
ε_{a}	Total soil porosity (air-filled and water-filled pores)	m^3 m^{-3}	ORCHIDEE LSM
T	Mean soil temperature	K	ORCHIDEE LSM
t	Time	s	ORCHIDEE LSM
z	Depth	m	ORCHIDEE LSM

Table A4. Constants for the empirical and mechanistic COS soil models.

Constant name	Description	Value	Unit	Reference
Empirical COS soil model				
k_{soil}	Constant to convert CO_2 production from respiration to COS uptake	1.2	pmol COS per µmol CO_2	Yi et al. (2007)
Mechanistic COS soil model				
C_a	Ambient air COS concentration when constant (500 ppt)	2.0437×10^{-8}	mol m^{-3}	
z_{max}	Maximum soil depth	0.09	m	ORCHIDEE LSM
pK_w	Dissociation constant of water	14	–	
ΔH_a	Thermodynamic parameter	40	kJ mol^{-1}	Ogée et al. (2016)
ΔH_d	Thermodynamic parameter	200	kJ mol^{-1}	Ogée et al. (2016)
ΔS_d	Thermodynamic parameter	660	J mol^{-1} K^{-1}	Ogée et al. (2016)
R	Ideal gas constant	8.314	J mol^{-1} K^{-1}	
$D_{0,a}(25\,°C, 1\,atm)$	Binary diffusivity of COS in the free air at 25 °C and 1 atm	1.27×10^{-5}	m^2 s^{-1}	Massman (1998)
$D_{0,l}(25\,°C)$	Binary diffusivity of COS in the free water at 25 °C	1.94×10^{-9}	m^2 s^{-1}	Ulshöfer et al. (1996)
Q_{10}	Multiplicative factor of the production rate for a 10 °C temperature rise	2.7	–	Meredith et al. (2018)
P_{ref}	Reference production term	10	pmol m^2 s^{-1}	

Appendix B: Locations and descriptions of the observation sites

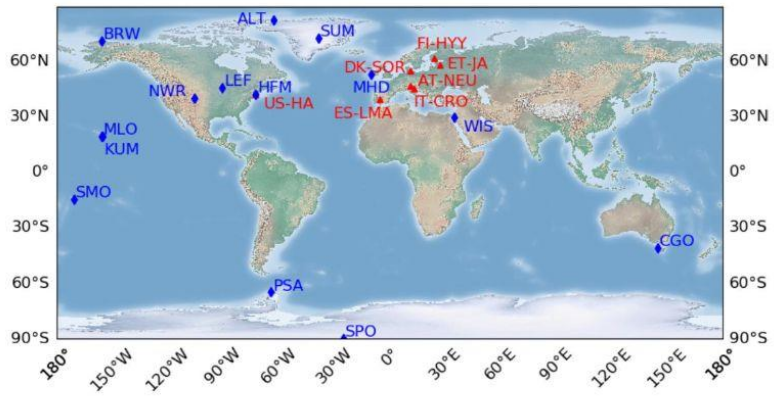

Figure B1. Locations of the observation sites for soil COS flux measurements (red) and atmospheric concentration measurements (blue).

Table B1. List of air sampling sites selected for evaluation of COS concentrations.

Site	Short name	Coordinates	Elevation (metres above sea level)	Comments
South Pole, Antarctica (United States)	SPO	90.0° S, 24.8° E	2810	
Palmer Station, Antarctica (United States)	PSA	64.77° S, 64.05° W	10.0	
Kennaook / Cape Grim, Australia	CGO	40.68° S, 144.69° E	164	Inlet is 70 m aboveground
Tutuila, American Samoa	SMO	14.25° S, 170.56° W	77	
Mauna Loa, United States	MLO	19.54° N, 155.58° W	3397	
Cape Kumukahi, United States	KUM	19.74° N, 155.01° W	3	
Weizmann Institute of Science at the Arava Institute, Ketura, Israel	WIS	29.96° N, 35.06° E	151	
Niwot Ridge, United States	NWR	40.04° N, 105.54° W	3475	
Harvard Forest, United States	HFM	42.54° N, 72.17° W	340	Inlet is 29 m aboveground
Wisconsin, United States	LEF	45.95° N, 90.28° W	868	Inlet is 396 m aboveground on a tall tower
Mace Head, Ireland	MHD	53.33° N, 9.9° W	18	
Utqiaġvik (formerly Barrow), United States	BRW	71.32° N, 155.61° W	8	
Summit, Greenland	SUM	72.6° N, 38.42° W	3200	
Alert, Canada	ALT	82.45° N, 62.51° W	195	

Table B2. Normalized standard deviations (NSDs) of the simulated concentrations by the observed concentrations. Within brackets are the Pearson correlation coefficients (r) between simulated and observed COS concentrations for the mechanistic and empirical approaches, calculated between 2011 and 2015 at selected NOAA stations. For each station, NSD and r values closest to one are in bold, and the farthest ones are in italic. The time series have been detrended beforehand and filtered to remove the synoptic variability (see Sect. 2.3.3).

	SMO	KUM	MLO	NWR	LEF	HFM	MHD	SUM	BRW	ALT
Mechanistic (oxic)	1.1 **(0.8)**	0.7 (0.7)	**0.9** (0.8)	*0.4* (0.4)	*0.2* (0.7)	*0.3* (0.8)	1.5 (0.2)	*0.4* (0.2)	**1.1** *(0.1)*	*0.8* *(0.1)*
Empirical (oxic)	**1.0** (0.7)	0.8 **(0.9)**	1.2 **(0.9)**	**0.8** (0.4)	0.5 **(0.9)**	0.6 **(0.9)**	1.5 **(0.4)**	0.5 (0.6)	*1.3* (0.3)	**0.9** **(0.4)**
Mechanistic (oxic + anoxic)	*1.2* (0.7)	*0.6* *(0.6)*	**0.9** *(0.7)*	0.5 *(0.1)*	*0.2* *(0.2)*	*0.3* *(0.5)*	**1.0** *(0.1)*	*0.4* *(0.0)*	*1.3* *(0.1)*	*0.8* *(0.1)*
Launois (oxic + anoxic)	1.1 *(0.6)*	**1.0** **(0.9)**	*1.4* **(0.9)**	1.4 **(0.7)**	**0.9** **(0.9)**	**0.8** **(0.9)**	*1.6* **(0.4)**	**0.6** **(0.7)**	1.2 **(0.4)**	**0.9** **(0.4)**

Appendix C: Soil COS production term for the mechanistic model

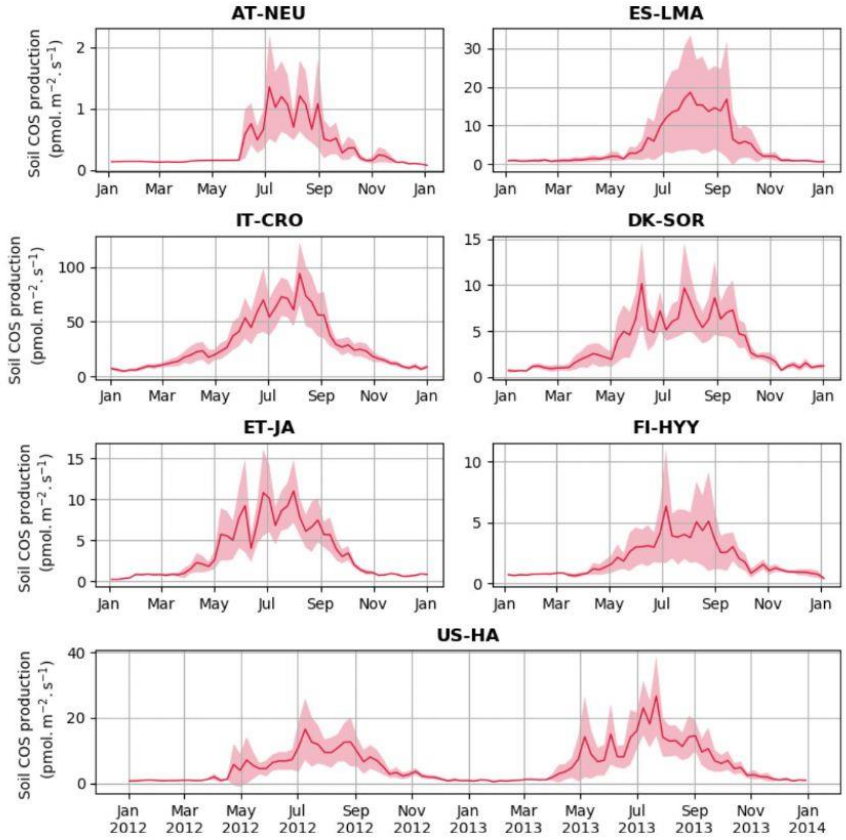

Figure C1. Seasonal cycles of soil COS production with weekly average production at AT-NEU, ES-LMA, IT-CRO, DK-SOR, ET-JA, FI-HYY, and US-HA. The shaded areas above and below the modelled curve represent the standard deviation over a week. Soil COS production was computed with a variable atmospheric COS concentration.

Appendix D: Global-scale soil COS fluxes

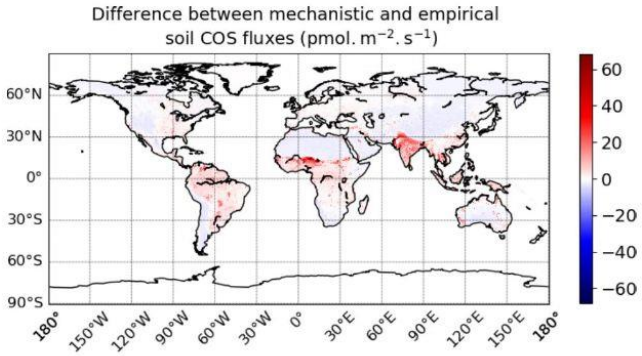

Figure D1. Mean difference between soil COS fluxes computed with the mechanistic and the empirical model over 2010–2019. The map resolution is $0.5° \times 0.5°$.

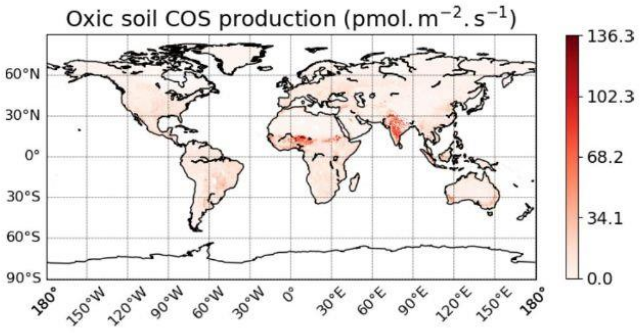

Figure D2. Mean spatial distribution of oxic soil COS production term over 2010–2019. The map resolution is 0.5° × 0.5°.

Appendix E: Prior versus post-optimization parameter values

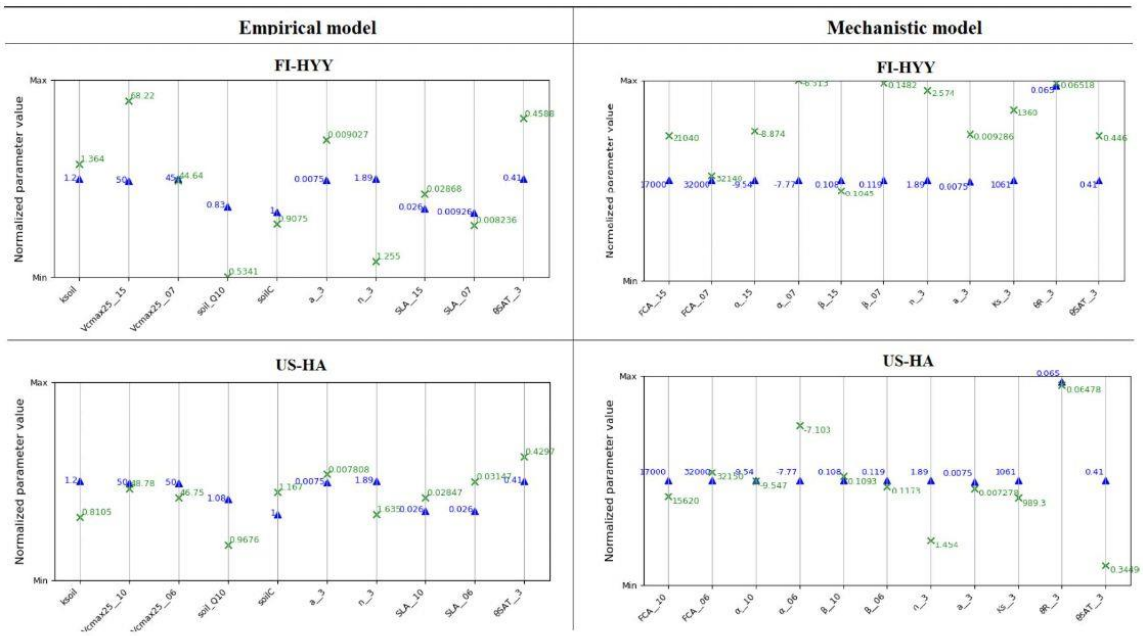

Figure E1. Comparison between prior and post-optimization parameter values at FI-HYY and US-HA. The y axis represents the normalization between the edges of the range of variation for each parameter. Prior values of the parameters are represented in blue, and post-optimization values are in green.

Code availability. The CMIP6 version of the ORCHIDEE model including the soil COS sub-models is available on request to the authors. The LMDZ model is available at http://svn.lmd.jussieu.fr/LMDZ/LMDZ6/ (Laboratoire de Météorologie Dynamique, 2021) under the CeCILL (CEA CNRS INRIA Logiciel Libre) v2 free software license.

Data availability. For FI-HYY, we used the 2015 soil chamber COS measurements described in Sun et al. (2017), which can be found at https://doi.org/10.15146/R39P4R or in Zenodo at https://doi.org/10.5281/zenodo.322936. For US-HA, we used the soil COS flux data derived from eddy covariance COS and CO_2 measurements and soil chamber CO_2 measurements conducted in 2012 and 2013 as described in Wehr et al. (2017). We used the COS flux data published in Kitz et al. (2020; https://doi.org/10.5281/zenodo.3664784, Kitz, 2020) and Spielmann et al. (2019a; https://doi.org/10.5281/zenodo.2586891, Spiel-

mann et al., 2019b) for AT-NEU in 2015, DK-SOR and ES-LMA in 2016, and IT-CRO in 2017.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-19-2427-2022-supplement.

Author contributions. CA, FM, MR, and PP conceived the research. JO advised regarding the spatialization of his mechanistic model. CA and FM coded the ORCHIDEE developments and made the simulations. MR transported all COS sinks and sources with the LMDZ model. FK, FMS, and GW provided the data for AT-NEU, ES-LMA, DK-SOR, IT-CRO, and ET-JA. WS provided the data for the FI-HYY site, and RW provided them for the US-HA site. NR provided code and guidance for the sensitivity analysis and data assimilation experiments. SB, JEC, MEW, DH, STL, US, and DM were consulted on their respective expertise.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors thank the reviewers for their insightful and useful comments which helped to improve this study. The authors are very grateful to everyone who participated in field data collection used in this study. We thank Vladislav Bastrikov for providing the ORCHIDAS code. We also acknowledge Nicolas Vuichard for providing the soil bulk density map used in ORCHIDEE simulations. Operation of the US-HA site is supported by the AmeriFlux Management Project with funding by the US Department of Energy's Office of Science (contract no. DE-AC02-05CH11231), and additionally it is a part of the Harvard Forest Long Term Ecological Research (LTER) site supported by the National Science Foundation (grant no. DEB-1832210). The field campaign at DK-SOR was supported by the Danish ICOS contribution (ICOS/DK) and by the Independent Research Fund Denmark (grant no. DFF-1323-00182).

Financial support. This research has been mainly supported by the 4C project of the European Commission's Horizon 2020 framework programme (grant no. 821003) and to a small extent by VERIFY (grant no. 776810).

Florian Kitz, Felix M. Spielmann, and Georg Wohlfahrt were supported by the Austrian Science Fund (FWF) (contract nos. P26931, P27176, P31669, and I03859) and the University of Innsbruck.

Review statement. This paper was edited by Sönke Zaehle and reviewed by two anonymous referees.

References

Aneja, V. P., Overton, J. H., and Aneja, A. P.: Emission survey of biogenic sulfur flux from terrestrial surfaces, JAPCA J. Air Waste Ma., 31, 256–258, https://doi.org/10.1080/00022470.1981.10465218, 1981.

Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., and Peylin, P.: 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), Geosci. Model Dev., 11, 4739–4754, https://doi.org/10.5194/gmd-11-4739-2018, 2018.

Belviso, S., Schmidt, M., Yver, C., Ramonet, M., Gros, V., and Launois, T.: Strong similarities between night-time deposition velocities of carbonyl sulphide and molecular hydrogen inferred from semi-continuous atmospheric observations in Gif-sur-Yvette, Paris region, Tellus B, 65, 20719, https://doi.org/10.3402/tellusb.v65i0.20719, 2013.

Belviso, S., Lebegue, B., Ramonet, M., Kazan, V., Pison, I., Berchet, A., Delmotte, M., Yver-Kwok, C., Montagne, D., and Ciais, P.: A top-down approach of sources and non-photosynthetic sinks of carbonyl sulfide from atmospheric measurements over multiple years in the Paris region (France), PLOS ONE, 15, e0228419, https://doi.org/10.1371/journal.pone.0228419, 2020.

Berkelhammer, M., Asaf, D., Still, C., Montzka, S., Noone, D., Gupta, M., Provencal, R., Chen, H., and Yakir, D.: Constraining surface carbon fluxes using in situ measurements of carbonyl sulfide and carbon dioxide, Global Biogeochem. Cy., 28, 161–179, https://doi.org/10.1002/2013GB004644, 2014.

Berry, J., Wolf, A., Campbell, J. E., Baker, I., Blake, N., Blake, D., Denning, A. S., Kawa, S. R., Montzka, S. A., Seibt, U., Stimler, K., Yakir, D., and Zhu, Z.: A coupled model of the global cycles of carbonyl sulfide and CO_2: A possible new window on the carbon cycle, J. Geophys. Res.-Biogeo., 118, 842–852, https://doi.org/10.1002/jgrg.20068, 2013.

Bird, B., Stewart, W., and Lightfoot, E.: Transport Phenomena, John Wiley and Sons, Chemical Engineering Department, University of Wisconsin-Madison, 2002.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J. L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M. A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J. Y., Guenet, B., Guez, L. E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J. B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J. Adv. Model. Earth Sy., 12, e2019MS002010, https://doi.org/10.1029/2019MS002010, 2020.

Brændholt, A., Ibrom, A., Larsen, K. S., and Pilegaard, K.: Partitioning of ecosystem respiration in a beech forest, Agr. Forest Meteorol., 252, 88–98, https://doi.org/10.1016/j.agrformet.2018.01.012, 2018.

Burkholder, J. B., Sander, S. P., Abbatt, J. P. D., Barker, J. R., Cappa, C., Crounse, J. D., Dibble, T. S., Huie, R. E., Kolb, C. E., Kurylo, M. J., Orkin, V. L., Percival, C. J., Wilmouth, D. M., and Wine, P. H.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation No. 19, JPL Publication 19-xx, Jet Propulsion Laboratory, Pasadena, http://jpldataeval.jpl.nasa.gov (last access: 9 May 2022), 2019.

Campbell, J. E., Carmichael, G. R., Chai, T., Mena-Carrasco, M., Tang, Y., Blake, D. R., Blake, N. J., Vay, S. A., Collatz, G. J., Baker, I., Berry, J. A., Montzka, S. A., Sweeney, C., Schnoor, J. L., and Stanier, C. O.: Photosynthetic control of atmospheric carbonyl sulfide during the growing season, Science, 322, 1085–1088, https://doi.org/10.1126/science.1164015, 2008.

Campolongo, F., Cariboni, J., and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environ. Modell. Softw., 22, 1509–1518, https://doi.org/10.1016/j.envsoft.2006.10.004, 2007.

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J. L., Ghattas, J., Grandpeix, J. Y., Idelkadi, A., Mellul, L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., Wang, F., and Zhao, Y.: Improved Near-Surface Continental Climate in IPSL-CM6A-LR by Combined Evolutions of Atmospheric and Land Surface Physics, J. Adv. Model. Earth Sy., 12, e2019MS002005, https://doi.org/10.1029/2019MS002005, 2020.

Chin, M. and Davis, D. D.: A reanalysis of carbonyl sulfide as a source of stratospheric background sulfur aerosol, J. Geophys. Res.-Atmos., 100, 8993–9005, https://doi.org/10.1029/95JD00275, 1995.

Dantec-Nédélec, S., Ottlé, C., Wang, T., Guglielmo, F., Maignan, F., Delbart, N., Valdayskikh, V., Radchenko, T., Nekrasova, O., Zakharov, V., and Jouzel, J.: Testing the capability of ORCHIDEE land surface model to simulate Arctic ecosystems: Sensitivity analysis and site-level model calibration, J. Adv. Model. Earth Sy., 9, 1212–1230, https://doi.org/10.1002/2016MS000860, 2017.

Davidson, C., Amrani, A., and Angert, A.: Tropospheric carbonyl sulfide mass balance based on direct measurements of sulfur isotopes, P. Natl. Acad. Sci. USA, 118, e2020060118, https://doi.org/10.1073/pnas.2020060118, 2021.

Deepagoda, T. K. K. C., Moldrup, P., Schjønning, P., de Jonge, L. W., Kawamoto, K., and Komatsu, T.: Density-Corrected Models for Gas Diffusivity and Air Permeability in Unsaturated Soil, Vadose Zone J., 10, 226–238, https://doi.org/10.2136/vzj2009.0137, 2011.

de Mello, W. Z. and Hines, M. E.: Application of static and dynamic enclosures for determining dimethyl sulfide and carbonyl sulfide exchange in Sphagnum peatlands: Implications for the magnitude and direction of flux, J. GEOPHYS. RES., 601–615, 1994.

Devai, I. and DeLaune, R. D.: Trapping Efficiency of Various Solid Adsorbents for Sampling and Quantitative Gas Chromatographic Analysis of Carbonyl Sulfide, Anal. Lett., 30, 187–198, https://doi.org/10.1080/00032719708002300, 1997.

Elliott, S., Lu, E., and Rowland, F. S.: Rates and mechanisms for the hydrolysis of carbonyl sulfide in natural waters, Environ. Sci. Technol., 23, 458–461, https://doi.org/10.1021/es00181a011, 1989.

El-Madany, T. S., Reichstein, M., Perez-Priego, O., Carrara, A., Moreno, G., Pilar Martín, M., Pacheco-Labrador, J., Wohlfahrt, G., Nieto, H., Weber, U., Kolle, O., Luo, Y. P., Carvalhais, N., and Migliavacca, M.: Drivers of spatio-temporal variability of carbon dioxide and energy fluxes in a Mediterranean savanna ecosystem, Agr. Forest Meteorol., 262, 258–278, https://doi.org/10.1016/j.agrformet.2018.07.010, 2018.

Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry, Atmos. Chem. Phys., 6, 2273–2319, https://doi.org/10.5194/acp-6-2273-2006, 2006.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth Syst. Sci. Data, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

Glatthor, N., Höpfner, M., Baker, I. T., Berry, J., Campbell, J. E., Kawa, S. R., Krysztofiak, G., Leyser, A., Sinnhuber, B. M., Stiller, G. P., Stinecipher, J., and von Clarmann, T.: Tropical sources and sinks of carbonyl sulfide observed from space, Geophys. Res. Lett., 42, 10082–10090, https://doi.org/10.1002/2015GL066293, 2015.

Goldberg, D. E.: Genetic algorithms in search, optimization, and machine learning, Addison-Wesley Publishing Company, 1989.

Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M. A., Walters, S., Lamarque, J. F., and Holland, E. A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, J. Geophys. Res.-Atmos., 109, D04314, https://doi.org/10.1029/2003jd003957, 2004.

Hauglustaine, D. A., Balkanski, Y., and Schulz, M.: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate, Atmos. Chem. Phys., 14, 11031–11063, https://doi.org/10.5194/acp-14-11031-2014, 2014.

Haupt, R. L. and Haupt, S. E.: Practical Genetic Algorithms, Wiley, 2004.

Helmig, D., Apel, E., Blake, D., Ganzeveld, L. N., Lefer, B. L., Meinardi, S., Swanson, A. L.: Release and uptake of

volatile inorganic and organic gases through the snowpack at Niwot Ridge, Colorado, Biogeochemistry, 95, 167–183, https://doi.org/10.1007/s10533-009-9326-8, 2009.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., de Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hörtnagl, L. and Wohlfahrt, G.: Methane and nitrous oxide exchange over a managed hay meadow, Biogeosciences, 11, 7219–7236, https://doi.org/10.5194/bg-11-7219-2014, 2014.

Hörtnagl, L., Bamberger, I., Graus, M., Ruuskanen, T. M., Schnitzhofer, R., Müller, M., Hansel, A., and Wohlfahrt, G.: Biotic, abiotic, and management controls on methanol exchange above a temperate mountain grassland, J. Geophys. Res.-Biogeo., 116, G03021, https://doi.org/10.1029/2011JG001641, 2011.

Hourdin, F. and Issartel, J. P.: Sub-surface nuclear tests monitoring through the CTBT xenon network, Geophys. Res. Lett., 27, 2245–2248, https://doi.org/10.1029/1999GL010909, 2000.

Hourdin, F., Rio, C., Grandpeix, J. Y., Madeleine, J. B., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L., Foujols, M. A., Mellul, L., Traore, A. K., Dufresne, J. L., Boucher, O., Lefebvre, M. P., Millour, E., Vignon, E., Jouhaud, J., Diallo, F. B., Lott, F., Gastineau, G., Caubel, A., Meurdesoif, Y., and Ghattas, J.: LMDZ6A: The Atmospheric Component of the IPSL Climate Model With Improved and Better Tuned Physics, J. Adv. Model. Earth Sy., 12, e2019MS001892, https://doi.org/10.1029/2019MS001892, 2020.

Kaisermann, A., Jones, S., Wohl, S., Ogée, J., and Wingate, L.: Nitrogen fertilization reduces the capacity of soils to take up atmospheric carbonyl sulphide, Soil Systems, 2, 62, https://doi.org/10.3390/soilsystems2040062, 2018.

Kanda, K. I., Tsuruta, H., and Minami, K.: Emission of dimethyl sulfide, carbonyl sulfide, and carbon bisulfide from paddy fields, Soil Sci. Plant Nutr., 38, 709–716, https://doi.org/10.1080/00380768.1992.10416701, 1992.

Kesselmeier, J., Teusch, N., and Kuhn, U.: Controlling variables for the uptake of atmospheric carbonyl sulfide by soil, J. Geophys. Res.-Atmos., 104, 11577–11584, https://doi.org/10.1029/1999JD900090, 1999.

Kettle, A. J., Kuhn, U., von Hobe, M., Kesselmeier, J., and Andreae, M. O.: Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks, J. Geophys. Res.-Atmos., 107, ACH 25-1–ACH 25-16, https://doi.org/10.1029/2002JD002187, 2002.

Kitz, F.: Data for "Soil COS exchange: a comparison of three European ecosystems", Zenodo [data set], https://doi.org/10.5281/zenodo.3664784, 2020.

Kitz, F., Gerdel, K., Hammerle, A., Laterza, T., Spielmann, F. M., and Wohlfahrt, G.: In situ soil COS exchange of a temperate mountain grassland under simulated drought, Oecologia, 183, 851–860, https://doi.org/10.1007/s00442-016-3805-0, 2017.

Kitz, F., Spielmann, F. M., Hammerle, A., Kolle, O., Migliavacca, M., Moreno, G., Ibrom, A., Krasnov, D., Noe, S. M., and Wohlfahrt, G.: Soil COS Exchange: A Comparison of Three European Ecosystems, Global Biogeochem. Cyc., 34, e2019GB006202, https://doi.org/10.1029/2019GB006202, 2020.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., and Nikinmaa, E.: CO_2 exchange and component CO_2 fluxes of a boreal Scots pine forest, Boreal Environ. Res., 14, 761–783, 2009.

Kooijmans, L. M. J., Cho, A., Ma, J., Kaushik, A., Haynes, K. D., Baker, I., Luijkx, I. T., Groenink, M., Peters, W., Miller, J. B., Berry, J. A., Ogée, J., Meredith, L. K., Sun, W., Kohonen, K.-M., Vesala, T., Mammarella, I., Chen, H., Spielmann, F. M., Wohlfahrt, G., Berkelhammer, M., Whelan, M. E., Maseyk, K., Seibt, U., Commane, R., Wehr, R., and Krol, M.: Evaluation of carbonyl sulfide biosphere exchange in the Simple Biosphere Model (SiB4), Biogeosciences, 18, 6547–6565, https://doi.org/10.5194/bg-18-6547-2021, 2021.

Kuai, L., Worden, J. R., Campbell, J. E., Kulawik, S. S., Li, K. F., Lee, M., Weidner, R. J., Montzka, S. A., Moore, F. L., Berry, J. A., Baker, I., Denning, A. S., Bian, H., Bowman, K. W., Liu, J., and Yung, Y. L.: Estimate of carbonyl sulfide tropical oceanic surface fluxes using aura tropospheric emission spectrometer observations, J. Geophys. Res., 120, 11012–11023, https://doi.org/10.1002/2015JD023493, 2015.

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, Geosci. Model Dev., 7, 2581–2597, https://doi.org/10.5194/gmd-7-2581-2014, 2014.

Laboratoire de Météorologie Dynamique: LMDZ model, Laboratoire de Météorologie Dynamique [code], http://svn.lmd.jussieu.fr/LMDZ/LMDZ6/, last access: 21 October 2021.

Lardy, R., Bellocchi, G., and Soussana, J. F.: A new method to determine soil organic carbon equilibrium, Environ. Modell. Softw., 26, 1759–1763, https://doi.org/10.1016/j.envsoft.2011.05.016, 2011.

Launois, T., Peylin, P., Belviso, S., and Poulter, B.: 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, Atmos. Chem. Phys., 15, 9285–9312, https://doi.org/10.5194/acp-15-9285-2015, 2015.

Lennartz, S. T., Marandino, C. A., von Hobe, M., Cortes, P., Quack, B., Simo, R., Booge, D., Pozzer, A., Steinhoff, T., Arevalo-Martinez, D. L., Kloss, C., Bracher, A., Röttgers, R., Atlas, E., and Krüger, K.: Direct oceanic emissions unlikely to account for the missing source of atmospheric carbonyl sulfide, Atmos. Chem. Phys., 17, 385–402, https://doi.org/10.5194/acp-17-385-2017, 2017.

Lennartz, S. T., Marandino, C. A., von Hobe, M., Andreae, M. O., Aranami, K., Atlas, E., Berkelhammer, M., Bingemer, H., Booge, D., Cutter, G., Cortes, P., Kremser, S., Law, C. S., Marriner, A., Simó, R., Quack, B., Uher, G., Xie, H., and Xu, X.: Marine carbonyl sulfide (OCS) and carbon disulfide (CS2): a compilation of measurements in seawater and the marine boundary layer, Earth Syst. Sci. Data, 12, 591–609, https://doi.org/10.5194/essd-12-591-2020, 2020.

Lennartz, S. T., Gauss, M., von Hobe, M., and Marandino, C. A.: Monthly resolved modelled oceanic emissions of carbonyl sulphide and carbon disulphide for the period 2000–2019, Earth Syst. Sci. Data, 13, 2095–2110, https://doi.org/10.5194/essd-13-2095-2021, 2021.

Liu, J., Geng, C., Mu, Y., Zhang, Y., Xu, Z., and Wu, H.: Exchange of carbonyl sulfide (COS) between the atmosphere and various soils in China, Biogeosciences, 7, 753–762, https://doi.org/10.5194/bg-7-753-2010, 2010.

Lopez-Sangil, L., Rousk, J., Wallander, H., and Casals, P.: Microbial growth rate measurements reveal that land-use abandonment promotes a fungal dominance of SOM decomposition in grazed Mediterranean ecosystems, Biol. Fert. Soils, 47, 129–138, https://doi.org/10.1007/s00374-010-0510-8, 2011.

Ma, X., Huang, J., Zhao, T., Liu, C., Zhao, K., Xing, J., and Xiao, W.: Rapid increase in summer surface ozone over the North China Plain during 2013–2019: a side effect of particulate matter reduction control?, Atmos. Chem. Phys., 21, 1–16, https://doi.org/10.5194/acp-21-1-2021, 2021.

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., Köhler, P., Gómez-Dans, J., and Disney, M.: Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data, Sci. Rep.-UK, 8, 1973, https://doi.org/10.1038/s41598-018-20024-w, 2018.

Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., Bastrikov, V., and MacBean, N.: Optimizing Carbon Cycle Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Mean Net CO_2 Flux and its Inter-Annual 2 Variability, J. Geophys. Res.-Biogeo., 126, e2021JG006400, https://doi.org/10.1029/2021JG006400, 2021.

Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K.-M., Commane, R., Wehr, R., Campbell, J. E., Belviso, S., Montzka, S. A., Raoult, N., Seibt, U., Shiga, Y. P., Vuichard, N., Whelan, M. E., and Peylin, P.: Carbonyl sulfide: comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach, Biogeosciences, 18, 2917–2955, https://doi.org/10.5194/bg-18-2917-2021, 2021.

Masaki, Y., Iizuka, R., Kato, H., Kojima, Y., Ogawa, T., Yoshida, M., Matsushita, Y., and Katayama, Y.: Fungal carbonyl sulfide hydrolase of trichoderma harzianum strain thif08 and its relationship with clade D β-carbonic anhydrases, Microbes Environ., 36, ME20058, https://doi.org/10.1264/jsme2.ME20058, 2021.

Maseyk, K., Berry, J. A., Billesbach, D., Campbell, J. E., Torn, M. S., Zahniser, M., and Seibt, U.: Sources and sinks of carbonyl sulfide in an agricultural field in the Southern Great Plains, P. Natl. Acad. Sci. USA, 111, 9064–9069, https://doi.org/10.1073/pnas.1319132111, 2014.

Masotti, I., Belviso, S., Bopp, L., Tagliabue, A., and Bucciarelli, E.: Effects of light and phosphorus on summer DMS dynamics in subtropical waters using a global ocean biogeochemical model, Environ. Chem., 13, 379–389, https://doi.org/10.1071/EN14265, 2016.

Massman, W. J.: A review of the molecular diffusivities of H_2O, CO_2, CH_4, CO, O_3, SO_2, NH_3, N_2O, NO, and NO_2 in air, O_2 and N_2 near STP, Atmos. Environ., 32, 1111–1127, 1998.

Meredith, L. K., Boye, K., Youngerman, C., Whelan, M., Ogée, J., Sauze, J., and Wingate, L.: Coupled biological and abiotic mechanisms driving carbonyl sulfide production in soils, Soil Systems, 2, 1–27, https://doi.org/10.3390/soilsystems2030037, 2018.

Meredith, L. K., Ogée, J., Boye, K., Singer, E., Wingate, L., von Sperber, C., Sengupta, A., Whelan, M., Pang, E., Keiluweit, M., Brüggemann, N., Berry, J. A., and Welander, P. v.: Soil exchange rates of COS and $CO^{18}O$ differ with the diversity of microbial communities and their carbonic anhydrase enzymes, ISME J., 13, 290–300, https://doi.org/10.1038/s41396-018-0270-2, 2019.

Millington, R. J. and Quirk, J. P.: Permeability of porous solids, T. Faraday Soc., 57, 1200–1207, https://doi.org/10.1039/TF9615701200, 1961.

Moldrup, P., Olesen, T., Komatsu, T., Yoshikawa, S., Schjønning, P., and Rolston, D. E.: Modeling Diffusion and Reaction in Soils: X. A Unifying Model for Solute and Gas Diffusivity in Unsaturated Soil, Soil Sci., 168, 2003.

Montzka, S. A., Calvert, P., Hall, B. D., Elkins, J. W., Conway, T. J., Tans, P. P., and Sweeney, C. S.: On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO_2, J. Geophys. Res.-Atmos., 112, D09302, https://doi.org/10.1029/2006JD007665, 2007.

Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33, 161–174, https://doi.org/10.1080/00401706.1991.10484804, 1991.

Noe, S. M., Kimmel, V., Hüve, K., Copolovici, L., Portillo-Estrada, M., Püttsepp, Ü., Jõgiste, K., Niinemets, Ü., Hörtnagl, L., and Wohlfahrt, G.: Ecosystem-scale biosphere-atmosphere interactions of a hemiboreal mixed forest stand at Järvselja, Estonia, Forest Ecol. Manage., 262, 71–81, https://doi.org/10.1016/j.foreco.2010.09.013, 2011.

Noe, S. M., Niinemets, Ü., Krasnova, A., Krasnov, D., Motallebi, A., Kängsepp, V., Jõgiste, K., Hõrrak, U., Komsaare, K., Mirme, S., Vana, M., Tammet, H., Bäck, J., Vesala, T., Kulmala, M., Petäjä, T., and Kangur, A.: SMEAR Estonia: Perspectives of a large-scale forest ecosystem – Atmosphere research infrastructure, Forestry Studies, 63, 56–84, https://doi.org/10.1515/fsmu-2015-0009, 2015.

Ogée, J., Sauze, J., Kesselmeier, J., Genty, B., Van Diest, H., Launois, T., and Wingate, L.: A new mechanistic framework to predict OCS fluxes from soils, Biogeosciences, 13, 2221–2240, https://doi.org/10.5194/bg-13-2221-2016, 2016.

Parazoo, N. C., Denning, A. S., Berry, J. A., Wolf, A., Randall, D. A., Kawa, S. R., Pauluis, O., and Doney, S. C.: Moist synoptic transport of CO_2 along the mid-latitude storm track, Geophys. Res. Lett., 38, L09804, https://doi.org/10.1029/2011GL047238, 2011.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Ribeca, A., van Ingen, C., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. v., Bonal, D., Bonnefond, J. M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K.

J., de Cinti, B., de Grandcourt, A., de Ligne, A., de Oliveira, R. C., Delpierre, N., Desai, A. R., di Bella, C. M., di Tommasi, P., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grünwald, T., Hammerle, A., Han, S., Han, X., Ulf Hansen, B., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hörtnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janouš, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J., Lion, M., Liska, A. J., Lohila, A., López-Ballesteros, A., López-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Lüers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Hatala Matthes, J., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S, Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K, Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosetto, M., Nouvellon, Y., Novick, K., Oechel, W., Eivind Olesen, J., Ourcival, J., Papuga, S. A., Parmentier, F., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, U., Raz-Yaseef, N., Rebmann, C., Reed, D., Resco de Dios, V., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sánchez-Cañete, E. P., Sanchez-Mejia, Z. M., Peter Schmid, H., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedlák, P., Serrano-Ortíz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci. Data, 7, 225, https://doi.org/10.1038/s41597-020-0534-3, 2020.

Peylin, P., Ciais, P., Denning, A. S., Tans, P. P., Berry, J. A., and White, J. W. C.: A 3-dimensional study of δ 18 O in atmospheric CO_2: contribution of different land ecosystems, Tellus B, 51, 642–667, https://doi.org/10.3402/tellusb.v51i3.16452, 1999.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, Geosci. Model Dev., 9, 3321–3346, https://doi.org/10.5194/gmd-9-3321-2016, 2016.

Pilegaard, K., Ibrom, A., Courtney, M. S., Hummelshøj, P., and Jensen, N. O.: Increasing net CO_2 uptake by a Danish beech forest during the period from 1996 to 2009, Agr. Forest Meteorol., 151, 934–946, https://doi.org/10.1016/j.agrformet.2011.02.013, 2011.

Protoschill-Krebs, G., Wilhelm, C., and Kesselmeier, J.: Consumption of Carbonyl Sulphide (COS) by Higher Plant Carbonic Anhydrase (CA), Atmos. Environ., 30, 3151–3156, 1996.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative, Geosci. Model Dev., 8, 2315–2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., and Maugis, P.: Evaluating and Optimizing Surface Soil Moisture Drydowns in the ORCHIDEE Land Surface Model at In Situ Locations, J. Hydrometeorol., 22, 1025–1043, https://doi.org/10.1175/JHM-D-20-0115.1, 2021.

Remaud, M., Chevallier, F., Cozic, A., Lin, X., and Bousquet, P.: On the impact of recent developments of the LMDz atmospheric general circulation model on the simulation of CO_2 transport, Geosci. Model Dev., 11, 4489–4513, https://doi.org/10.5194/gmd-11-4489-2018, 2018.

Remaud, M., Chevallier, F., Maignan, F., Belviso, S., Berchet, A., Parouffe, A., Abadie, C., Bacour, C., Lennartz, S., and Peylin, P.: Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling, Atmos. Chem. Phys., 22, 2525–2552, https://doi.org/10.5194/acp-22-2525-2022, 2022.

Reynolds, C. A., Jackson, T. J., and Rawls, W. J.: 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 Resour. Res., 36, 3653–3662, https://doi.org/10.1029/2000WR900130, 2000.

Sandoval-Soto, L., Stanimirov, M., von Hobe, M., Schmitt, V., Valdes, J., Wild, A., and Kesselmeier, J.: Global uptake of carbonyl sulfide (COS) by terrestrial vegetation: Estimates corrected by deposition velocities normalized to the uptake of carbon dioxide (CO_2), Biogeosciences, 2, 125–132, https://doi.org/10.5194/bg-2-125-2005, 2005.

Sauze, J., Ogée, J., Maron, P.-A., Crouzet, O., Nowak, V., Wohl, S., Kaisermann, A., Jones, S., Wingate, L., Ee, O., and Jones, S. P.: The interaction of soil phototrophs and fungi with pH and their impact on soil CO_2, $CO^{18}O$ and OCS exchange, Soil Biol. Biochem., 115, 371–382, https://doi.org/10.1016/j.soilbio.2017.09.009, 2017.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional

sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Smith, K., Jakubzick, C., Whittam, T., and Ferry, J.: Carbonic anhydrase is an ancient enzyme widespread in prokaryotes, P. Natl. Acad. Sci. USA, 96, 15184–15189, 1999.

Smith, K. S. and Ferry, J. G.: Prokaryotic carbonic anhydrases, FEMS Microbiol. Rev., 24, 335–366, https://doi.org/10.1111/j.1574-6976.2000.tb00546.x, 2000.

Spielmann, F. M., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., Ibrom, A., El-Madany, T. S., Gerdel, K., Moreno, G., Kolle, O., Karl, T., Peressotti, A., and Delle Vedove, G.: Gross Primary Productivity of Four European Ecosystems Constrained by Joint CO_2 and COS Flux Measurements, Geophys. Res. Lett., 46, 5284–5293, https://doi.org/10.1029/2019GL082006, 2019a.

Spielmann, F. M., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., Ibrom, A., El-Madany, T., Gerdel, K., Moreno, G., Kolle, O., Karl, T., Peressotti, A., and Delle Vedove, G.: Dataset for "Gross primary productivity of four European ecosystems constrained by joint CO_2 and COS flux measurements", Zenodo [data set], https://doi.org/10.5281/zenodo.2586891, 2019b.

Spielmann, F. M., Hammerle, A., Kitz, F., Gerdel, K., and Wohlfahrt, G.: Seasonal dynamics of the COS and CO_2 exchange of a managed temperate grassland, Biogeosciences, 17, 4281–4295, https://doi.org/10.5194/bg-17-4281-2020, 2020.

Stinecipher, J. R., Cameron-Smith, P. J., Blake, N. J., Kuai, L., Lejeune, B., Mahieu, E., Simpson, I. J., and Campbell, J. E.: Biomass Burning Unlikely to Account for Missing Source of Carbonyl Sulfide, Geophys. Res. Lett., 46, 14912–14920, https://doi.org/10.1029/2019GL085567, 2019.

Sun, W., Maseyk, K. S., Juarez, S., Lett, C., and Seibt, U. H.: Soil-atmosphere carbonyl sulfide (COS) exchange in a tropical rainforest at La Selva, Costa Rica, AGU Fall Meeting Abstracts, 2014, B41C-0075, 2014.

Sun, W., Maseyk, K., Lett, C., and Seibt, U.: A soil diffusion–reaction model for surface COS flux: COSSM v1, Geosci. Model Dev., 8, 3055–3070, https://doi.org/10.5194/gmd-8-3055-2015, 2015.

Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., Levula, J., Keskinen, H., and Seibt, U.: Dataset for "Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland", Dryad [data set], https://doi.org/10.15146/R39P4R, or Zenodo [data set], https://doi.org/10.5281/zenodo.322936, 2017.

Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., Levula, J., Keskinen, H., and Seibt, U.: Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland, Atmos. Chem. Phys., 18, 1363–1378, https://doi.org/10.5194/acp-18-1363-2018, 2018.

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., Andrews, A. E., Lang, P. M., Neff, D., Dlugokencky, E., Miller, J. B., Montzka, S. A., Miller, B. R., Masarie, K. A., Biraud, S. C., Novelli, P. C., Crotwell, M., Crotwell, A. M., Thoning, K., and Tans, P. P.: Seasonal climatology of CO_2 across north america from aircraft measurements in the NOAA/ESRL global greenhouse gas reference network, J. Geophys. Res., 120, 5155–5190, https://doi.org/10.1002/2014JD022591, 2015.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Soc. Ind. Appl. Math., 354 pp., 2005.

Terrenoire, E., Hauglustaine, D., Cohen, Y., Cozic, A., Valorso, R., Lefèvre, F., and Matthes, S.: Impact of present and future aircraft NO_x and aerosol emissions on atmospheric composition and associated direct radiative forcing of climate, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2022-222, in review, 2022.

Thoning, K. W., Tans, P. P., and Komhyr, W. D.: Atmospheric Carbon Dioxide at Mauna Loa Observatory 2. Analysis of the NOAA GMCC Data, 1974–1985, J. Geophys. Res., 8549–8565, 1989.

Tootchi, A., Jost, A., and Ducharne, A.: Multi-source global wetland maps combining surface water imagery and groundwater constraints, Earth Syst. Sci. Data, 11, 189–220, https://doi.org/10.5194/essd-11-189-2019, 2019.

Ulshöfer, V. S., Flijck, O. R., Uher, G., and Andreae, M. O.: Photochemical production and air-sea exchange of sulfide in the eastern Mediterranean Sea, Marine Chemistry, Mediterranean Sea, Mar. Chem., 25–39, 1996.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res.-Biogeo., 112, https://doi.org/10.1029/2006JG000293, 2007.

Van Diest, H. and Kesselmeier, J.: Soil atmosphere exchange of carbonyl sulfide (COS) regulated by diffusivity depending on water-filled pore space, Biogeosciences, 5, 475–483, https://doi.org/10.5194/bg-5-475-2008, 2008.

van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J., 44, 892–898, 1980.

Wania, R., Ross, I., and Prentice, I. C.: Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1, Geosci. Model Dev., 3, 565–584, https://doi.org/10.5194/gmd-3-565-2010, 2010.

Watts, S. F.: The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide, Atmos. Environ, 34, 761–779, https://doi.org/10.1016/S1352-2310(99)00342-8, 2000.

Wehr, R., Commane, R., Munger, J. W., McManus, J. B., Nelson, D. D., Zahniser, M. S., Saleska, S. R., and Wofsy, S. C.: Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake, Biogeosciences, 14, 389–401, https://doi.org/10.5194/bg-14-389-2017, 2017.

Weiner, T., Gross, A., Moreno, G., Migliavacca, M., Schrumpf, M., Reichstein, M., Hilman, B., Carrara, A., and Angert, A.: Following the Turnover of Soil Bioavailable Phosphate in Mediterranean Savanna by Oxygen Stable Isotopes, J. Geophys. Res.-Biogeo., 123, 1850–1862, https://doi.org/10.1029/2017JG004086, 2018.

Whelan, M. E. and Rhew, R. C.: Carbonyl sulfide produced by abiotic thermal and photodegradation of soil organic matter from wheat field substrate, J. Geophys. Res.-Biogeo., 120, 54–62, https://doi.org/10.1002/2014JG002661, 2015.

Whelan, M. E., Min, D. H., and Rhew, R. C.: Salt marsh vegetation as a carbonyl sulfide (COS) source to the atmosphere, Atmos. Environ., 73, 131–137, https://doi.org/10.1016/j.atmosenv.2013.02.048, 2013.

Whelan, M. E., Hilton, T. W., Berry, J. A., Berkelhammer, M., Desai, A. R., and Campbell, J. E.: Carbonyl sulfide exchange in soils for better estimates of ecosystem carbon uptake, Atmos. Chem. Phys., 16, 3711–3726, https://doi.org/10.5194/acp-16-3711-2016, 2016.

Whelan, M. E., Lennartz, S. T., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L. M. J., Hilton, T. W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K.-F., Yakir, D., Zumkehr, A., Katayama, Y., Ogée, J., Spielmann, F. M., Kitz, F., Rastogi, B., Kesselmeier, J., Marshall, J., Erkkilä, K.-M., Wingate, L., Meredith, L. K., He, W., Bunk, R., Launois, T., Vesala, T., Schmidt, J. A., Fichot, C. G., Seibt, U., Saleska, S., Saltzman, E. S., Montzka, S. A., Berry, J. A., and Campbell, J. E.: Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles, Biogeosciences, 15, 3625–3657, https://doi.org/10.5194/bg-15-3625-2018, 2018.

Wilhelm, E., Battino, R., and Wilcock, R. J.: Low-pressure solubility of gases in liquid water, Chem. Rev., 77, 219–262, https://doi.org/10.1021/cr60306a003, 1977.

Wohlfahrt, G., Brilli, F., Hörtnagl, L., Xu, X., Bingemer, H., Hansel, A., and Loreto, F.: Carbonyl sulfide (COS) as a tracer for canopy photosynthesis, transpiration and stomatal conductance: Potential and limitations, Plant Cell Environ., 35, 657–667, https://doi.org/10.1111/j.1365-3040.2011.02451.x, 2012.

Wu, J., Larsen, K. S., van der Linden, L., Beier, C., Pilegaard, K., and Ibrom, A.: Synthesis on the carbon budget and cycling in a Danish, temperate deciduous forest, Agr. Forest Meteorol., 181, 94–107, https://doi.org/10.1016/j.agrformet.2013.07.012, 2013.

Yan, Y., Li, R., Peng, L., Yang, C., Liu, C., Cao, J., Yang, F., Li, Y., and Wu, J.: Emission inventory of carbonyl sulfide (COS) from primary anthropogenic sources in China, Environ. Pollut., 247, 745–751, https://doi.org/10.1016/j.envpol.2019.01.096, 2019.

Yang, F., Qubaja, R., Tatarinov, F., Stern, R., and Yakir, D.: Soil–atmosphere exchange of carbonyl sulfide in a Mediterranean citrus orchard, Atmos. Chem. Phys., 19, 3873–3883, https://doi.org/10.5194/acp-19-3873-2019, 2019.

Yi, Z., Wang, X., Sheng, G., Zhang, D., Zhou, G., and Fu, J.: Soil uptake of carbonyl sulfide in subtropical forests with different successional stages in south China, J. Geophys. Res.-Atmos., 112, D08302, https://doi.org/10.1029/2006JD008048, 2007.

Zeebe, R. E.: On the molecular diffusion coefficients of dissolved CO_2, HCO_3^-, and CO_3^{2-} and their dependence on isotopic mass, Geochim. Cosmochim. Ac., 75, 2483–2498, https://doi.org/10.1016/j.gca.2011.02.010, 2011.

Zhang, J., Wang, L., and Yang, Z.: Emission of Biogenic Sulfur Gases from the Microbial Decomposition of Cystine in Chinese Rice Paddy Soils Environmental Contamination and Toxicology, B. Environ. Contam. Tox., 850–857, 2004.

Zobler, L.: A World Soil File for Global Climate Modelling, NASA Technical Memorandum 87802, NASA Goddard Institute for Space Studies, New York, USA, 1986.

Zumkehr, A., Hilton, T. W., Whelan, M., Smith, S., Kuai, L., Worden, J., and Campbell, J. E.: Global gridded anthropogenic emissions inventory of carbonyl sulfide, Atmos. Environ., 183, 11–19, https://doi.org/10.1016/j.atmosenv.2018.03.063, 2018.

3.3 Improving ecosystem COS flux representation in LSMs

Since the implementation of ecosystem COS fluxes in ORCHIDEE, recent studies on vegetation and soil COS exchanges have revealed opportunities to improve the representation of these COS fluxes. These studies offer valuable insights, by providing new constraints on COS fluxes and by identifying missing processes that could be integrated in the COS models.

In the vegetation COS uptake model from Berry et al. (2013), the internal conductance to COS that represents the mesophyll conductance and carbonic anhydrase (CA) activity, is assumed proportional to the maximum carboxylation rate of Rubisco (V_{cmax}). This assumption implies that the temperature response of Rubisco is applied to CA when simulating vegetation COS uptake. In the SiB4 LSM, Cho et al. (2023) recently introduced a novel function to describe the temperature response of CA, incorporating its specific temperature optimum. This new approach still expresses the internal conductance to COS as proportional to the V_{cmax} at 25°C ($V_{cmax,25}$) of Rubisco, but it replaces Rubisco's temperature response function with that of CA. Optimizing model parameters related to the stomatal and internal conductance using COS flux observations at two sites, the Hyytiälä and Harvard forests, Cho et al. (2023) found that CA has lower temperature optimums than Rubisco. Using the optimized parameter values to simulate global vegetation COS flux in SiB4 decreased the COS uptake in regions where air temperatures exceed 25°C, predominantly in tropical areas, and increased the uptake in regions with temperatures below 25°C. In ORCHIDEE, this new CA temperature response has been implemented and its impact on the simulated vegetation COS fluxes is discussed in Section 4.2, as well as the potential distinction between mesophyll conductance and the enzyme activity in the vegetation COS model, as it is the case for CO_2 in ORCHIDEE (Abadie et al., 2023).

In ORCHIDEE, the contribution of understory vegetation to vegetation COS uptake is not accounted for due to the lack of representation of understory vegetation. This could be addressed by using LSMs that integrate a finer representation of vegetation structure and dynamics, such as CLM-FATES (Fisher et al., 2015).

In addition to leaf COS absorption by vascular plants, non-vascular plants and lichens also contribute to ecosystem COS exchanges. These processes are not yet represented in LSMs. Bryophytes and lichens can absorb COS even in the dark, potentially making a significant contribution depending on moisture conditions (Gimeno et al., 2017; Sun et al., 2018; Rastogi et al., 2018). These organisms have also been reported to emit COS driven by temperature (Gimeno et al., 2017). Understanding the mechanisms related to bryophyte and lichen COS fluxes and integrating them into LSMs could help scale up their contribution and assess their importance at the global scale.

Some vascular plants also have the potential to emit COS (Bloem et al., 2012; Geng & Mu, 2006; Maseyk et al., 2014). Experimental studies have reported COS production by agricultural crops, such as wheat fields at specific growth stages (Maseyk et al., 2014), or following fungal infection in oilseed rape (Bloem et al., 2012). The ecosystem COS fluxes simulated in ORCHIDEE have been used to investigate COS emission in agroecosystems in Central France (Belviso et al., 2022a), as presented in Section 5.2.1. This also highlights the potential role of plant-fungi interactions in the

phyllosphere for ecosystem COS exchanges, which requires future research and is not yet represented in LSMs.

Regarding the representation of soil COS exchanges, two mechanistic soil COS flux models have been proposed by Sun et al. (2015) and Ogée et al. (2016). For now, the model from Ogée et al. (2016) has been preferably implemented in LSMs such as SiB4 and ORCHIDEE (Kooijmans et al., 2021; Abadie et al., 2022) due to its lower complexity. Sun et al. (2015) requires resolving COS diffusion, consumption, and emission by discretizing the soil column into several layers, whereas Ogée et al. (2016) derived a simplified analytical solution assuming a soil column with uniform temperature, soil moisture, and porosity and steady-state conditions. Therefore, implementing the model from Sun et al. (2015) in LSMs, which would alter the COS concentration profile in the soil column, could yield different estimates of soil COS contribution at the global scale and help evaluate the uncertainty related to the choice of soil COS flux representation.

In addition to the existing simple empirical soil COS uptake models (Kettle et al., 2002; Berry et al., 2013; Launois et al., 2015), Whelan et al. (2022) proposed a new empirical soil COS exchange model including an uptake and a production term. This model is based on biome-specific response curves to describe soil COS fluxes parametrized from field and upscaled lab incubation experiments, using only information on biome type, soil moisture, and surface temperature as input data. Contrary to the typically used exponential oxic soil COS production in response to soil temperature, Whelan et al. (2022) considered a logistic function to define a maximum soil COS production for each biome. While this could refine soil COS production implemented in LSMs, its impact on simulated soil COS fluxes might be limited, as its significance is more pronounced under conditions not typically found in vegetated ecosystems. New estimates of anoxic soil COS production are also provided based on an upscaling of a function of soil temperature fitted on *in situ* salt marsh measurements. This new wetland estimate, representative of high COS emissions from saline environments, yields a COS production four times higher than the one estimated in ORCHIDEE, which better represents freshwater environments (Abadie et al., 2022). This highlights a possible future improvement of wetland COS flux representation in LSMs by distinguishing between saline and freshwater environments.

Regarding processes of soil COS flux that have been recently investigated and are not yet accounted for in LSMs, Kitz et al. (2023) highlighted the contributions of living roots and the rhizosphere to soil COS exchanges. Their experimental study with young beech trees found that living roots have a contrasting impact on soil COS fluxes throughout the season. During the growing season, the presence of living roots significantly reduced soil COS emissions compared to bare soil, attributed to the CA enzyme in roots that consumes COS. However, during periods of vegetative inactivity, living roots increased soil COS emissions. The variation in soil COS fluxes due to living roots may also be influenced by changes in the rhizosphere microbial community, which can be affected by rhizodeposition. This study calls for further investigation into the underlying mechanisms of living root contribution to soil COS exchanges. Understanding these processes could lead to their integration in LSMs that include soil COS models. Additionally, the study used an artificial, highly nutrient-rich soil, and future research should focus on assessing the impact of living roots on soil COS exchanges in natural soils to enhance the applicability of these findings.

Finally, several studies have highlighted the impact of using constant versus variable atmospheric COS concentrations on simulated ecosystem COS exchanges (Kooijmans et al., 2021; Abadie et al., 2022). Consequently, spatially and temporally varying atmospheric COS concentrations are used as input data in LSMs. For example, SiB4 uses 3-hourly COS concentrations at a 4°x6° spatial resolution (Kooijmans et al., 2021), and ORCHIDEE uses 3-hourly concentrations at a 3.75°x1.9° spatial resolution (Abadie et al., 2023). This approach allows for accounting for the impact of high-resolution variations in COS concentrations on simulated soil and vegetation COS fluxes. However, these approaches do not consider the influence of biospheric COS fluxes on atmospheric COS concentrations. To address this limitation, future simulations of ecosystem COS fluxes could benefit from coupling LSMs with atmospheric circulation models, providing a comprehensive view of the interactions and feedback between atmospheric COS concentrations and surface fluxes.

4 USING COS AND CO_2 FLUX OBSERVATIONS TO CONSTRAIN GPP AND PLANT TRANSPIRATION IN ORCHIDEE

4.1 Assimilating COS flux observations in ORCHIDEE

The implementation of vegetation and soil COS models in ORCHIDEE (Sections 3.1 and 3.2) enables to simulate ecosystem COS flux, allowing the assimilation of eddy covariance COS flux data, which has been collected over several years at a few sites (Kohonen et al., 2020; Vesala et al., 2022; Wehr et al., 2017), to constrain ORCHIDEE parameters. Although atmospheric COS concentration measurements are available at more locations than eddy covariance COS flux data (Section 2.2), these concentrations reflect the influence of all COS budget components. To optimize ORCHIDEE parameters using atmospheric COS concentrations, a comprehensive understanding of all other components influencing these concentrations at the measurement sites is required. In contrast, assimilating eddy covariance COS flux data focuses on the biospheric contribution to the COS budget, providing a stronger constraint on parameters that also determine GPP and LE. Indeed, COS flux data has been found to provide information on stomatal diffusion (Berkelhammer et al., 2020; Kooijmans et al., 2017; Wehr et al., 2017). Therefore, assimilating biospheric COS flux data can offer a new constraint on the stomatal conductance related parameters in ORCHIDEE, informing both GPP and plant transpiration, which are coupled through stomatal diffusion.

Assimilating eddy covariance COS flux data also enables to calibrate the biospheric COS models implemented in LSMs (Kooijmans et al., 2021; Maignan et al., 2021; Abadie et al., 2022; Chen et al., 2023). Several atmospheric inversion modeling studies using atmospheric COS concentration observations have identified a missing biospheric COS sink in Northern high latitudes (Ma et al., 2021; Remaud et al., 2022). Then, Vesala et al. (2022) developed a parametric model of vegetation COS fluxes calibrated against eddy covariance COS fluxes at the Hyytiälä boreal forest, which also indicated increased COS uptake in high latitude boreal evergreen needleleaf biomes compared to the uptake simulated in the SiB4 LSM. Therefore, calibrating the COS model parameters in ORCHIDEE through data assimilation techniques allows us to constrain the simulated biospheric COS fluxes and assess whether this approach can reduce the missing COS sink in Northern high latitudes, bringing us closer to closing the global COS budget (Section 1.3.3).

Consequently, eddy covariance COS fluxes from the Hyytiälä site, which constitute the longest time series of ecosystem COS flux measurements, have been assimilated in ORCHIDEE. This study is the first to assimilate COS flux data in ORCHIDEE. The data assimilation approach, along with the evaluation of the optimized ecosystem COS flux, GPP, and LE, and their upscaling over the entire boreal evergreen needleleaf biome, are presented in my second publication as the lead author in the following section.

4.2 Additional constraint on GPP and plant transpiration from COS flux observations for boreal forests

JGR Biogeosciences

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 Abadie[1], Fabienne Maignan[1], Marine Remaud[1], Kukka-Maaria Kohonen[2,3], Wu Sun[4], Linda Kooijmans[5], Timo Vesala[2,6], Ulli Seibt[7], Nina Raoult[1], Vladislav Bastrikov[1], Sauveur Belviso[1], and Philippe Peylin[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France, [2]Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland, [3]Department of Environmental Systems Science, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland, [4]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA, [5]Meteorology and Air Quality, Wageningen University and Research Centre, Wageningen, The Netherlands, [6]Institute for Atmospheric and Earth System Research/Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland, [7]Department 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 CO_2 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 Hyytiälä 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 Hyytiälä 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 Hyytiälä 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 (CO_2) 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 CO_2 uptake because COS and CO_2 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 CO_2 uptake measured at Hyytiälä boreal forest. We find that this joint assimilation improves the simulated plant CO_2 uptake, as well as transpiration because of the strong link between COS, CO_2 and H_2O fluxes through stomatal diffusion. Scaling up this assimilation framework to evergreen needleleaf boreal forests, we find that assimilating ecosystem COS flux and plant CO_2 uptake data increases the vegetation COS uptake for this biome, but not plant CO_2 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 CO_2 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

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 CO_2 during photosynthesis. Inside the leaves, COS is presumed to be totally hydrolyzed by the carbonic anhydrase (CA), an enzyme also involved in CO_2 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 CO_2 measurements. In addition, COS helps to constrain stomatal diffusion, which determines the coupling between CO_2 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 CO_2 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 CO_2 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 (CS_2) (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 CO_2 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 parameterization 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 CO_2 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 CO_2 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

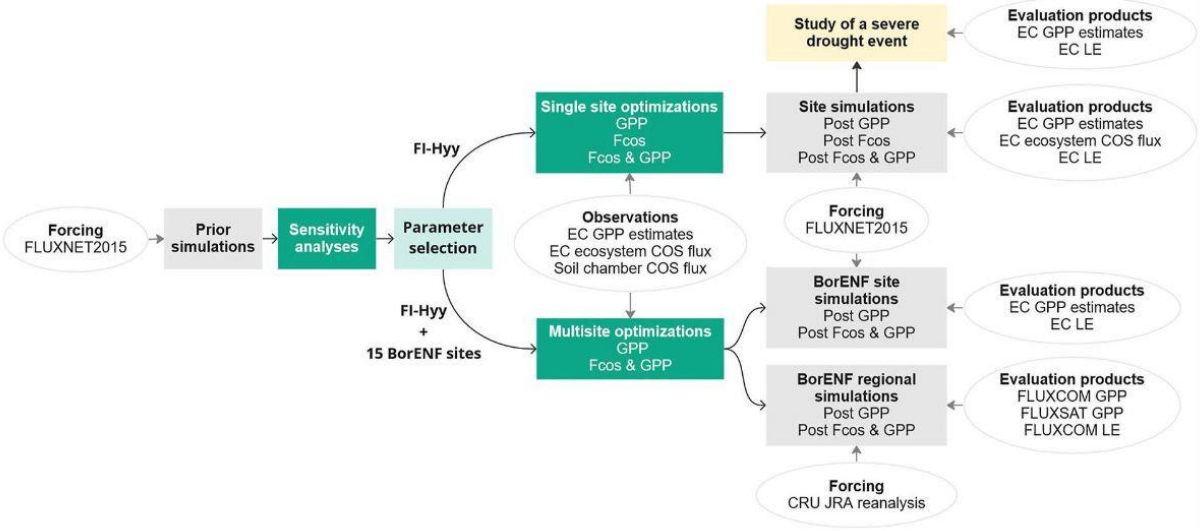

Figure 1. Overview of the workflow, showing the input data and main steps of this study. Actions are represented in boxes, whereas input data used as forcing, observations, and evaluation products, are in circles. EC, eddy covariance; FI-Hyy, Hyytiälä forest in Finland; BorENF, Boreal evergreen needleleaf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

Figure 2. Mean annual air temperatures range from −1.4° 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 Hyytiälä 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

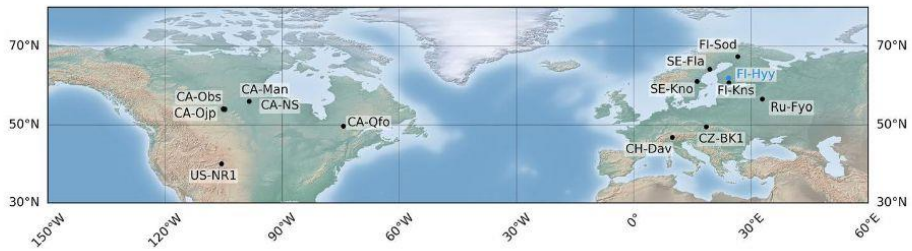

Figure 2. Location of the 16 studied sites. The CA-NS label includes CA-NS1, CA-NS2, and CA-NS5. FI-Hyy is represented in blue as it is the only site with COS observations, while the other sites are represented in black. The background map corresponds to the "shadedrelief" map (http://www.shadedrelief.com) 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 wintertime. 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),

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

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) (μmol m^{-2} s^{-1}), R_d the day respiration (μmol m^{-2} s^{-1}), C_i the intercellular CO_2 partial pressure (μmol m^{-2} s^{-1}), C_{i*} the base CO_2 compensation point in the absence of R_d (μmol m^{-2} s^{-1}), and f_{VPD} the function describing the effect of VPD, defined as,

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

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).

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

with $F_{\mathrm{COS,veg}}$ the vegetation COS uptake flux (pmol m^{-2} s^{-1}), [COS] the atmospheric COS concentration (ppt), $g_{b,\mathrm{COS}}$, $g_{s,\mathrm{COS}}$, and $g_{i,\mathrm{COS}}$ the laminar boundary layer, stomatal, and internal COS conductances (mol m^{-2} s^{-1}), respectively. The internal conductance $g_{i,\mathrm{COS}}$ 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 (μmol m^{-2} s^{-1}) (Badger & Dean Price, 1994; Berry et al., 2013). Therefore, $g_{i,\mathrm{COS}}$ is expressed as proportional to Vmax,

$$g_{i,\mathrm{COS}} = \alpha \cdot V_{\mathrm{max}} \tag{4}$$

where α is a parameter, the value of which depends on the photosynthetic pathway. Typical values of α are 0.0012 mol μmol^{-1} for C3 and 0.013 mol μmol^{-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 Ogee et al. (2016),

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

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

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

where $F_{\mathrm{COS,anoxicsoil}}$ is the COS flux from anoxic soils (mol m^{-2} s^{-1}), P_{ref} is the reference production term (mol m^{-2} 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),

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

with $\boldsymbol{x_b}$ the a priori vector of parameters, $\boldsymbol{x}$ the optimized vectors of parameters, $\boldsymbol{y}$ the observations, $H(\boldsymbol{x})$ the corresponding model outputs, and $\boldsymbol{R}$ and $\boldsymbol{B}$ the prior error covariance matrices for the observations (including measurement and model errors) and the parameters, respectively. $\boldsymbol{R}$ and $\boldsymbol{B}$ 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 $\boldsymbol{R}$ 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 $\boldsymbol{B}$ 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 CO_2 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	Optimized variables: GPP	Optimized variables: Ecosystem COS flux and GPP
α	Proportionality constant used in the calculation of the internal conductance to COS	0.001200 [0.000720, 0.001680]	mol µmol^{-1}	x		x
acclim$_{Jmax,a}$	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]	J K^{-1} mol^{-1}	x	x	x
acclim$_{Vcmax,a}$	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]	J K^{-1} mol^{-1}	x	x	x
SLA	Specific leaf area, involved in the calculation of leaf biomass	0.00926 [0.00695, 0.01157	m^2 gC^{-1}	x	x	x
Vcmax$_{25}$	Maximum rate of Rubisco activity-limited carboxylation at 25°C	45. [34., 56.]	µmol m^{-2} 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
gb$_{ref}$	Leaf bulk boundary layer conductance	0.04 [0.03, 0.05]	m s^{-1}	x	x	x
Tphoto$_{min}$	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]	m^2 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 ±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_{COS} ½	Post GPP & F_{COS} ⅓
GPP (10 BorENF sites)	1	½	⅔
Ecosystem COS fluxes (FI-Hyy)	0	½	⅓

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 pmol m^{-2} 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 ± 1.2 pmol m^{-2} 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 pmol m^{-2} 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_{COS} ½"). 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 ⅓ (scenario "Post GPP & F_{COS} ⅓"; 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.

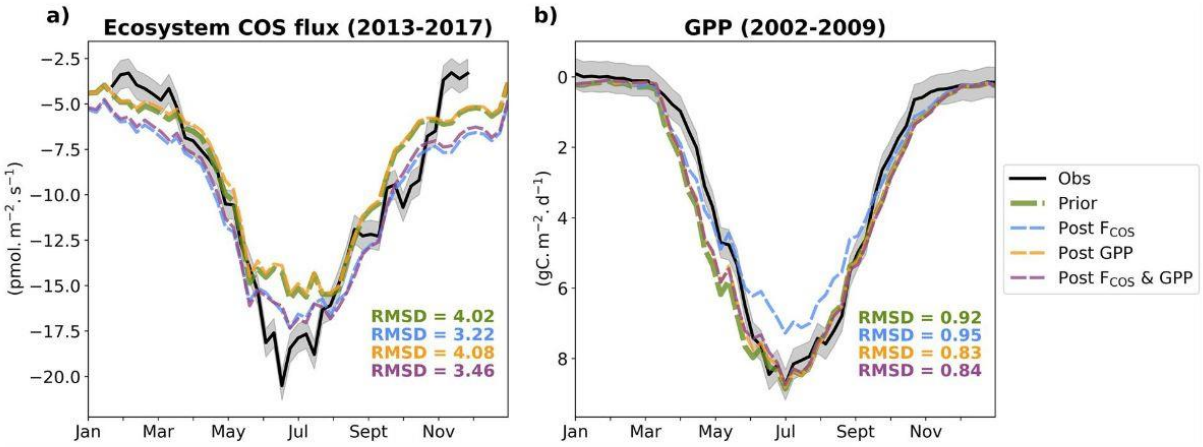

Figure 3. Mean seasonal cycles of the weekly ecosystem COS flux (pmol m^2 s^{-1}) (a) and GPP (gC m^2 d^{-1}) (b) at FI-Hyy. The weekly averages are computed considering both nighttime and daytime data for the eddy covariance observations and simulated fluxes. Note that the *y*-axis has been inverted 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 pmol m^2 s^{-1} over the period against −8.8 pmol m^2 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 pmol m^2 s^{-1}), but is slightly improved by the optimizations that assimilate COS flux observations (amplitudes near 12.5 pmol m^2 s^{-1} against 11.7 pmol m^2 s^{-1} after the GPP-only assimilation and 11.9 pmol m^2 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 gC m^2 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 gC m^2 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 pmol m^2 s^{-1} in 2015, −21.9 pmol m^2 s^{-1} in 2016) while it was underestimated in the prior simulation (−17.6 pmol m^2 s^{-1} in 2015, −19.5 pmol m^2 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

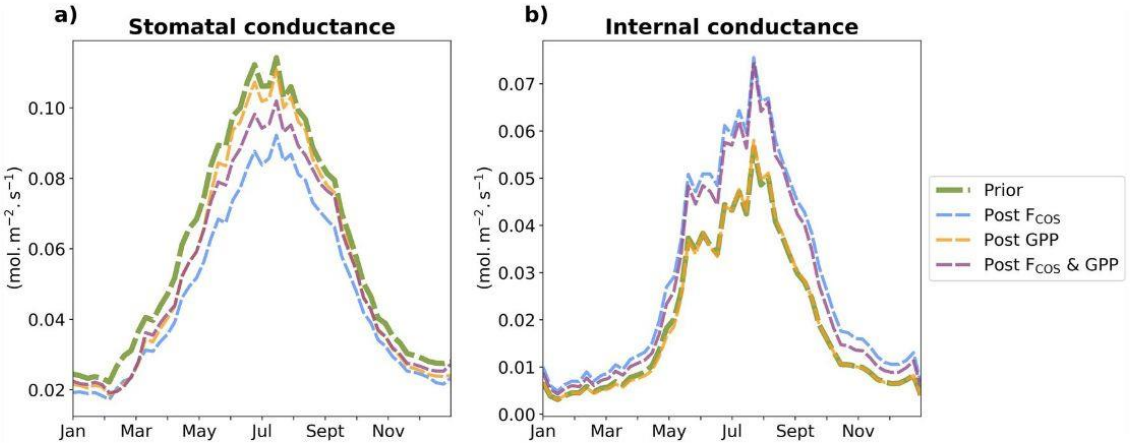

Figure 4. Mean seasonal cycles of the weekly modeled stomatal (a), and internal (b) conductances to COS (mol m^2 s^{-1}) at FI-Hyy between 2013 and 2017. Note that the posterior seasonal cycles of the internal conductance in the "Prior" and "Post F_{COS}" simulations overlap with each other.

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 $Vcmax_{25}$ and α parameters (Figure S3 in Supporting Information S1) explains this higher internal conductance as it relates to Vcmax 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 hight percentage of reduction in uncertainties are $Vcmax_{25}$ (66%–85% reduction) and $acclim_{Vcmax,a}$ (the intercept of the linear regression representing the acclimation to temperature of the entropy term for Vcmax) (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 gb_{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 $Tphoto_{min}$ (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_{COS} & 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_{COS} & GPP") to 45% (8.37

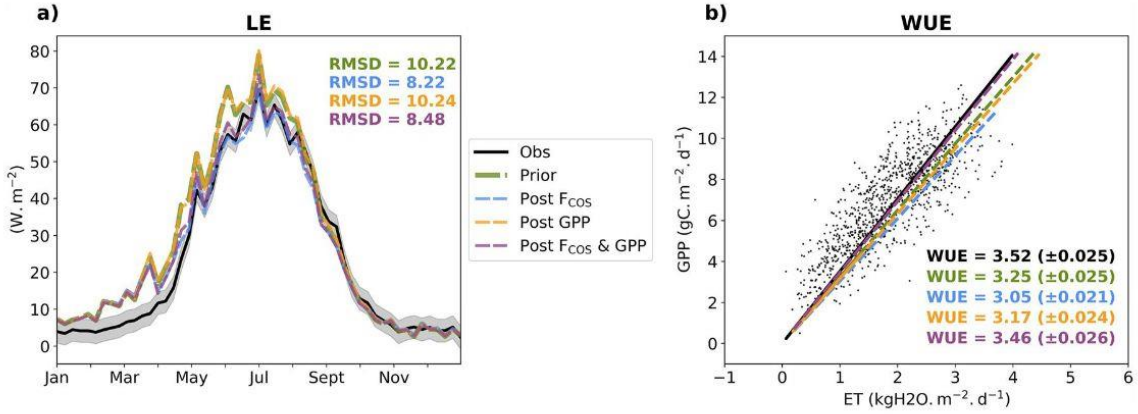

Figure 5. Mean seasonal cycle of the weekly observed and modeled LE (W m^{-2}) over 2002–2009 (a) and summer (May to August) WUE (gC kgH_2O^{-1}) (b) at FI-Hyy. For LE, the RMSD values computed over the weekly fluxes are given for the prior and the post-optimization simulations in the corresponding color (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 (±0.025) gC kgH_2O^{-1} and 3.17 (±0.024) gC kgH_2O^{-1}, respectively, which are below the observationally derived WUE of 3.52 (±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 (±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 (±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

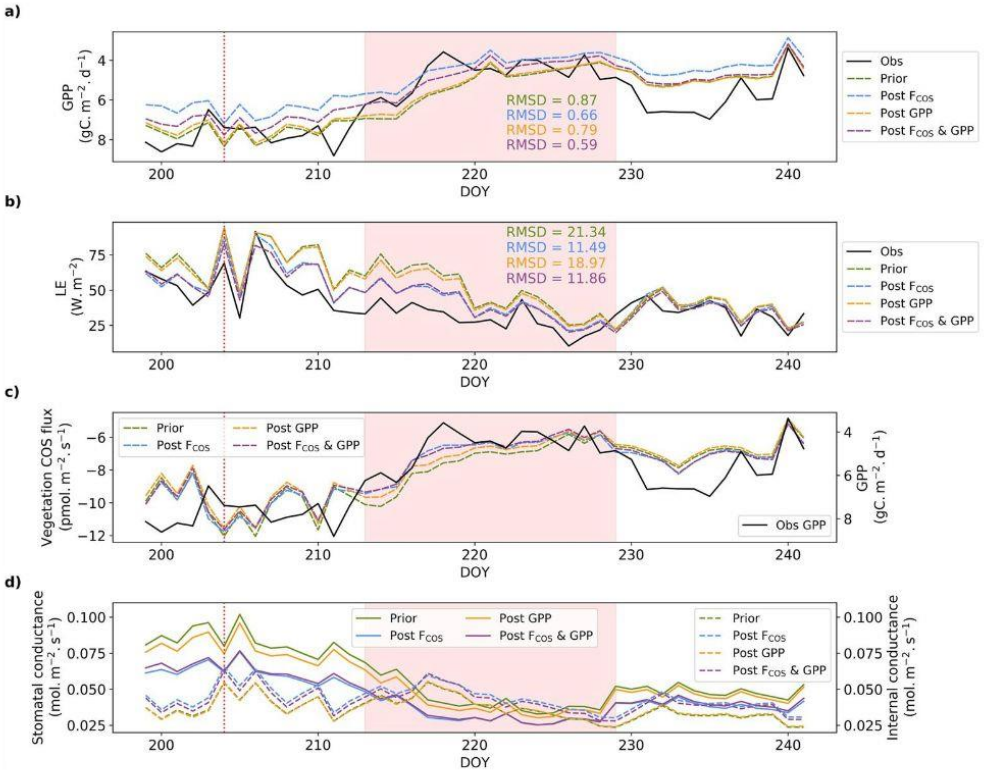

Figure 6. Daily averages of observed and modeled GPP (gC m^2 d^{-1}) (a) and LE (W m^{-2}) (b) at FI-Hyy during a severe drought event in 2006. Note that the *y*-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) to facilitate the comparison with vegetation COS uptake. The comparison between the modeled vegetation COS fluxes (pmol m^2 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 (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

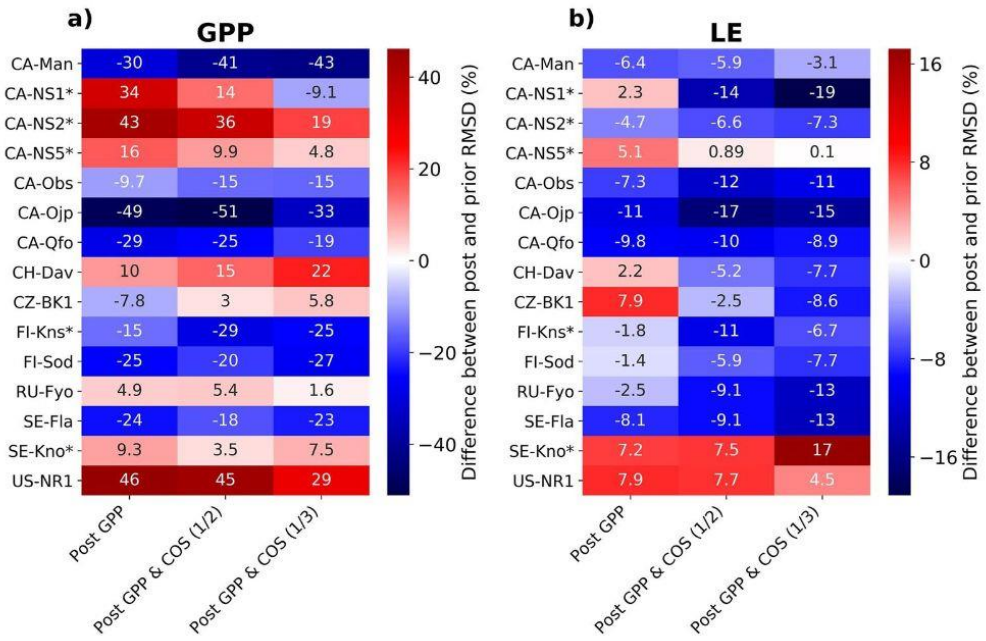

Figure 7. Relative difference (%) in RMSDs of weekly GPP (a) and LE (b) between prior simulations and after multi-site optimizations at each of the 15 validation sites. A negative difference (in blue) means a decrease in RMSD after multi-site optimization. In all multi-site optimization scenarios, GPP 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 assimilaton 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_{COS} ½" and "Post GPP & F_{COS} ⅓") 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: $acclim_{Jmax,a}$, 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

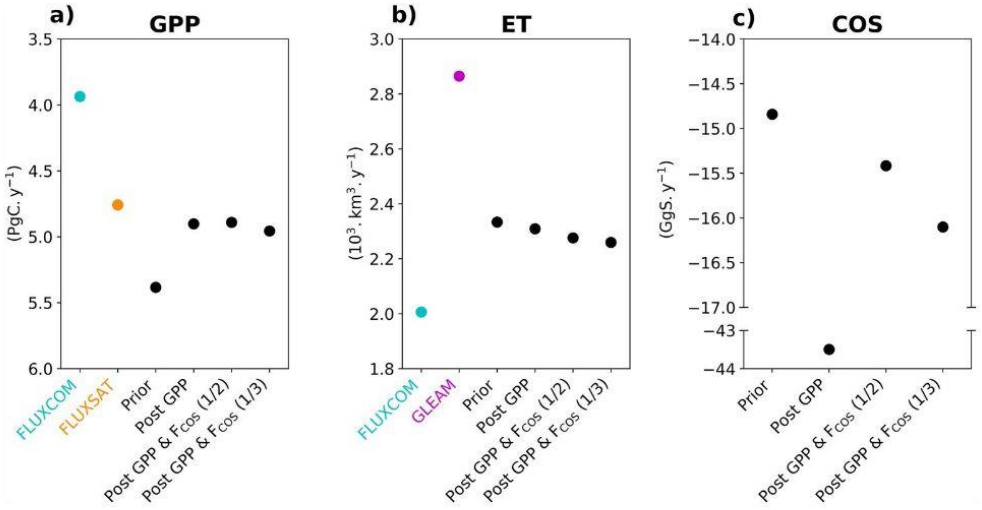

Figure 8. GPP (a), ET (b), and vegetation COS fluxes (c) in the prior simulation and each of the multi-site optimization scenarios computed over land grid cells that have a fractional coverage of BorENF higher than 40%, averaged between 2006 and 2015. For GPP and ET, the fluxes from the evaluation products (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 ½ or ⅓, 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} ½" and "Post GPP & F_{COS} ⅓") 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 ½ for COS observations and by 7% with a weight of ⅓ for COS observations.

For both GPP and LE, assimilating GPP and COS observations using a weight of ⅓ for the latter yields post-optimization GPP and LE that are closer to the observations than using a weight of ½ for COS observations in the assimilation. This can be due to the larger decrease in A1 and increase in $Vcmax_{25}$ for the joint assimilation with a weight of ⅓ for COS data compared to the assimilation using a weight of ½ 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 10^3 km^3 yr^{-1}) and GLEAM (2.86 10^3 km^3 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 ½") and 9% ("Post F_{COS} & GPP ⅓") compared to the prior. Surprisingly, the optimization that assimilates only GPP data leads to a substantial increase in vegetation COS uptake to 43.50 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 ± 0.008 gC kgH_2O^{-1}) compared with FLUXCOM WUE (2.0 ± 0.007 gC kgH_2O^{-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 ± 0.009 gC kgH_2O^{-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., $acclim_{Jmax,a}$).

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 gC m^2 d^{-1}, which is lower than the FLUXNET GPP of 6.6 gC m^2 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 $Vcmax_{25}$ and α. When both $Vcmax_{25}$ and α need to be increased, the optimization that assimilates only COS observations increases α more than $Vcmax_{25}$ ($Vcmax_{25}$ = 46.6 µmol m^2 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 $Vcmax_{25}$ more than α ($Vcmax_{25}$ = 54 µmol m^2 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 $Vcmax_{25}$ needed to reproduce GPP cannot compensate for the decrease in the stomatal conductance resulting from the decrease in A1 (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 (gb_{ref}) and in the parameter that represents thermal acclimation of Vcmax ($acclim_{Vcmax,a}$), 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 $acclim_{Vcmax,a}$, which affects the internal conductance to COS through Vcmax 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 ⅓"). 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 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 (photosynthetically active radiation, vapor pressure deficit, air temperature, and leaf area index) to simulate vegetation

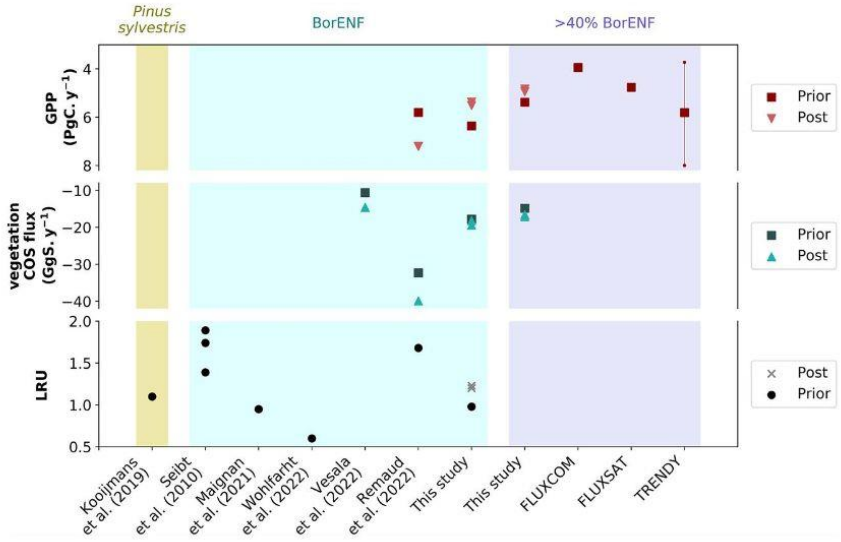

Figure 9. Synthesis of LRU values and vegetation COS and GPP fluxes from different studies focusing on boreal biomes. Note that the *y*-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) so that both GPP and vegetation COS fluxes representing uptakes 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% (light 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 C_3 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 ⅓") 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 µmol m^{-2} 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 ½" and "Post F_{COS} & GPP ⅓") for all C_3 PFTs as the values of these two parameters were initially defined according to the photosynthetic pathways (C_3 or C_4 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 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 (Vcmax) 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 $Vcmax_{25}$ 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 $g_{i,COS}$ (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 $g_{i,COS}$ (RMSD of 3.6 pmol m^2 s^{-1} against 3.2 pmol m^2 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 CO_2 assimilation. This would imply to distinguish between the mesophyll conductance and CA activity in $g_{i,COS}$. 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 subniveal 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 flasks 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, CO_2, and H_2O 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 H_2O, CO_2, 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 CO_2 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 Hyytiälä 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 Hyytiälä 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), GreenGrass, 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 Hyytiälä 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, E., 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

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 CO_2: 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 CO_2 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., Doughty, 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 CO_2 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

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., Ciais, 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 CO_2. *Plant Physiology*, *171*(1), 280–293. https://doi.org/10.1104/pp.15.01990

Farquhar, G. D., von CaemmererCaemmerer, S. S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 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 CO_2 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 CO_2 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 CO_2 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 CO_2 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., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(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 CO_2 flux and its inter-annual 2 variability.

Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K. M., Commane, 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/jsme2.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 CO_2. *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

Ogée, 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

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 (CO_2). *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 CO_2 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 CO_2 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 CO_2 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., Foltýnová, 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-15-3625-2018ï

Wohlfahrt, G., Brilli, 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

4.3 Recent advances in COS flux data assimilation in LSMs and prospective future experiments

Since the above DA study of ecosystem COS fluxes from the Hyytiälä forest in ORCHIDEE, other studies have implemented a representation of ecosystem COS fluxes in LSMs and assimilated COS flux data to refine the simulated biospheric COS fluxes, as well as GPP, LE, sensible heat flux, and soil moisture (Zhu et al., 2024a, 2024b).

Zhu et al. (2024a) developed the adjoint-based Nanjing University Carbon Assimilation System (NUCAS) to assimilate multiple COS flux observations in the Boreal Ecosystem Productivity Simulator (BEPS). A representation of biospheric COS fluxes has been integrated in BEPS, with vegetation COS uptake simulated using the same mechanistic approach as implemented in ORCHIDEE and SiB4, as proposed by Berry et al. (2013). However, unlike ORCHIDEE and SiB4, BEPS does not base soil COS exchanges on the mechanistic model from Ogée et al. (2016). Instead, BEPS adopts a simplified approach to simulate soil COS exchanges (Whelan et al., 2016). This approach models soil COS exchanges as the sum of an abiotic component, represented by an exponential function of soil temperature, and a biotic component, which depends on soil temperature and moisture.
COS flux data from seven sites (Neustift grassland in Austria, Sorø forest in Denmark, Las Majadas del Tietar savanna in Spain, Hyytiälä forest in Finland, a soybean site in Italy, Harvard forest in the United States, and Wind River experimental forest in the United States) were assimilated to optimize 76 parameters in BEPS using a gradient-based optimization approach (BFGS). Among these, five parameters related to photosynthesis, transpiration, and soil water transport were efficiently optimized following COS flux DA. Single-site COS flux DA reduced the mean GPP RMSD from 8.22 μmol m^{-2} s^{-1} to 6.38 μmol m^{-2} s^{-1}. Notably, this improvement in simulated GPP was achieved by assimilating independent COS flux data, without incorporating any GPP data. However, results varied across sites, with some degradation of the simulated GPP at low-stature vegetation sites. Consistent with the COS flux DA in ORCHIDEE, the simulated LE also showed improvement with a mean RMSD reduction from 94.69 W m^{-2} to 79.69 W m^{-2}, and a slight improvement of about 5% in the simulated sensible heat flux. Additionally, COS flux DA helped correct the prior overestimation of the soil moisture decline rate in summer, although significant errors remained due to a weaker link between ecosystem COS flux and soil moisture parameters compared to photosynthesis-related parameters.

Building on this work, Zhu et al. (2024b) continued exploring the potential of COS flux assimilation to constrain the simulated GPP in BEPS. In this second study, they used COS flux observations from the same seven sites as Zhu et al. (2024a). However, instead of employing a gradient-based approach, they used a Monte Carlo-based parameter calibration method to optimize nine parameters associated with photosynthesis, soil hydrology, stomatal diffusion, and energy balance. The results indicated that both the simulated ecosystem COS flux and GPP were most sensitive to the maximum Rubisco carboxylation rate at 25°C and the leaf nitrogen content, while less sensitive to soil hydrology parameters as noted by Zhu et al. (2024a). The Ball-Berry model parameter for stomatal conductance intercept was critical for COS flux simulation but had a

minimal impact on simulated GPP. Across all sites, the mean RMSD reduction in ecosystem COS fluxes was 32.09%, and a GPP RMSD reduction of 35.42% was achieved following COS flux DA.

These two studies assimilate only COS flux observations to constrain the carbon, water, or energy fluxes in BEPS. However, Zhu et al. (2024a) noted that relying solely on COS flux data can result in equifinality issues (Beven, 1993) after the optimization, where different combinations of optimized parameter values yield similar reductions in the cost function. They also emphasize that assimilating only COS flux data can degrade the simulated GPP at some sites and lead to overshooting of some photosynthesis-related parameters. Consequently, Zhu et al. (2024a) suggest that it may be necessary to combine COS flux data with other observational data streams during DA to better constrain the carbon, water, and energy fluxes. This finding aligns with the results from COS flux DA in ORCHIDEE, where the joint assimilation of COS flux and GPP data improved both simulated variables, while using only COS flux data degraded the simulated GPP at the Hyytiälä site.

Despite growing interest in using COS flux observations to constrain stomatal conductance, photosynthesis, and water and energy fluxes (Abadie et al., 2023; Cho et al., 2023; Chen et al., 2023; Zhu et al., 2024a, 2024b), all studies have highlighted the limitations of this approach due to the scarcity of COS flux observations. To address this challenge, a second constraint could be added by performing a joint assimilation of two data streams: ecosystem COS flux and atmospheric COS concentration data, to directly optimize model parameters. Indeed, COS concentration observations from the NOAA network and other measurement towers, such as the GIF site near Paris in France, offer additional information at a large spatial scale, including around 15 sites with some providing data spanning more than 20 years, which complements the *in situ* COS flux data (Section 2.2). An example of such an approach, which assimilates both surface fluxes and atmospheric concentration data, has been conducted in the ORCHIDEE-LMDZ LSM-atmosphere model by Bacour et al. (2023). This study used *in situ* NEE and LE estimates, NDVI satellite retrievals, and CO_2 concentrations. To our knowledge, DA of COS flux for parameter optimization (Abadie et al., 2023; Cho et al., 2023; Chen et al., 2023; Zhu et al., 2024a, 2024b) or DA of COS concentrations for surface flux estimation through atmospheric inversions (Ma et al., 2021; Remaud et al., 2022) have been conducted in LSMs, but no joint assimilation has yet been performed. This approach could exploit the complementarities between the information embedded in ecosystem COS fluxes, which provide a more direct constraint on vegetation activity, and atmospheric COS concentrations, which account for all components of the global COS budget.

5 USING ATMOSPHERIC COS CONCENTRATIONS TO CONSTRAIN THE GLOBAL COS AND CO_2 BUDGETS

5.1 Overview of the studies using atmospheric COS concentrations

While the primary focus of my PhD work was using COS flux observations to constrain simulated ecosystem COS fluxes, GPP, and plant transpiration in ORCHIDEE, I also participated in several studies that used atmospheric COS and CO_2 concentration data to inform surface COS fluxes and terrestrial vegetation CO_2 exchanges. Investigating atmospheric COS concentrations can provide an additional constraint on the COS fluxes simulated in ORCHIDEE, supplementing parameter optimization and process evaluation against *in situ* COS flux data. This approach also allows for the assessment of other components of the global COS budget beyond the vegetation and soil COS fluxes.

The LSCE team studying COS benefited from a complementary blend of expertise, resulting in several publications involving both experimental and modeling approaches. Notably, Sauveur Belviso (permanent researcher) led the experimental work on COS concentrations, monitoring atmospheric COS at the GIF and Trainou stations in France, and conducting COS concentration measurements in agroecosystems near LSCE. Marine Remaud (postdoctoral researcher) contributed her expertise in atmospheric transport modeling with the LMDZ model, allowing to transport the soil and vegetation COS fluxes simulated in ORCHIDEE, along with other components of the global COS budget to simulate atmospheric COS concentrations. She also performed atmospheric inversion modeling to optimize and inform surface COS and CO_2 fluxes at the global scale using atmospheric COS and CO_2 concentration measurements. Then, Antoine Berchet and Isabelle Pison (both permanent researchers) used the Flexible Particle model (FLEXPART) to assess surface COS fluxes, including the ecosystem COS fluxes simulated in ORCHIDEE. Contrary to the LMDZ atmospheric circulation model, the FLEXPART Lagrangian particle dispersion model tracks the trajectories of particles as they move through the atmosphere (Pisso et al., 2019). FLEXPART can be operated both forward and backward in time. In forward mode, particles are released from one or more sources to determine concentrations. Backward simulations are employed to examine the history of air parcels impacting a specific location, such as an atmospheric measurement site, which serves as a receptor.

Consequently, the simulated vegetation and soil COS fluxes in ORCHIDEE have been used in multiple studies ranging from local to global scales. The publications resulting from these collaborations are summarized in Figure 9 below.

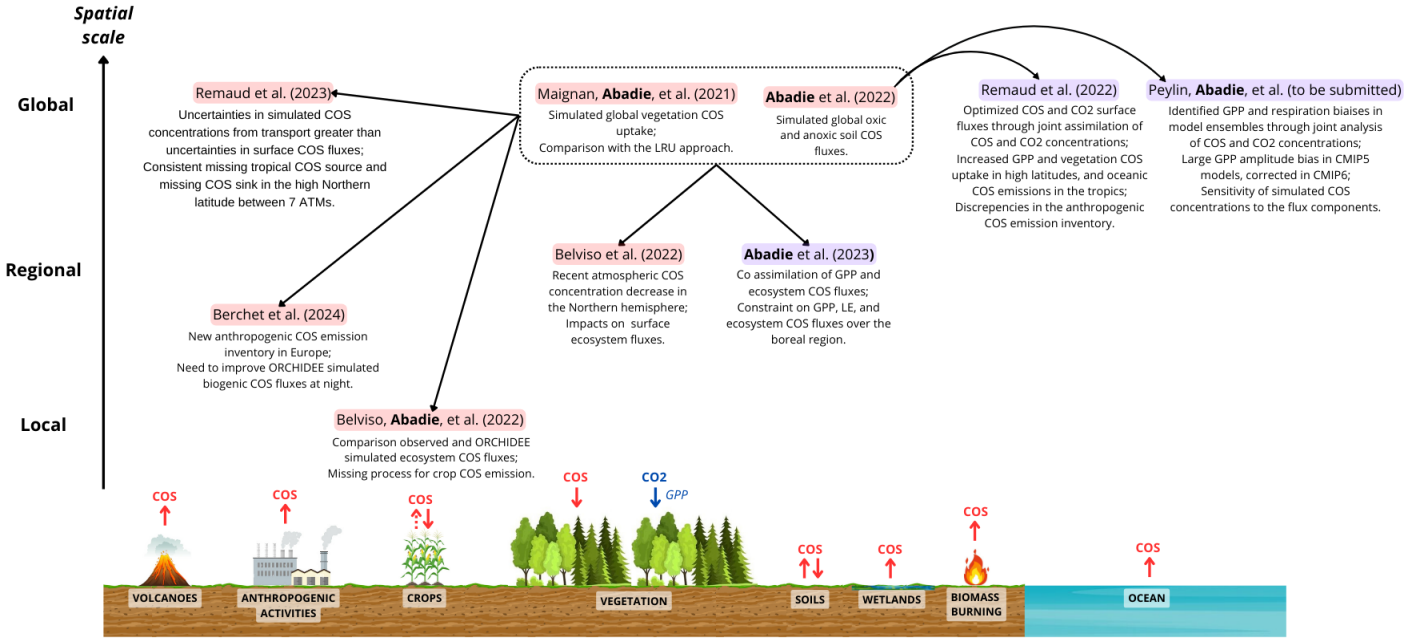

Figure 9: Overview of the studies involving COS surface fluxes I contributed to. The main COS surface flux components of the global COS budget are represented (in red), while only GPP is shown for CO_2 (in blue). Studies highlighted in light red focus on COS while studies highlighted in light purple focus on both COS and CO_2.

5.2 Applications for the regional and global COS budget

5.2.1 Agroecosystem contribution to COS emissions

While vegetation COS flux is considered to be a one-way flux from the atmosphere to the leaf interior, where it is completely hydrolyzed by CA (Stimler et al., 2010), some field and controlled environment experiments have highlighted the potential of vascular plants to emit COS (Bloem et al., 2012; Geng & Mu, 2006; Maseyk et al., 2014). However, this emission process is not represented in LSMs, and there has been no large-scale estimation of vegetation's role in COS production. As a result, the contribution of vegetation to COS emissions is not accounted for in the global COS budget.

Therefore, in Belviso et al. (2022a) (Appendix 9.2), the vegetation and soil COS fluxes simulated in ORCHIDEE were evaluated against field observations at two agroecosystems in central France. My contributions to this work included conducting the ORCHIDEE simulations for these sites and participating in the analysis and presentation of the results, specifically comparing the observed and simulated fluxes. This study investigated potential COS emissions from croplands that are not yet considered in the model. Three independent approaches were used to compare the model with observations.

First, monitoring nighttime ratios of vertical mole fraction gradients of COS and CO_2 during the summer months in 2019 and 2020 at a rural tall tower site near Orléans suggested a weakening of nocturnal net COS uptake during peak growing season, recovering in August. In contrast, the simulated nocturnal COS and CO_2 ecosystem fluxes in ORCHIDEE maintained a constant proportional relationship throughout the summer. This discrepancy indicated that the model's nocturnal vegetation COS uptake driven by residual nighttime conductance and soil COS exchanges was insufficient to reproduce the observed patterns at this rural site.

Secondly, the radon tracer method was applied at the GIF site near Paris from 2014 to 2021. In this approach, ^{222}Radon (^{222}Rn), a tracer emitted by soil, is used to estimate nocturnal ecosystem COS fluxes based on measurements of ^{222}Rn and COS mixing ratios (Belviso et al., 2020). The results indicated a biogenic COS source peaking in late June and early July. Comparison with the ORCHIDEE simulations revealed that these observed summer COS emissions were not represented in the model, highlighting a missing COS source in the ORCHIDEE simulations (Figure 10).

Then, surveys in spring and summer 2019, 2020, and 2021 were conducted using flask-air samples upwind and downwind of wheat and rapeseed fields in GIF to determine horizontal COS concentration gradients. These surveys revealed that rapeseed shifted from COS uptake to emission in early summer, while wheat's uptake significantly decreased, showing that rapeseed was a larger COS source than wheat at the plot scale. This suggests that the biogenic process not yet included in ORCHIDEE likely corresponds to COS production from rapeseed crops during their ripening and senescence stages. This COS production may be associated with a specific pathway in Brassicaceae species, involving isothiocyanates and thiocyanate ions (Blaževića et al., 2020). However, further research is needed to fully understand plant COS production pathways.

Finally, an empirical function was fitted between the simulated and observed ecosystem COS fluxes to estimate the potential COS emissions by rapeseed (Figure 10). Extrapolating this function across the global rapeseed harvested area of 32 million hectares in 2015 resulted in an estimated emission of 0.84 ± 0.13 GgS y^{-1}. This contribution is therefore minimal compared to anthropogenic COS emissions for example, which are estimated to range from +220 to +580 GgS y^{-1} (Zumkehr et al., 2018).

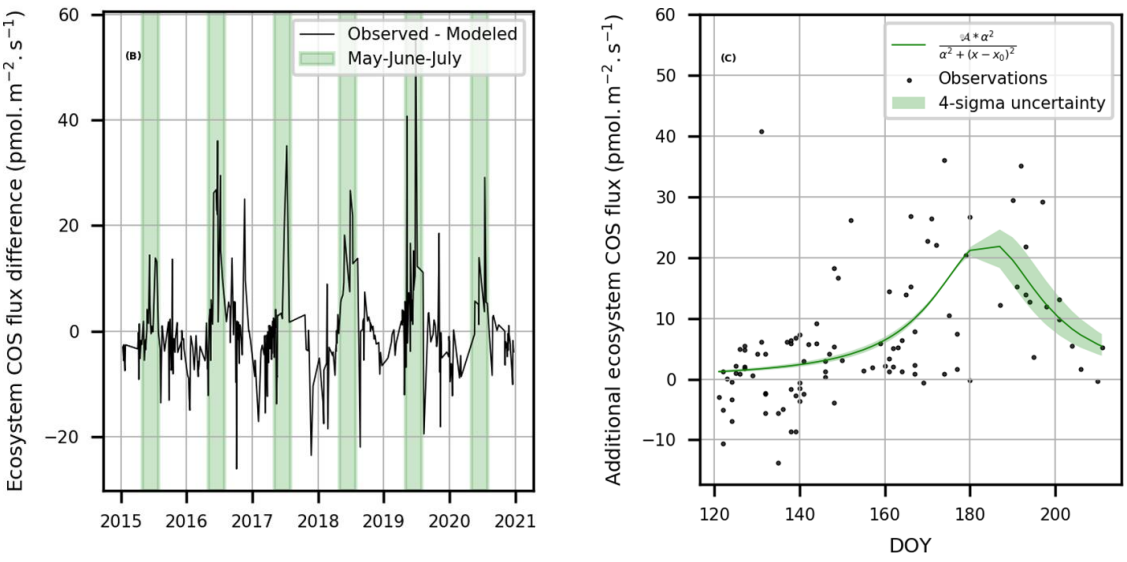

Figure 10: Comparison between observed and simulated nighttime fluxes at GIF. Left panel: Difference between observed and simulated ecosystem COS fluxes between 2015 and 2020. Right panel: Data from the left panel between May and July, gathered across years, fitted with the polynomial function presented in the legend, which estimates the dynamics of an additional source of COS not represented in ORCHIDEE. The light green shading around the fitting function represents the 4-sigma uncertainty.

Therefore, this study highlighted a missing process in ecosystem COS flux models implemented in LSMs, which accounts for vegetation's potential to produce COS. However, this source is of secondary importance compared to other COS sources for the global COS budget.

5.2.2 Evaluation of COS anthropogenic and biogenic emissions from atmospheric COS observations

Additional uncertainties persist beyond those related to the potential of crops to emit COS, such as for the vegetation, soil, or anthropogenic activity components, which are estimated to significantly contribute to the global COS budget (Whelan et al., 2018). While bottom-up modeling approaches provide estimates of surface COS fluxes (Berry et al., 2013; Kooijmans et al., 2021; Maignan et al., 2021; Abadie et al., 2022), further constraints on these fluxes can be achieved through atmospheric transport and simulated COS concentration evaluated against COS concentration data (Ma et al., 2021; Remaud et al., 2022).

Berchet et al. (in review) aimed to evaluate the biogenic COS fluxes simulated in ORCHIDEE and the anthropogenic emissions from two inventories, the one from Zumkehr et al. (2018) and a home-made inventory based on Belviso et al. (2023), in Western Europe. The home-made inventory incorporates data on CS_2 emissions from the viscose industry and CO_2 emissions from coal power plants in Europe. This evaluation involved using the FLEXPART model to transport all COS surface fluxes contributing to atmospheric COS concentration in this region and comparing the simulated concentrations with COS concentration measurements recorded at the GIF site in France over more than five years. In this study, I provided the biogenic COS fluxes from vegetation and soil COS flux simulation in ORCHIDEE, and contributed to the interpretation of the results.

The findings suggest a strong overestimation of COS anthropogenic emissions by the Zumkehr et al. (2018) inventory, especially in the Paris area despite the absence of coal power plant and major viscose industry, which cannot agree with the observed COS concentrations at GIF. In contrast, the new home-made inventory led to simulated atmospheric COS concentrations that were consistent with the observations.
Additionally, this study assessed the nighttime and daytime variations, defined as the 12-hour day/night differences in measured or simulated COS concentrations. This analysis highlighted an underestimation of the simulated biogenic COS uptake in ORCHIDEE at night in winter and spring. This underscores the need to refine ecosystem COS flux at night, attributed to both soil COS exchanges and vegetation COS uptake through a residual nighttime stomatal conductance.

Therefore, the transport of surface COS fluxes and the evaluation of the resulting atmospheric concentrations can reveal structural deficiencies in COS process representation or highlight the need for improved parameterization. Moveover, in contrast to the NOAA network providing monthly or bi-weekly flask samples at several measurement stations (Section 2.2), the GIF COS concentration time series offers a higher data frequency with hourly records, enabling a constraint on the surface fluxes at a diurnal time resolution.

5.2.3 Impact of recent change in atmospheric COS concentration on surface COS fluxes

While atmospheric CO_2 concentrations have continuously risen since pre-industrial times due to anthropogenic emissions, with atmospheric COS levels generally mirroring this trend, a decline in COS concentrations has been observed since 2015 (Li et al., 2024). Changes in atmospheric COS concentrations affect surface COS exchanges, impacting their contribution to the global COS budget. In particular, vegetation COS uptake presents a first order relationship with atmospheric COS concentration (equation 2). At the global scale, a first assessment of the impacts of the recent decrease in atmospheric COS concentrations on soil COS fluxes simulated in ORCHIDEE was conducted in Abadie et al. (2022). This assessment revealed that the reduction in atmospheric COS concentrations led to a decrease of about 25 GgS in the simulated net soil COS uptake between 2016 and 2019.

In Western Europe, Belviso et al. (2022b) monitored the trend in atmospheric COS concentrations measured at the GIF site in France from August 2014 to November 2021, and investigated its implications for biospheric COS fluxes using the COS fluxes simulated in ORCHIDEE. I provided the vegetation and soil COS fluxes simulated by ORCHIDEE for this study and contributed to the representation and analysis of the results. This study found a significant decrease in the seasonal cycle amplitude of COS concentrations, dropping by 27 ppt over six years, with the spring maximum declining more sharply than the autumn minimum. This decline aligns with tropospheric COS trends attributed to reduced anthropogenic emissions.
At the GIF site, the drop in atmospheric COS concentrations resulted in a 10% decrease in the simulated vegetation COS uptake in ORCHIDEE between 2016 and 2021. Across the entire Northern hemisphere, this decline in atmospheric COS concentration led to an 8% reduction in the simulated vegetation COS sink between 2016 and 2019 (Figure 11). Since vegetation COS uptake is the main driver of seasonal variations in atmospheric COS concentration in the Northern hemisphere, this suggests that the reduced plant COS uptake contributes to the decrease in the seasonal cycle amplitude of COS concentrations in recent years.

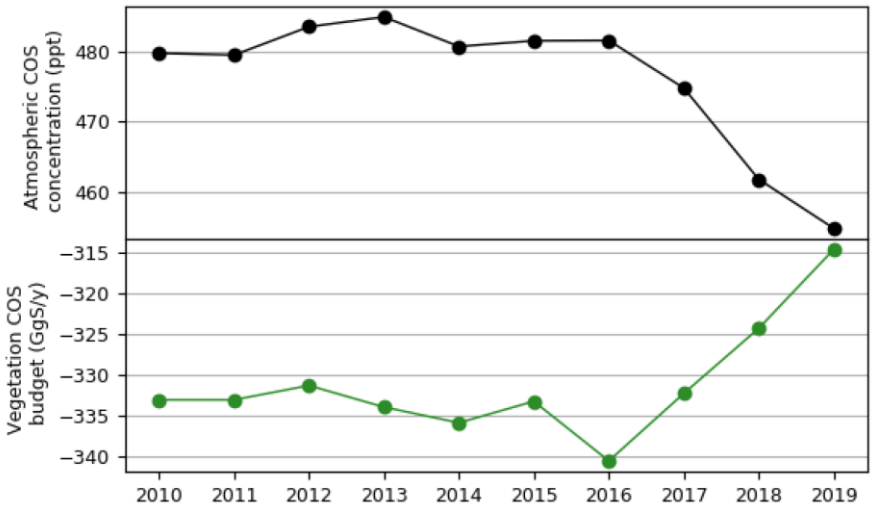

Figure 11: Evolution the mean annual atmospheric COS concentration and mean annual plant COS uptake computed with a monthly variable atmospheric COS concentration, in the Northern hemisphere between 2010 and 2019.

The importance of considering seasonal variations in atmospheric COS concentration to compute biospheric COS fluxes in LSMs has been highlighted previously, in particular to account for the impact of COS concentration drawdown during the vegetation growing season (Abadie et al., 2022; Kooijmans et al., 2021). Belviso et al. (2022b) emphasized the need to consider trends in atmospheric COS concentration when simulating vegetation COS fluxes, as these trends significantly affect the role of plants in the COS sink within the global COS budget.

5.2.4 Impact of atmospheric COS transport models

The studies presented in the previous sections have highlighted some of the uncertainties in the simulated surface COS fluxes. Additionally, uncertainties in the simulated atmospheric COS concentrations arise from the transport of COS fluxes by atmospheric transport models (ATMs). These uncertainties must be evaluated when using COS concentration measurements to constrain surface COS or CO_2 fluxes, such as to determine potential missing COS sinks or sources that could help closing the global COS budget.

Remaud et al. (2023) introduced the first atmospheric tracer transport model intercomparison project for COS (TransCom-COS), involving seven ATMs that simulated tropospheric COS concentrations based on state-of-the-art surface COS fluxes. In this study, I provided the ORCHIDEE simulated vegetation and soil COS fluxes for transport by the ATMs, in addition to those simulated in SiB4. This work aimed to explore transport uncertainties in simulated COS concentrations and evaluate surface COS fluxes in relation with the current unbalanced global COS budget. Simulated COS concentrations were compared with measurements from 15 NOAA network stations between 2012 and 2018.

The latitudinal gradients of observed and simulated COS concentrations were analyzed for boreal winter and summer (Figure 12). This analysis helped disentangle the contributions of oceanic and

biospheric COS fluxes, which show strong seasonality. In winter, all ATMs underestimated simulated COS concentrations in the tropics by 50 ppt, suggesting a missing oceanic COS source in this region. During summer, the ATMs failed to capture the observed latitudinal gradient, particularly overestimating simulated COS concentrations above 40°N by 50 ppt, likely due to a missing biospheric COS sink. These findings are consistent with previous studies by Ma et al. (2021) and Remaud et al. (2022). Significant transport uncertainties were also observed above 40°N in boreal summer, with a spread between ATMs exceeding 60 ppt.

Then, the evaluation of seasonal amplitude in COS concentrations revealed a large spread between ATMs, reaching 50 ppt at 6 of the 15 stations, which is half of the observed amplitude (100 ppt). Comparing simulated COS concentrations with ORCHIDEE and SiB4 biospheric fluxes or two different oceanic flux estimates demonstrated that changes in the biospheric or oceanic components typically had a lesser effect on the seasonal amplitude of simulated COS concentrations than the variability observed across different ATMs.

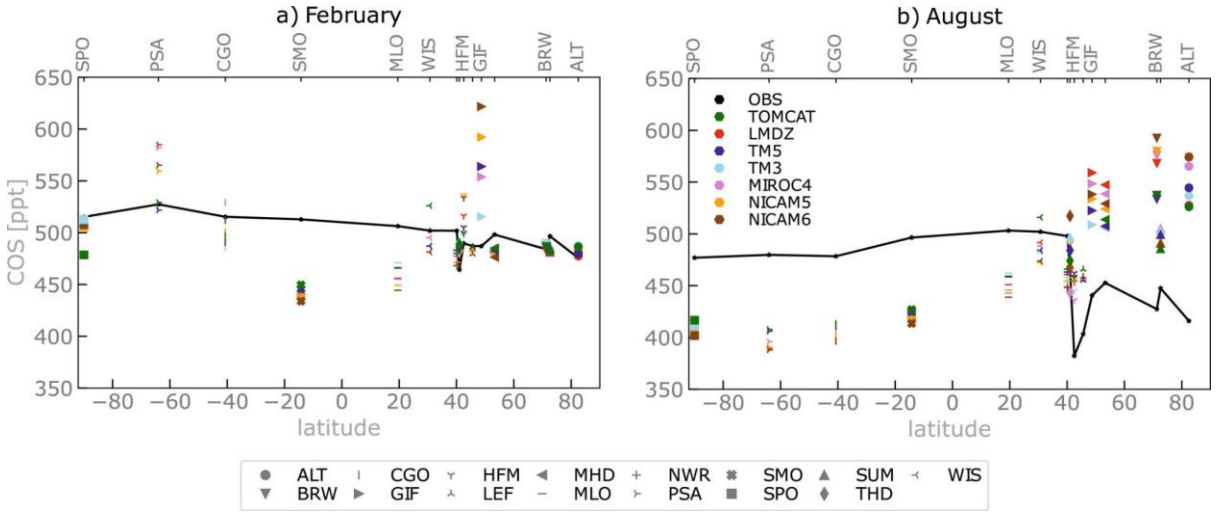

Figure 12: Comparison of the latitudinal variations of COS concentrations simulated by seven ATMs using a control scenario of state-of-the-art surface COS fluxes (colored dots) with the observations (black line) for February (left), and August (right) at 15 NOAA stations over the years 2012-2018. The simulated COS concentrations have been shifted such that the means are the same as the mean of the observations (~500 ppt). The time series of COS mixing ratio have been detrended and filtered to remove the synoptic variability beforehand. In August, the value at the GIF station simulated by the TOMCAT ATM was removed as it was an outlier (value above 800 ppt). For the same reason, the COS values at the GIF station simulated by TOMCAT (800 ppt) and LMDZ (around 700 ppt) have been removed in February. The KUM station has been removed, which is co-located in longitude and latitude with the MLO station, for the sake of simplicity. Removing the KUM station does not affect the results.

Therefore, while this study has highlighted significant uncertainties in the simulated COS concentrations due to the transport of surface COS fluxes, the various ATM simulations consistently identify discrepancies in surface COS fluxes, specifically with a missing COS source in the tropics and a missing COS sink in the Northern high latitudes.

On the other hand, directly assimilating COS flux observations to constrain COS and CO_2 surface fluxes eliminates the impact of transport uncertainties on COS concentrations. Assimilating COS

fluxes can also offer a more direct constraint as it involves fewer components of the global COS budget compared to assimilating COS concentrations, as discussed in Section 4.1. However, atmospheric COS concentrations provide information on a larger scale than COS flux observations.

5.3 Valorisation of atmospheric COS and CO_2 spatio-temporal gradients

5.3.1 Evaluation of simulated GPP and terrestrial ecosystem respiration from model intercomparison projects using atmospheric COS and CO_2 data

The studies by Belviso et al. (2022a, 2022b), Berchet et al. (in review), and Remaud et al. (2023) discussed in the previous section have advanced our understanding of the COS budget, from investigating local-scale COS flux processes to examining the impact of atmospheric transport models on COS concentrations. Improving the quantification of COS sinks and sources is crucial for using COS concentration observations to inform on GPP beyond the ecosystem scale. However, using previous estimates of COS components, Launois et al. (2015) utilized atmospheric concentration measurements to evaluate trends, and seasonal and spatial variations in GPP as simulated by three LSMs. This approach has been revised (based on some of the work conducted in this PhD) and further expanded to create a comprehensive evaluation framework based on COS concentration observations for a broader range of LSMs, incorporating the current knowledge on COS component contributions.

In Peylin et al. (to be submitted in early September 2024) presented in Appendix 9.3, we evaluated GPP and terrestrial ecosystem respiration across three model ensembles (Coupled Model Intercomparison Projects Phase 5 (CMIP5), CMIP6, Trend in the Land Carbon Cycle Project Version 7 (TRENDY-V7)) using a combined analysis of atmospheric COS and CO_2 concentrations. This evaluation framework is based on the premise that CO_2 concentrations reflect both GPP and respiration fluxes, whereas COS concentrations offer a more direct constraint on GPP. This joint analysis aims to identify potential biases in the seasonal amplitude and/or phase of GPP and/or respiration for individual models, encompassing a total of 64 models, as well as differences across the intercomparison experiments.
This evaluation uses atmospheric COS and CO_2 concentration measurements from the NOAA network, focusing on two stations capturing the influence of the Northern hemisphere fluxes: Mauna Loa in the Hawaiian region (MLO) and Alert in Canada (ALT). These measured atmospheric concentrations are compared to simulated concentrations resulting from the transport of all COS or CO_2 flux components using the LMDZ version 3 transport model. The vegetation COS uptake for each model is determined using the LRU relationships based on the simulated GPP from each LSM. Therefore, the simulated concentrations only differ between models by the vegetation COS flux transported (based on model GPP) for the COS concentrations, and by the net biome productivity (NBP) simulated in each model and transported for the CO_2 concentrations. Finally, we evaluate the sensitivity of the simulated COS concentration to a change in LRU values, or in the

soil or ocean contributions. In this work, Philippe Peylin initiated and conceptualized the approach, while I provided the ORCHIDEE simulated soil COS fluxes for the transport, ran the atmospheric transport simulations, and created the figures. Then, we both contributed equally to writing the article.

The identification of seasonal amplitude biases involves comparing simulated COS and CO_2 amplitudes normalized by observed ones (Figure 14). If a model overestimates or underestimates both COS and CO_2 amplitudes, it suggests a misfit likely influenced by GPP, which directly affects simulated concentrations of both gases (in a coherent and similar way, see the LRU equation in Figure 3). Conversely, if a model underestimates COS amplitude while overestimating CO_2, or vice versa, it indicates potential biases in both GPP and respiration fluxes, as solely adjusting the GPP flux alone cannot simultaneously correct the simulated amplitudes of COS and CO_2.
The seasonal phase analysis is based on the COS versus CO_2 correlations between the simulated and observed concentrations. A model with low correlations for both gases suggests a predominance of GPP flux in the simulated phase biases. Alternatively, a model showing high COS correlation but not CO_2 implies a respiration bias alongside accurate GPP flux. Conversely, a model with high CO_2 correlation but low COS correlation may indicate biases in both GPP and respiration phases, with respiration compensating for a GPP bias.

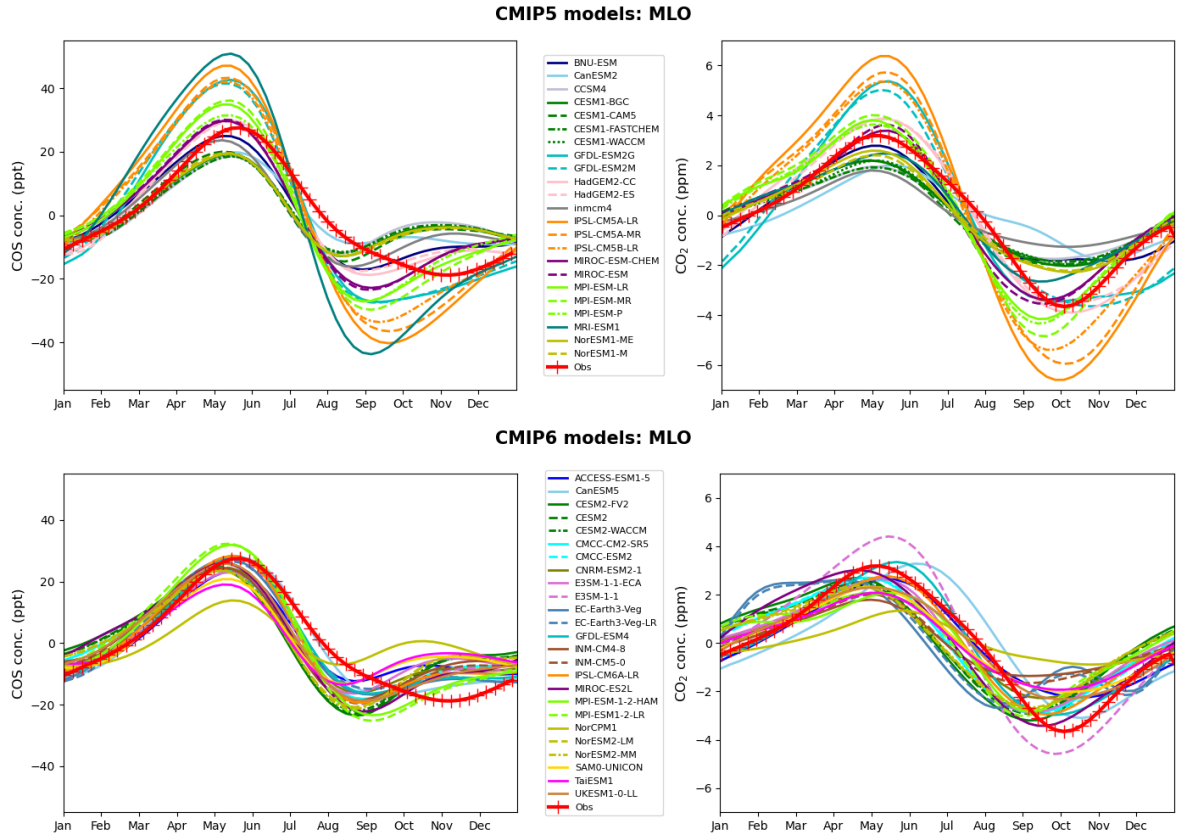

Figure 13: Detrended mean seasonal cycles of atmospheric COS (left) and CO_2 (right) concentrations simulated at Mauna Loa station (MLO) for all models of CMIP5 (top) and CMIP6 (bottom) ensembles. The observations are represented by red crosses.

The evaluation of the simulated COS and CO_2 concentrations reveals a significant spread in their seasonal amplitudes across the CMIP5 model ensemble, ranging from 29 ppt to 94 ppt for COS

and from 3.1 ppm to 13 ppm for CO_2, related to the large variability in simulated GPP among the LSMs (Figure 13). Furthermore, the joint analysis demonstrates a positive linear relationship between the normalized seasonal amplitudes of COS and CO_2 concentrations simulated by CMIP5 models, underscoring the influential role of GPP on these variations (Figure 14). While some LSMs accurately replicate observed seasonal amplitudes, a majority of CMIP5 models exhibit a bias in GPP seasonal cycle amplitude for the Northern hemisphere.

The CMIP6 ensemble exhibits a noteworthy reduction in the variability of COS and CO_2 seasonal amplitudes compared to CMIP5, ranging from 21 ppt to 57 ppt for COS and from 2.2 ppm to 9 ppm for CO_2, indicating increased model convergence (Figure 13). However, across most CMIP6 models, both CO_2 and COS seasonal amplitudes are slightly underestimated (Figure 14), suggesting a general tendency towards lower simulated GPP seasonal amplitudes in these models.

Assessing the last intercomparison model ensemble, TRENDY-V7, reveals a wide range in simulated COS and CO_2 seasonal amplitudes, between 30 ppt and 74 ppt for COS and between 1.9 ppm and 7.7 ppm for CO_2, despite using the same meteorological forcing for all models.

Then, considering the contributions of the various COS components used in this study, the global COS budget remains unbalanced, with an average deficit in COS sources ranging from 276 GgS y^{-1} to 530 GgS y^{-1} across intercomparison exercises. Adjusting lower LRU values could potentially balance the COS budget by reducing the vegetation sink. However, solely lowering LRU values would not suffice, as it would decrease the simulated COS seasonal amplitude, potentially causing all models to deviate from observed values. Therefore, other COS processes would need to display a more pronounced seasonal cycle in phase with that of GPP.

Finally, sensitivity tests exploring the simulated COS concentrations with varying estimates of soil and ocean contributions demonstrated significant impacts on the global COS budget for both components, although the soil component showed a lesser effect on the seasonal amplitude. This highlights the need for improved understanding and estimation of COS fluxes from oceans and soils.

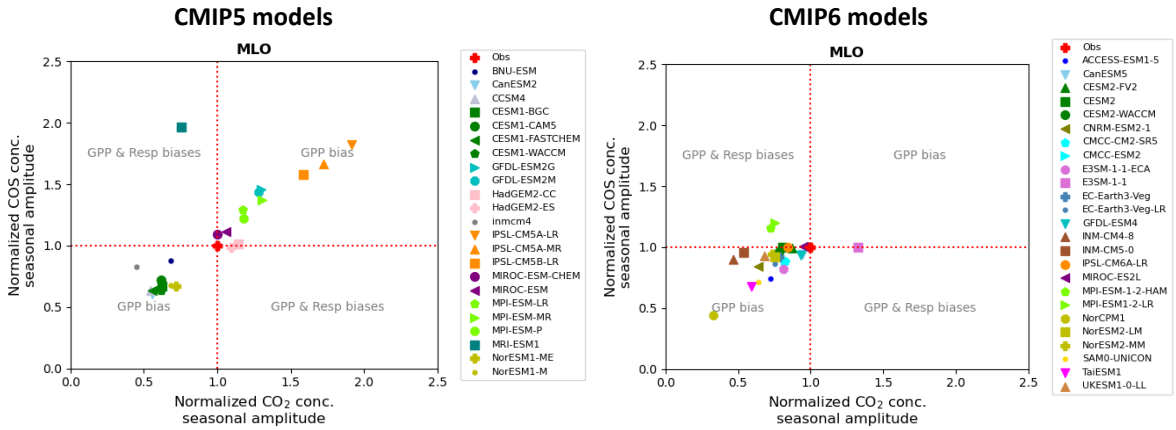

Figure 14: Scatter plots of CO_2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO for CMIP5 (left) and CMIP6 (right) models using the reference scenario. The observed amplitude is identified by the red cross.

Despite these sources of uncertainty, this study has developed a promising framework to evaluate simulated ecosystem CO_2 gross fluxes (GPP and ecosystem respiration) from various model

ensembles used in intercomparison exercises, using COS and CO_2 concentration measurements. This work highlighted some individual model discrepancies that were also confirmed by previous studies (i.e., the ORCHIDEE model used in the CMIP5 exercice exhibited an excessively large seasonal amplitude in GPP at Northern high latitudes, a discrepancy that was largely corrected in CMIP6, see Figure 14). This evaluation framework can be applied to the latest TRENDY versions and future CMIP and TRENDY exercises, enabling the monitoring of simulated GPP and respiration biases over successive model intercomparison experiments. Additionally, the approach should be regularly updated to incorporate future advancements in the characterization of COS flux components.

5.3.2 Inferring COS and CO_2 surface fluxes from inverse modeling of atmospheric COS and CO_2 data

As seen in the previous section, atmospheric transport models can be used to evaluate potential biases in LSM GPP estimates by transporting the different COS and CO_2 surface flux components and comparing the resulting simulated concentrations against concentration observations. In contrast, atmospheric inversion modeling approaches enable optimization of the surface fluxes in addition to identifying discrepancies in the spatial and temporal dynamics of these fluxes.

Remaud et al. (2022) conducted a joint assimilation of atmospheric COS and CO_2 concentration data into the LMDZ atmospheric transport model to constrain surface COS fluxes, GPP, and plant respiration, using scaling factors applied to an ensemble of few large ecosystem-based regions. They assimilated atmospheric COS and CO_2 concentration observations at 15 sites from the NOAA network. This approach not only improved estimates of large-scale GPP and plant respiration fluxes but also addressed the challenge of the global COS budget closure, which has been hindered by significant uncertainties in its components. The inversion framework uses simulated ORCHIDEE GPP and plant transpiration fluxes, optimizing each flux for the 15 ORCHIDEE PFTs. Vegetation COS flux is linked to GPP through the empirical LRU relationship, distinguishing values for C_3 and C_4 plants. In this study, I have performed the ORCHIDEE simulation providing the fluxes that have been transported and optimized. Then, the simulated concentrations are evaluated against independent data with MIPAS satellite retrievals, HIAPER Pole-to-Pole Observations (HIPPO) airborne measurements, and surface measurements over Japan and France.
The resulting optimized COS budget suggests a global biospheric uptake of -800 GgS y^{-1} between 2008 and 2019, with enhanced absorption observed in high latitudes (Figure 15), mainly attributed to a vegetation sink of -620 GgS y^{-1}. Optimized oceanic COS emissions have been doubled to 530 GgS y^{-1}, with the increase primarily concentrated in tropical regions, while emissions in high latitudes were revised downward compared to the prior. This suggests that the contribution of DMS to higher COS emissions in the tropics might be significant, given that measurements of COS and CS_2 in seawater do not support such an increase in oceanic sources (Lennartz et al., 2017, 2020). The decrease in oceanic emissions in high latitudes may indicate an overestimation of direct COS emissions and COS emissions via CS_2 in this region.
In terms of the global CO_2 budget, the inversion substantially increased net vegetation uptake in high latitudes, nearly tripling the previous estimate. GPP in this region was boosted by almost 2

GtC y^{-1}, while the respiration flux remained unchanged. Conversely, in tropical regions, both respiration and GPP showed a tendency to decrease.

The evaluation against independent concentration data showed that the system effectively corrects previously underestimated GPP in high latitudes. However, the decrease in biospheric uptake in the Amazon following the inversion contradicts MIPAS COS retrievals, highlighting the limited observational constraints in this critical region. Finally, comparisons with surface measurements in Japan and France indicate inaccuracies in the prior anthropogenic COS emission inventory from Zumkehr et al. (2018), with too strong sources over Europe and Japan, and too weak sources in the Eastern edge of China. These findings highlight the necessity for an improved inventory and supports the new one proposed by Belviso et al. (2023) for Europe.

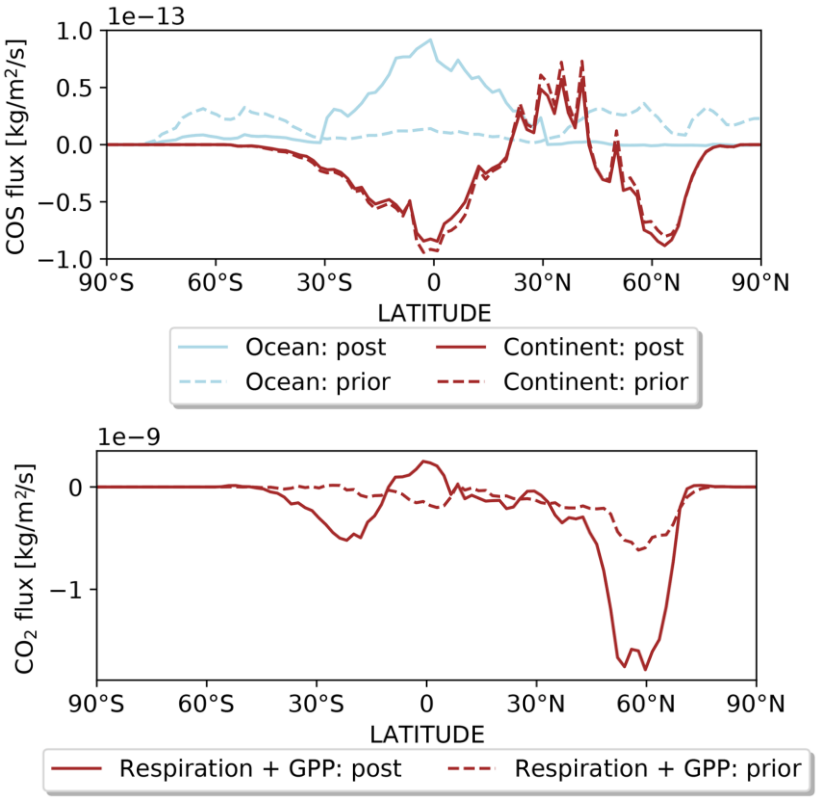

Figure 15: Latitudinal distribution of the prior (dashed line) and posterior fluxes (full line): for the continental (red) and oceanic components (blue) of the COS budget (top), and for the net CO_2 fluxes from terrestrial vegetation (bottom). Terrestrial vegetation fluxes are the sum of GPP and respiration fluxes. The fluxes have been averaged over the years 2009-2019.

Therefore, although this atmospheric inversion modeling study does not directly constrain ORCHIDEE model parameters and process representations that determine the simulated biospheric COS and CO_2 fluxes, it highlights discrepancies in these fluxes and guides future improvement efforts.

6 IMPROVING THE SIMULATED RESPONSE OF GPP AND PLANT TRANSPIRATION TO DROUGHT EVENTS

6.1 Current limitations in the representation of vegetation physiological response to drought events in ORCHIDEE

During the investigation of COS potential to improve the simulation of GPP and plant transpiration in ORCHIDEE, a particular attention was given to the constraint introduced by COS (through a data assimilation approach) when evaluating a drought event at the Hyytiälä site (Section 4.2). This assessment revealed that while DA experiments incorporating COS data improved both simulated GPP and LE, parameter optimization through DA was insufficient to accurately represent this drought event in ORCHIDEE. This finding underscores the structural insufficiency in ORCHIDEE's representation of vegetation responses to drought events, which is critical for carbon, water, and energy flux simulation as drought events are expected to be more frequent and intense with climate change. Consequently, the final part of this PhD work focuses on improving the simulation of GPP and plant transpiration under soil moisture stress conditions.

Issues in representing vegetation response to soil moisture decrease in ORCHIDEE have been previously reported across different versions, leading to contrasting conclusions (Rebel et al., 2012; De Pue et al., 2022; MacBean et al., 2020). For example, Rebel et al. (2012) identified an underestimation of vegetation response to droughts in an earlier ORCHIDEE model version that lacked a multi-layer soil hydrology component. In semi-arid ecosystems, MacBean et al. (2020) also demonstrated that incorporating a discretized soil hydrology scheme with 11 soil layers in ORCHIDEE improved the simulated evapotranspiration through a more accurate representation of soil moisture in the upper layers compared to a simpler 2-layer bucket scheme. However, more recently, using the ORCHIDEE version from the CMIP6 exercise (Boucher et al., 2020; Cheruy et al., 2020), De Pue et al. (2022) found an oversensitivity of the simulated LE and GPP to soil moisture decrease, which strongly correlates with errors in the simulated LAI.

In the ORCHIDEE version used in this work (Section 2.1), the physiological vegetation response to drought is modeled by an empirical function ($f_{water\ stress}$) that limits stomatal conductance, mesophyll conductance, and the Rubisco carboxylation and RuBP regeneration rates as soil moisture decreases (Figure 16). In the default version of the model, $f_{water\ stress}$ decreases linearly when soil moisture drops below a threshold corresponding to no stress conditions ($\theta_{no\ stress}$) until it reaches a minimum threshold at the wilting point (θ_{WP}), where the stomata close as plants can no longer extract water from the soil. These thresholds, $\theta_{no\ stress}$ and θ_{WP}, are determined by soil texture. This $f_{water\ stress}$ function is computed for each soil layer except the topsoil layer considered without root (~1 mm deep), weighted by the root profile, and integrated over the soil column to determine the soil moisture stress applied for each PFT.

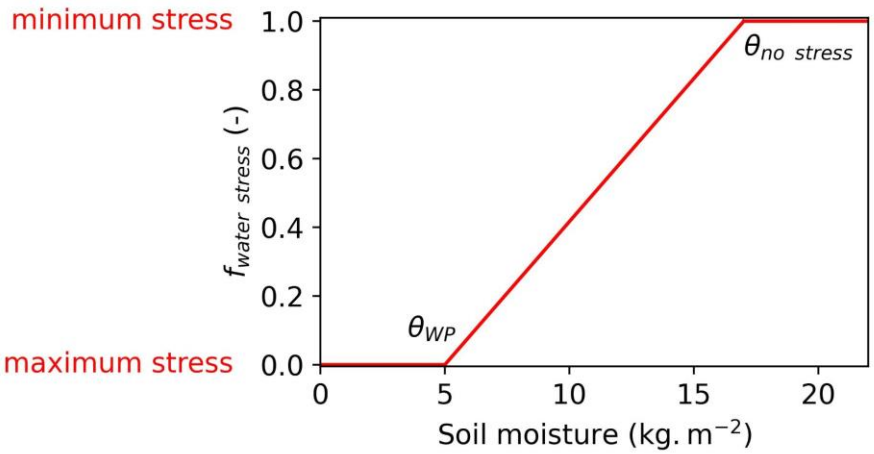

Figure 16: Variations of the function ($f_{water\ stress}$) that controls stomatal closure when soil moisture (kg.m^{-2}) decreases, as implemented by default in ORCHIDEE. Here, its variations are illustrated considering a minimum stress for soil moisture above 18 kg.m^{-2} ($\theta_{no\ stress}$) and a maximum stress below 5 kg.m^{-2} (θ_{WP}).

Therefore, the default simulated response to drought in ORCHIDEE depends solely on soil moisture availability and soil texture, which determines $\theta_{no\ stress}$ and θ_{WP}. Additionally, the response varies between tree and herbaceous PFTs, as the maximum root depth is set to 2 m for trees and 1 m for herbaceous species. This is a significant limitation as several studies have shown that vegetation responses to soil moisture stress vary across different biomes (McDowell et al., 2008; Chaves et al., 2003; Li et al., 2023). Consequently, the default simulated drought response in ORCHIDEE has neither been defined nor calibrated for different PFTs, failing to account for the diversity of vegetation responses to soil moisture stress. Then, other climatic factors beyond soil moisture, such as VPD and atmospheric CO_2 levels, have been found to influence physiological vegetation responses to drought (Swann et al. 2016; Stovall et al., 2019; Li et al., 2023), which can lead to variations in drought response within biomes.

Furthermore, the default drought response implementation only considers the immediate effects of soil moisture changes on vegetation physiology, without integrating potential mechanisms of plant plasticity. Indeed, several studies have demonstrated that plants can acclimate and adapt to environmental changes over short to long timescales, which can influence their responses to drought events (Marchin et al., 2016; Grossiord et al., 2017).

Finally, another potential limitation of the physiological vegetation response to drought in ORCHIDEE is the application of the same function to stomatal conductance, mesophyll conductance, and the Rubisco carboxylation and RuBP regeneration rates. This approach assumes a coordinated response between stomatal and non-stomatal limitations under soil moisture stress conditions. Indeed, Drake et al. (2017) indicates that both types of limitations occur during droughts. However, the relative importance of mesophyll and Rubisco activity limitations is still debated and may vary depending on the species and the duration of stress (Cano et al., 2014; Gimeno et al., 2019). As a result, differentiating the responses of stomatal conductance, mesophyll

conductance, and photosynthetic efficiency to soil moisture stress remains challenging to implement in LSMs and is associated with significant uncertainty.

The following section details the last focus of this PhD dedicated to improving vegetation physiological response to soil moisture stress, with the draft of the article soon to be submitted (before the end of September 2024) to *Global Biogeochemical Cycles*. In this study, another definition of the $f_{water\ stress}$ function is tested, which has been previously implemented in ORCHIDEE based on the work of Meridja (2011). Such function replaces the linear dependence of $f_{water\ stress}$ to soil moisture with an exponential bounded function. Although this new definition has not yet been calibrated for the different PFTs in ORCHIDEE and therefore shares the same limitations as the default $f_{water\ stress}$ function shown in Figure 16, it was originally developed by Meridja (2011) to better account for the diversity of plant sensitivity to soil moisture decrease. The study in the following section aims to calibrate this new function for several PFTs and proposes incorporating an acclimation mechanism in the simulated vegetation physiological response to soil moisture stress.

6.2 Refining vegetation physiological response to drought in ORCHIDEE by incorporating acclimation to atmospheric moisture stress

Vegetation acclimation to atmospheric moisture stress during drought events over Europe

Camille Abadie[1], Fabienne Maignan[1], Cédric Bacour[1], and Philippe Peylin[1]

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

Abstract

Improving the simulated response of Gross Primary Production (GPP) and plant transpiration to soil moisture stress in Land Surface Models (LSMs) is crucial for accurate climate simulations, especially as drought events are expected to increase. In this study, we used the ORCHIDEE LSM to investigate how the simulated physiological response of vegetation to soil moisture stress can be refined, focusing on Europe, which has experienced significant droughts. We used *in situ* GPP and latent heat flux (LE) data from more than 40 sites across various biomes, along with data assimilation techniques, to improve GPP and plant transpiration representations. This work shows that the speed of stomatal closure under soil moisture stress can be refined by incorporating vegetation acclimation to long-term vapor pressure deficit (VPD) conditions. This new drought response results in a greater reduction in GPP root mean square deviation than a response based solely on biome type, achieving an 18% improvement at the site scale, whereas the biome-type-

only version shows no improvement. However, the two model versions show similar performance in simulating LE, with an 8 to 9% improvement at the site scale but a slight degradation at the regional level (using data-driven benchmarks). Projections until 2100 show that the model incorporating VPD acclimation results in the smallest increase in soil moisture stress, 22% less than the model based solely on vegetation-type-dependent responses. This study underscores the importance of better understanding potential acclimation mechanisms and how to implement them in LSMs to enhance the accuracy of climate projections.

1. Introduction

Europe has been experiencing an increase in climate anomalies due to climate change, leading to more frequent and severe drought events in recent years (Bastos et al., 2020). The period from 2018 to 2020 is particularly notable for its significant impacts on ecosystems, as well as substantial social and economic losses (Conradt et al., 2023). Soil moisture stress is often considered the primary factor affecting plant functioning during droughts (McDowell et al., 2008). The decrease of soil moisture typically induces plant water stress, causing stomatal closure to prevent water loss, which in turns also limits carbon uptake and can reduce plant productivity. However, the combined effects of various climate change drivers on plant responses to drought are complex and difficult to disentangle. While warm conditions can enhance photosynthesis and plant transpiration (Dreyer et al., 2001; Way et al., 2015; Dusenge et al., 2018; Urban et al., 2017), further depleting soil moisture, the rising temperatures also increase vapor pressure deficit (VPD) (Grossiord et al., 2020), promoting stomatal closure to conserve water (Monteith, 1995; Oren et al., 1999). High VPD conditions usually coincide with soil moisture stress and amplifies its effect (Grossiord et al., 2020). Given VPD's significant role in drought-induced plant mortality (Breshears et al., 2013; Stovall et al., 2019), it is crucial to consider its current and future impacts on vegetation responses to drought.

Land surface models (LSMs) have been developed to investigate ecosystems functioning and are now increasingly used to project their possible responses to climate change. However, LSMs often struggle to accurately represent the diverse responses of vegetation to extreme events such as droughts and to account for different response timescales. For instance, several LSMs, including the Organising Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE), have been found to overestimate the frequency of hydrological droughts due to their over-sensitivity to short-term precipitation variability (Prudhomme et al., 2011; Tallaksen & Stahl, 2014). As a result, vegetation carbon uptake and evapotranspiration simulated in ORCHIDEE were found to be overly sensitive to droughts (De Pue et al., 2022).

Accurately representing stomatal response to drought is particularly critical, as it controls the coupling between carbon uptake and transpiration (Cowan & Farquhar, 1977; Kozlowski & Pallardy, 2002; Prentice et al., 2014; Marchin et al., 2016), determining vegetation water stress strategies that can be illustrated by the isohydricity concept (Tardieu & Simmoneau, 1998). According to this concept, isohydric species close their stomata to maintain steady leaf water potential under water stress, whereas anisohydric species keep their stomata open longer, resulting in a drop in leaf water potential. LSMs are evolving to include more complex processes to improve the realism of vegetation functioning representation, but the diversity of stomatal conductance response mechanisms remains challenging to implement. Short-term stomatal

regulation, such as the impact of instantaneous changes in VPD on stomatal conductance (Medlyn et al., 2011; Yin & Struik, 2009), is commonly accounted for in LSMs, including ORCHIDEE. However, on longer timescales, several studies have also highlighted the potential for stomatal conductance to acclimate to environmental conditions (Marchin et al., 2016; Kutsch et al., 2001).

In this study, we have refined the representation of vegetation physiological response to soil moisture stress in the ORCHIDEE LSM by comparing two formulations: one that only incorporates a biome-type dependency and another that also accounts for vegetation potential to acclimate to long-term VPD conditions. These formulations were evaluated for their ability to accurately simulate gross primary production (GPP) and latent heat flux (or evapotranspiration) (LE) during drought events. In doing so, we address the following questions:

1. Is a biome dependency sufficient to simulate vegetation response to soil moisture stress?
2. What is the impact of accounting for long term VPD conditions in the simulated vegetation response to drought?
3. How does accounting for the influence of environmental conditions in vegetation response to drought impact plant carbon and water exchanges under future climate?

The refinement of the vegetation physiological response to soil moisture stress in ORCHIDEE - including a formulation based solely on biome type and another that incorporates vegetation potential to acclimate to long-term VPD - and the calibration of these formulations are supported by assimilating *in situ* GPP and LE data. The impact of these updated ORCHIDEE versions on the simulated GPP and LE is then evaluated at both site and regional scales across Europe using *in situ* data and satellite-based evaluation products. Finally, projections through 2100 are conducted to assess the effects of these different vegetation responses to soil moisture stress under future climate conditions.

2. Materials and methods

2.1. The ORCHIDEE LSM

2.1.1. General model description

The ORCHIDEE LSM is the land component of the Earth System Model developed at the Institut Pierre Simon Laplace (Boucher et al., 2020; Cheruy et al., 2020). ORCHIDEE simulates carbon, water, and energy exchanges between land surfaces and the atmosphere from a daily time step for slow processes such as carbon allocation, to a half hourly time step for fast processes such as photosynthesis (Krinner et al., 2005). Vegetation is grouped into 14 plant functional types (PFTs) based on similar phenology, photosynthetic metabolisms, structure, and bioclimatic range, and a last PFT represents bare soil. PFT maps are prescribed annually using European Space Agency (ESA) Climate Change Initiative (CCI) land cover products (Lurton et al., 2020). The soil profile is decomposed into 11 soil layers with increasing depth from the top to the bottom of the soil column at a 2 m depth, and soil hydraulics are computed in each layer. Vertical water fluxes within the soil are modeled using the Richards equation across the discretized soil column and a free drainage condition is set at the column's base (de Rosnay et al., 2002; D'Orgeval et al., 2008). Soil textures, which determine soil characteristics such as porosity and hydraulics, are described using 12

classes from a global soil map based on the Food and Agriculture Organization of the United Nations/United States Department of Agriculture (FAO/USDA) texture classification (Reynolds et al., 2000). Global yearly mean near-surface CO_2 concentrations are provided by the TRENDY (Trends in the land carbon cycle) model intercomparison project (Sitch et al., 2015).

The ORCHIDEE version used in this study builds upon the model used for the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Boucher et al., 2020; Lurton et al., 2020). However, this version features updates compared to the CMIP6 version, such as revisions to the radiative transfer model (RTM) described below, the representation of vegetation physiological response to soil moisture stress based on the work of Meridja (2011), and the representation of the root profile (see section 2.1.2).

This ORCHIDEE version includes a RTM that distinguishes between direct and diffuse light within the canopy, based on the RTM developed by Zhang et al. (2020). The canopy is discretized in a maximum of 20 leaf area index (LAI) layers, with finer layers at the top to thicker layers at the bottom of the canopy. The calculation of the absorbed light, photosynthesis, and stomatal conductance per PFT are carried out for each LAI layer and for sunlit and shaded leaves separately. Photosynthesis and stomatal conductance are described following Yin & Struik (2009), based on the photosynthesis model of Farquhar et al. (1980) for C_3 species and Collatz et al. (1992) for C_4 species. In this model, an empirical function accounts for the effect of instantaneous change in VPD on stomatal conductance at each time step (Yin & Struik, 2009; Text S1). The maximum rates of Rubisco carboxylation and RuBP regeneration depend on leaf age (Ishida et al., 1999; Krinner et al., 2005), and on an implicit leaf nitrogen content, which decreases in the lower canopy (Johnson & Thornley, 1984). Both rates also respond to air temperature, following a modified Arrhenius function that accounts for acclimation to growth temperature, as described by Kattge & Knorr (2007). Additionally, a limitation is imposed on the maximum rates of Rubisco carboxylation and RuBP regeneration to account for the downregulation of the productivity under elevated CO_2 levels (Sellers et al., 1996; Bounoua et al., 1999, 2010). Then, photosynthesis and stomatal conductance are integrated over all LAI layers to compute GPP at the canopy level and the total canopy stomatal conductance.

Plant transpiration for each PFT is calculated at the canopy level, driven by the saturated moisture gradient between the surface and air at 2 m, and regulated by aerodynamic, leaf boundary layer, and total canopy stomatal conductances.

LE is the sum of various processes at the grid cell level, including plant transpiration, bare soil evaporation, canopy interception and evaporation, snow sublimation, and floodplain evaporation. Each process contributing to LE is driven by the saturated moisture gradient between the surface and air at 2 m and is regulated by aerodynamic, boundary layer conductances, and a specific conductance for each process, similar to how canopy stomatal conductance regulates transpiration. A resistance to soil evaporation is implemented following Sellers et al. (1992) to limit soil evaporation when soil moisture decreases (Text S1).

The aerodynamic conductance affecting both plant transpiration and carbon uptake is detailed in Su et al. (2001), which updates the approach of Massman (1999) based on localized near-field Lagrangian theory.

2.1.2. Representation of vegetation response to soil moisture stress

In ORCHIDEE, vegetation response to soil moisture stress is represented by an empirical function that limits stomatal conductance, mesophyll conductance, and Ribulose-1,5-bisphosphate carboxylase/oxygénase (Rubisco) carboxylation and ribulose-1,5-bisphosphate (RuBP) regeneration rates. This function is computed for each soil layer, except the topsoil layer (~1 mm deep) as it is considered without root and does not contribute to plant transpiration. In this model version, the soil moisture stress response is defined as an empirical exponential function $f_{water\ stress}$ ([0,1], unitless) that depends on soil water status and vegetation sensitivity to water stress following Meridja (2011):

$$f_{water\ stress,l} = exp\left(-\alpha_{water\ stress} \cdot \frac{(\theta_{FC,l}-\theta_{WP,l})}{(\theta_{no\ stress,l}-\theta_{WP,l})} \cdot \frac{(\theta_{no\ stress,l}-\theta_l)}{(\theta_l-\theta_{WP,l})}\right) \cdot root_{density,l} \quad \text{(Eq. 1)}$$

with θ_l the soil moisture (kg.m^{-2}) in a given soil layer l, and $\theta_{FC,l}$ and $\theta_{WP,l}$ the soil moisture at field capacity and wilting point (kg.m^{-2}) in the soil layer l, respectively. $\theta_{no\ stress,l}$ is the soil moisture threshold above which there is no water stress (kg.m^{-2}), defined as $\theta_{no\ stress,l} = \theta_{WP,l} + p_{max} \cdot (\theta_{FC,l} - \theta_{WP,l})$, with p_{max} the coefficient ([0,1], unitless) that determines this threshold. $f_{water\ stress,l}$ is bounded between 0 (when $\theta_l < \theta_{WP,l}$) and 1 (when $\theta_l > \theta_{no\ stress,l}$). Then, $\alpha_{water\ stress}$ is the water stress sensitivity coefficient which was initially considered as a constant ($\alpha_{water\ stress}$ = 1., unitless) in ORCHIDEE (see red curve in Figure 4), but is now defined as PFT-dependent in agreement with Meridja (2011) ([0.05-10.]). This coefficient determines the speed of vegetation response to decreasing soil moisture between $\theta_{no\ stress,l}$ and $\theta_{WP,l}$. Therefore, this coefficient could indicate either a more isohydric behavior, characterized by rapid stomatal closure under soil moisture stress, or a more anisohydric behavior, marked by slower stomatal closure under the same conditions. $root_{density,l}$ is the normalized root length fraction in the soil layer l ([0,1], unitless) (the fraction of the root length density in the layer divided by the total root length in the root zone) that weights the water stress function, and defined according to the available water in each layer (see Text S1). Finally, $f_{water\ stress,l}$ is summed over all soil layers to determine a response to soil moisture stress for each PFT.

Here, we aim to refine the representation of $f_{water\ stress}$ using data assimilation techniques, with a particular focus on the $\alpha_{water\ stress}$ coefficient, which has not yet been calibrated in ORCHIDEE. Specifically, we investigate whether $\alpha_{water\ stress}$ should depend solely on the PFT or if it should also vary with environmental factors.

2.1.3. Simulations

2.1.3.1. Simulations over the present period

For any simulation, we first performed a spin-up phase to stabilize all carbon pools and equilibrate net ecosystem production in the absence of any disturbance (Lardy et al., 2011). This phase requires cycling over the available years in the meteorological forcing data for 340 years with a constant atmospheric CO_2 concentration corresponding to a pre-industrial level. The spin-up is followed by a transient phase that introduces disturbances such as increasing atmospheric CO_2 concentration, and land use and climate change. The transient phase also cycles over the available

forcing years for 60 years until the first year of the meteorological data. Then, the simulations over the recent years are performed at a daily time step.
Site simulations (section 2.2) were forced by local micro-meteorological half-hourly measurements from the ICOS Warm Winter 2020 network (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022). Simulations over Europe were forced by 6-hourly CRU JRA reanalysis at a 0.5° spatial resolution over 2000-2021 (University of East Anglia Climatic Research Unit Japanese Reanalysis; Friedlingstein et al., 2020).

2.1.3.2. Projections

Projections over Europe were conducted following the shared socioeconomic pathway scenario 3 (SSP) with an additional radiative forcing of 7 $W.m^{-2}$ by the year 2100, called hereafter SSP370. This scenario is one of the scenarios considered in the Scenario Model Intercomparison Project (ScenarioMIP) (O'Neill et al., 2016), which is part of CMIP6. It was chosen as it represents a medium-to-high challenge scenario for mitigation and adaptation, combining relatively high societal vulnerability and forcing with important aerosol emissions and land use change. We used meteorological forcing datasets from the UK ESM (UKESM1) (Alistair et al., 2019) at a 2° spatial resolution, covering 1850 to 2014 for the historical period and 2015 to 2100 for projections.
A spin-up phase was first performed as described in the previous section, by cycling over the 10 first years of the historical forcing file for 340 years. Then, simulations at a daily time step were performed from 1850 to 2014 for the historical period, followed by projections until 2100. Note that for the projection simulations only, the PFT map is not updated each year and is fixed to the one of 2015.

2.2. Studied area and data

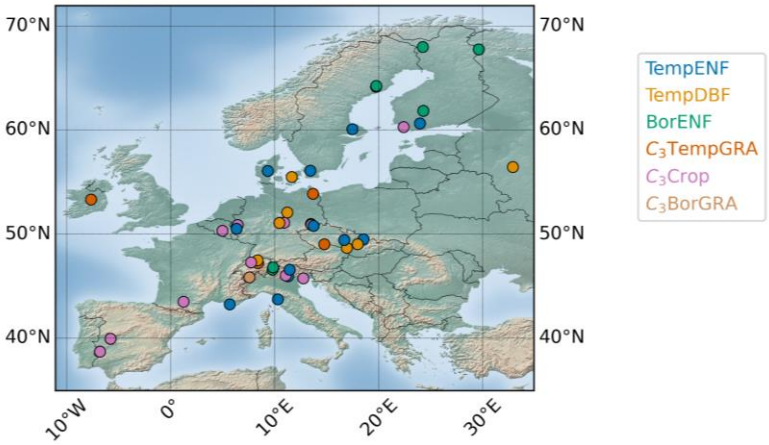

Figure 1: Location of the ICOS Warm Winter sites with a dominant PFT fraction of at least 50%. The legend colors represent the dominant PFT at each site or pixel. TempENF = Temperate evergreen needleleaf forest; TempDBF = Temperate deciduous broadleaf forest; BorENF = Boreal evergreen needleleaf forest; C_3TempGRA = Temperate natural grassland (C_3); C_3Crop = Crops (C_3); C_3BorGRA = Boreal natural grassland (C_3).

The model evaluation and parameterisation relies on flux measurements from the Integrated Carbon Observation System (ICOS) Warm Winter 2020 database, which captures the recent drought years experienced in Europe (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022). Among the 73 available sites, we kept 53 sites having a fraction of the dominant PFT of at least 50% at a 0.0125° spatial resolution using ESA-CCI land cover maps combined with Köppen-Geiger climate zone map at 0.25°. Then, we removed 7 sites for which the relative mean absolute difference (rMAD, the mean absolute difference divided by the mean of the observed variable) or the correlation between the ICOS GPP estimates and ORCHIDEE prior simulation was respectively higher than 1.25 or lower than 0.6, in 2017 or 2018 (the 2 years considered for data assimilation, see section 2.3). Indeed, we considered that these values point to model errors that cannot be corrected only through parameter optimization, but are related to structural issues in process implementation, such as missing processes to account for site disturbances.
This selection resulted in 46 sites (Figure 1, Table S1), encompassing six PFTs: Temperate evergreen needleleaf forest (TempENF, 14 sites), Temperate deciduous broadleaf forest (TempDBF, 7 sites), Boreal evergreen needleleaf forest (BorENF, 7 sites), Temperate natural C_3 grassland (C_3TempGRA, 5 sites), C_3 Crops (C_3Crop, 12 sites), and C_3 Boreal natural grassland (C_3BorGRA, 1 site).

2.2.1. Site scale data

We used *in situ* GPP and LE eddy covariance (EC)-based estimates from the ICOS Warm Winter 2020 database (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022). We considered daily data from 2017 up to 2020. GPP is estimated based on the daytime partitioning method (Lasslop et al., 2010). We applied a correction to the ICOS LE measurements to ensure energy balance closure each day, as it is the case in the ORCHIDEE model at each time step. LE was corrected using the flux-data-qaqc Python package (Volk et al., 2021) following the Bowen Ratio method (Bowen, 1926).

2.2.2. Regional scale evaluation products

We used three GPP products to evaluate the simulated GPP over Europe. First, we considered the FLUXSAT version 2.0 GPP, estimated from a light use efficiency approach using satellite data and calibrated with FLUXNET2015 (Pastorello et al., 2020) EC fluxes (Joiner et al., 2018). The second GPP product is FLUXCOM-X-BASE (Nelson et al., 2024), based on a machine learning approach using predictor variables such as remotely sensed vegetation indices and land surface temperatures from the National Aeronautics and Space Administration (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) along with meteorological variables. FLUXCOM-X-BASE also incorporates EC fluxes from various networks including FLUXNET 2015, ICOS Drought 2018, ICOS Warm Winter 2020, or Ameriflux. Then, we used the SIF-GPP product developed for the Sen4GPP project (Duveiller et al., 2023) in which GPP is estimated based on empirical linear relationships with Solar-Induced Fluorescence (SIF) estimates from the European Space Agency (ESA) Sentinel-5p Tropospheric Monitoring Instrument (TROPOMI) (Guanter et al., 2021) and calibrated per biome.
For LE, FLUXCOM-X-BASE also provides land evapotranspiration estimates (Nelson et al., 2024).

We evaluated the leaf area index (LAI) simulated in ORCHIDEE against three global LAI datasets. The two first LAI datasets (named here "PROBA-V" and "OLCI") are based on a neural network approach using top of canopy reflectances from SPOT-4 and -5, PROBA-V, and Sentinel-3 OLCI (Baret et al., 2013; Camacho et al., 2013). In addition to the reflectance data, the neural network is calibrated using LAI estimates derived from MODIS and CYCLOPES LAI products for the PROBA-V LAI dataset. Then, the neural network for the OLCI LAI dataset is trained using the PROBA-V LAI dataset. For these two datasets, LAI uncertainty is provided, which corresponds to the root mean squared deviation (RMSD) between 10-day composite and daily values. The PROBA-V dataset is available until April 2020 while the OLCI dataset is available until 2022. The last LAI dataset (named here "MODIS") is based on the reprocessing of MODIS version 6.1 LAI products MCD15A2H, MOD15A2H, and the MCD12Q1 land cover type product (Yuan et al., 2011; Lin et al., 2023).
The regional scale evaluation was conducted considering a 0.5° spatial resolution for all data products, between 2018 and 2020. The LAI products are available at a monthly timestep, while we considered 8-day averages for the GPP and LE products.

2.2.3. Regional scale ecoregion classification and drought identification

We used monthly standardized precipitation-evapotranspiration index (SPEI) (Vicente-Serrano et al., 2010) data from the SPEIbase v.2.9 provided by Beguería et al. (2023) at a 0.5° spatial resolution. SPEI is a meteorological drought index based on a water balance that includes the effect of both precipitation and temperature through potential evapotranspiration (PET), here computed using the FAO-56 Penman-Monteith method. We used SPEI to identify grid cells with drought conditions for each year over Europe, selecting the ones for which SPEI is lower than -1 (mild to extreme drought, see Table S2) for at least one month between June and August. The selected stressed areas are represented in Figure S1.
Then, we defined four ecoregions in Europe to evaluate the simulated LAI, based on a simplification of the original Köppen-Geiger climate classification at a 0.5° spatial resolution provided by Beck et al. (2023) built over the period 1991-2020. The four resulting ecoregions are referred to as "dry", "temperate", "cold", and "boreal". Their spatial distribution as well as details about the climate classification grouping are provided in Figure S2.
Finally, for projection analysis, four ecoregions have also been defined based on the same climate classification grouping. These classifications are derived from Beck et al. (2023) under the SSP370 scenario over 2070-2100, using a 2° spatial resolution, consistent with the projections performed with ORCHIDEE (Figure S3).

2.3. Data assimilation framework

To improve GPP and LE response simulated in ORCHIDEE, and in particular during drought events, we optimized the main parameters of the processes influencing these two variables using data assimilation (DA) techniques, as described in the following sections.

2.3.1. Parameter selection following sensitivity analyses

Prior to conducting DA, we ran sensitivity analyzes (SA) to select the most influential parameters for simulating GPP and LE. This preliminary step enables to limit the number of parameters included in the DA procedure, which reduces its computational cost and the risk of overfitting.
We performed SA at each selected ICOS site for GPP and LE in 2018 to identify the most important parameters during a year characterized by drought events in Europe. We used the Morris method (Morris, 1991), which is a cost efficient qualitative SA method that provides a ranking of the parameters to which GPP or LE are the most sensitive. We tested 128 parameters involved in GPP or LE representation considering a range of variation defined based on expert and physical knowledge for each parameter. Based on the SA results, we finally selected 20 parameters, most of them being PFT-dependent, mainly related to photosynthesis, stomatal conductance, biomass, soil hydrology and moisture stress vegetation response. These parameters are presented in Table S3.

2.3.2. Optimization tool

Following SA, we optimized the 20 selected parameters using the ORCHIDEE DA system (ORCHIDAS, https://orchidas.lsce.ipsl.fr/) that has been described in details in Bastrikov et al. (2018) and used in many studies (Kuppel et al., 2012; Peylin et al., 2016; Mahmud et al., 2021; MacBean et al., 2022; Bacour et al., 2023). ORCHIDAS allows finding a combination of parameter values that minimizes a cost function quantifying the mismatch between the observed and simulated variables as well as the distance between optimized parameters and their prior values, accounting for errors associated with all these components. Assuming Gaussian distributions for model-data and parameters errors, the cost function expresses as (Tarantola, 2005),

$$J(x) = \frac{1}{2} \cdot [(H_{GPP}(x) - y_{GPP})^T \cdot R_{GPP}^{-1} \cdot (H_{GPP}(x) - y_{GPP}) + (H_{LE}(x) - y_{LE})^T \cdot R_{LE}^{-1} \cdot (H_{LE}(x) - y_{LE}) + (x - x_b)^T \cdot B^{-1} \cdot (x - x_b)] \quad \text{(Eq. 2)}$$

with $y_{GPP/LE}$ the observed GPP or LE, and $H_{GPP/LE}(x)$ the corresponding simulated variable. x_b is the prior and x is the optimized vector of parameters. $R_{GPP/LE}$ and B are the prior error covariance matrices for the observations and the parameters, respectively. As in previous studies, we considered only diagonal elements for $R_{GPP/LE}$ and B, and defined the parameter error in B as 15% of the parameter range of variation while the observations errors were set to the RMSD between observed GPP and LE data and the prior model simulations (Kuppel et al., 2012; Bacour et al., 2023).
We performed DA using the genetic algorithm (GA) method (Goldberg, 1989; Haupt & Haupt, 2004; Santaren et al., 2014), which is a global search method that reduces the risk of falling into a local minima of $J(x)$ during its minimization. We ran the algorithm for 25 iterations with a population of 32 individuals to ensure its convergence.

2.3.3. Optimization experiments

Since most ORCHIDEE parameters are specified for each PFT, all DA experiments involve grouping the ICOS sites based on the dominant PFT fraction. Indeed, conducting multi-site optimizations per PFT allows us to derive a set of generic parameters for each PFT, ensuring representation of diverse conditions across multiple sites rather than specific conditions at a single site (Kuppel et al., 2012; Macbean et al., 2022). 42 out of the 46 sites were used in the optimizations, as the years used for DA (2017 and 2018) were not available for 4 sites, which were used for evaluation only (Table S1).

We conducted two successive DA experiments (Figure 2) to both support the development of a new drought-acclimated modeling scheme in ORCHIDEE (see section 3.1) and to evaluate its performance.

We performed a first optimization of the selected parameters (section 2.3.1, Table S3) by assimilating *in situ* GPP and LE data either in 2017 or in 2018 to account for different environmental conditions (2018 exhibited more pronounced soil water stress than 2017, Figure S4). Following this step, we analyzed the change in parameter values after optimization, with a specific focus on the parameter that determines the speed of stomatal closure $\alpha_{water\ stress}$ (section 2.1.2). After this step, a new definition of $\alpha_{water\ stress}$ is proposed, which incorporates vegetation potential to acclimate to long-term VPD conditions (see section 3.1), resulting in a drought-acclimated version of ORCHIDEE (ORCHIDEE-ACCLIM). This version is compared to the standard version (ORCHIDEE-STANDARD) in which $\alpha_{water\ stress}$ is a constant calibrated per PFT.

In the second DA step, *in situ* GPP and LE data were assimilated to optimize each of the two ORCHIDEE versions considering both 2017 and 2018, seeking for more generic parameter values. Indeed, we aim at obtaining combinations of optimized parameters that enhance the model performance during drought years without compromising performance in other periods.

Each optimization experiment involved assimilating GPP and LE data at daily intervals using 7-day moving averages. This approach helps to reduce the influence of high-frequency fluctuations, which are more challenging to capture and may introduce noise, while still retaining the impact of longer-term events such as droughts that develop over at least a week. This ensures that the minimization of $J(x)$ (eq. 2) focuses on significant variations rather than short-term noise.

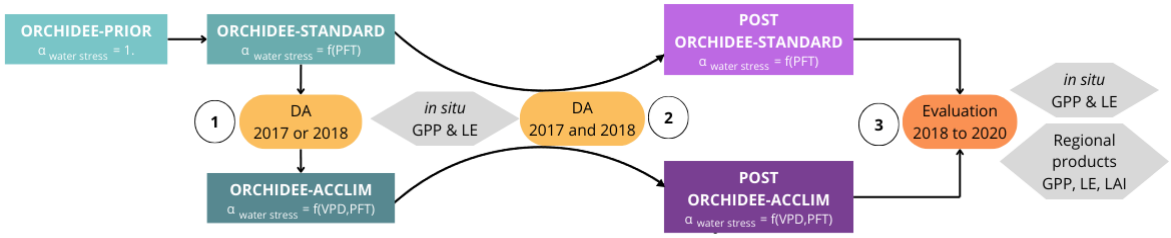

Figure 2: Overview of the DA and evaluation framework. Step 1 corresponds to the first DA experiment performed in either 2017 or 2018 that leads to the ORCHIDEE-ACCLIM model version with the redefinition of $\alpha_{water\ stress}$ as described in eq. 3. Step 2 is the second DA experiment using both years to optimize each model version (ORCHIDEE-STANDARD and ORCHIDEE-ACCLIM). The third step corresponds to the evaluation of each optimized model version against *in situ* and regional evaluation data.

To evaluate the simulated variables after DA, in addition to using 2019 and 2020 as independent years not included in the optimization, we also considered the year 2018. This inclusion allows for more sites to be evaluated at the site scale, as flux data were not available until 2020 at several ICOS Warm Winter 2020 sites.

3. Results

3.1. Defining long term VPD effect on vegetation response to soil moisture stress

The first optimization step (section 2.3.3) improved the simulated GPP RMSD by 15%/18% and LE by 28%/30% in 2017/2018 compared to ORCHIDEE-PRIOR (Table S4).
We analyzed the distribution of the optimized values of $\alpha_{water\ stress}$ (speed of stomatal closure during soil moisture stress) for each PFT depending on various environmental variables, and investigated if these optimized values could be defined as a function of environmental drivers, such as air and soil temperature, radiation, VPD over the assimilation period. To do so, several fitting functions were tested (not shown). The only significant relationship demonstrated was between mean annual VPD and $\alpha_{water\ stress}$ (Figure 3), with a coefficient of determination (R^2) of 0.92 when all other tested fitting functions and environmental drivers, including multilinear regressions, only led to a R^2 lower than 0.2. R^2 was calculated with the r2_score function from the sklearn.metrics module in Python.

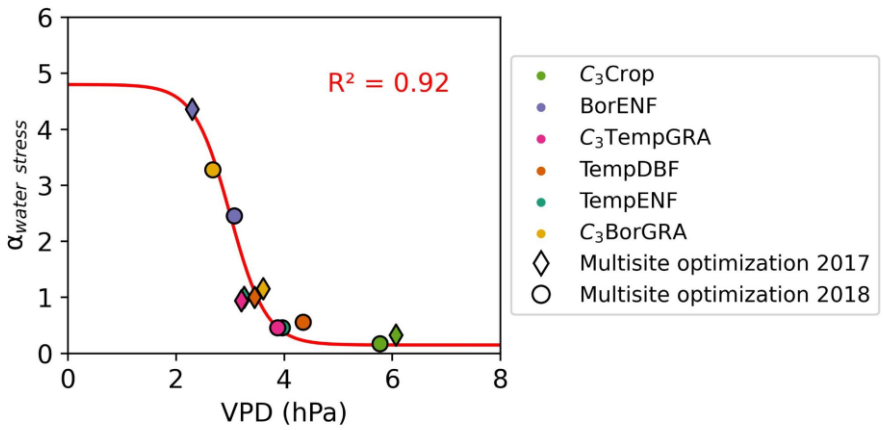

Figure 3: Optimized $\alpha_{water\ stress}$ values (-) obtained after the several multi-site assimilations versus yearly averaged VPD (hPa) over all sites grouped by dominant PFT in 2017 and 2018. The marker colors represent the dominant PFT while the marker shapes correspond to the year. The sigmoid regression function providing the best model-data fit is represented in red, with the corresponding coefficient of determination (R^2) on the top right.

Figure 3 shows the optimized values of $\alpha_{water\ stress}$ for the years 2017 and 2018 plotted against the mean annual VPD across all sites grouped by dominant PFT. Considering all PFTs, $\alpha_{water\ stress}$ decreases with increasing mean annual VPD values. Lower $\alpha_{water\ stress}$ values mean that stomata remain open for longer periods when soil moisture decreases, indicating increased vegetation resistance to droughts under long-term high VPD conditions. C_3Crop show the lowest optimized

$\alpha_{water\ stress}$ values, which are associated with the highest mean annual VPD conditions both in 2017 and 2018. For the other PFTs, different mean annual VPD values are found between the two years, with higher optimized values of $\alpha_{water\ stress}$ for the year with lowest VPD conditions. Notably, the two boreal PFTs (BorENF and C_3BorGRA) show the highest optimized $\alpha_{water\ stress}$ values. These highest $\alpha_{water\ stress}$ values illustrate that vegetation is less resistant to soil moisture stress under long-term low VPD conditions.

Across all PFTs, the best model-data fit was obtained with the following sigmoid function (with coefficients calibrated using the curve_fit function from the scipy.optimize module), which expresses the variation of $\alpha_{water\ stress}$ with rolling mean annual VPD ($VPDyear$) (hPa), yielding an R^2 of 0.92,

$$\alpha_{water\ stress} = b + \frac{(a-b)}{\left(1.+\ exp\left(VPDyear - \frac{c}{2}\right)^d\right)} \quad \text{(Eq. 3)}$$

with a = 4.8, b = 0.15, c = 6, and d = 3 four empirical coefficients.

This function was used instead of considering a constant value of $\alpha_{water\ stress}$ for each PFT (ORCHIDEE-STANDARD version) to derive a model version called hereafter ORCHIDEE-ACCLIM. The impact of $\alpha_{water\ stress}$ derived from this function for a range of rolling mean annual VPD conditions on the water stress function $f_{water\ stress}$ is illustrated in Figure 4.

The four empirical coefficients of this function were ultimately optimized in a DA experiment, along with the other model parameters (DA step 2 in Figure 2). Recognizing that various vegetation types may exhibit distinct responses to drought events despite similar environmental conditions (Schmied et al., 2023), we opted to define and optimize these four parameters for each PFT. Therefore, this new definition of $\alpha_{water\ stress}$ was initially observed across the different PFTs, and we assume that this behavior also applies within each PFT.

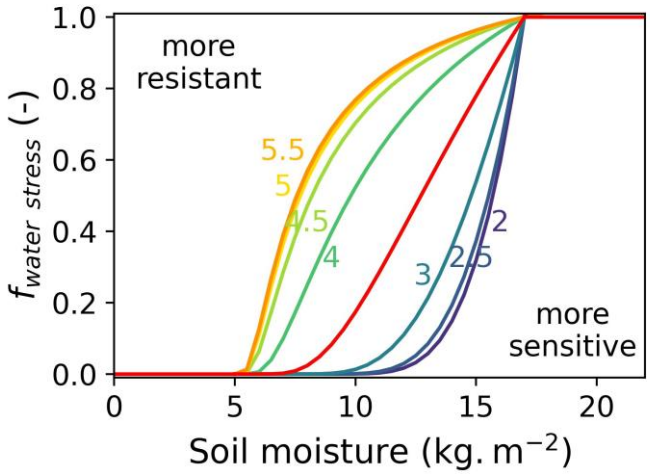

Figure 4: ORCHIDEE water stress function $f_{water\ stress}$ (-) variations with soil moisture (kg.m^{-2}), illustrated with a minimum stress for soil moisture above 18 kg.m^{-2} and a maximum stress below 5 kg.m^{-2}. The ORCHIDEE-STANDARD version is represented in red (corresponding to $\alpha_{water\ stress}$ = 1), while the ORCHIDEE-ACCLIM version is represented by all other colors associated with the VPD value (hPa) used to compute $\alpha_{water\ stress}$.

3.2. Evaluation of optimized simulated GPP and LE during drought events

We here assess the performance of the two optimized model versions (POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM) following the multi-site DA over the whole 2017-2018 period (step 2 in Figure 2), which corresponds to step 3 evaluation in Figure 2.

3.2.1. Site scale evaluation

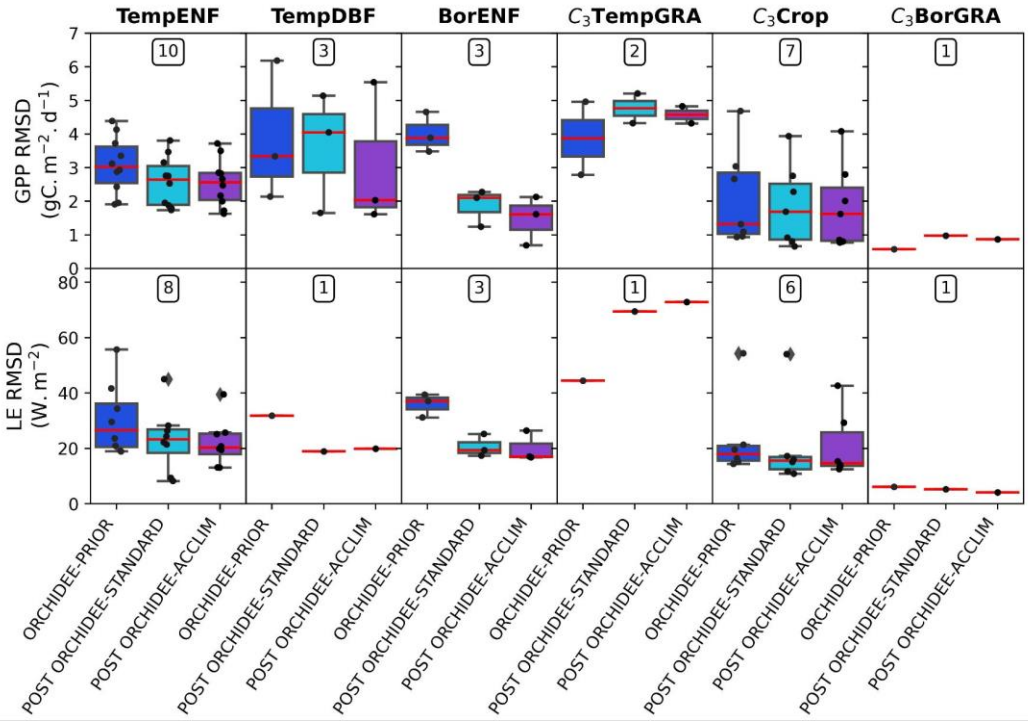

Figure 5: Evaluation of the simulated GPP (gC.m^{-2}.d^{-1}) (first row) and LE (W.m^{-2}) (second row) at all Warm Winter 2020 sites for which $f_{water\ stress}$ reaches a value below 0.85 between May and September in 2018, 2019, or 2020, for ORCHIDEE-PRIOR (dark blue), and the two optimized versions: POST ORCHIDEE-STANDARD (light blue) and POST ORCHIDEE-ACCLIM (purple). This evaluation is based on RMSD values between the daily observed and simulated variables. The number of sites included in the GPP or LE evaluation for each PFT is indicated in the top center box. Note that the number of sites can differ between the GPP and LE evaluations due to uneven data gaps.

Figure 5 shows the GPP and LE RMSD values per PFT for ORCHIDEE-PRIOR and the two optimized versions, POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM, over the period 2018-2020. The RMSD values are computed for the entire period from May to September, but only for sites where $f_{water\ stress}$ falls below 0.85 at any point during this period, which indicates significant soil moisture stress conditions. Concerning GPP, the median RMSD decreases for the two optimized model versions compared to ORCHIDEE-PRIOR for TempENF and BorENF, while only the POST ORCHIDEE-ACCLIM version improves the median RMSD for TempDBF. However, note that DA degrades the simulated GPP with the two optimized model versions for all herbaceous PFTs (C_3TempGRA, C_3BorGRA, C_3Crop) (see discussion in section 4.3). Across all PFTs, POST ORCHIDEE-ACCLIM reduces the median GPP RMSD by 18% (Table S5), while no improvement is found with

POST ORCHIDEE-STANDARD. For LE, both optimized model versions reduce the median RMSD for all PFTs except C_3TempGRA. The two optimized model versions lead to a similar reduction in the median RMSD across all PFTs by 8 to 9% (Table S5).
Interestingly, when considering all 46 sites and the entire seasonal cycles from 2018 to 2020, POST ORCHIDEE-STANDARD improves the median GPP RMSD by 10% and LE by 19% across all PFTs, while POST ORCHIDEE-ACCLIM leads to a lower improvement with 3% for GPP and 15% for LE (Table S6). Therefore, the higher performance of POST ORCHIDEE-ACCLIM in simulating GPP compared to POST ORCHIDEE-STANDARD is found specifically when evaluating sites with significant soil moisture stress conditions between May and September (Figure 5).

3.2.2. Regional scale evaluation

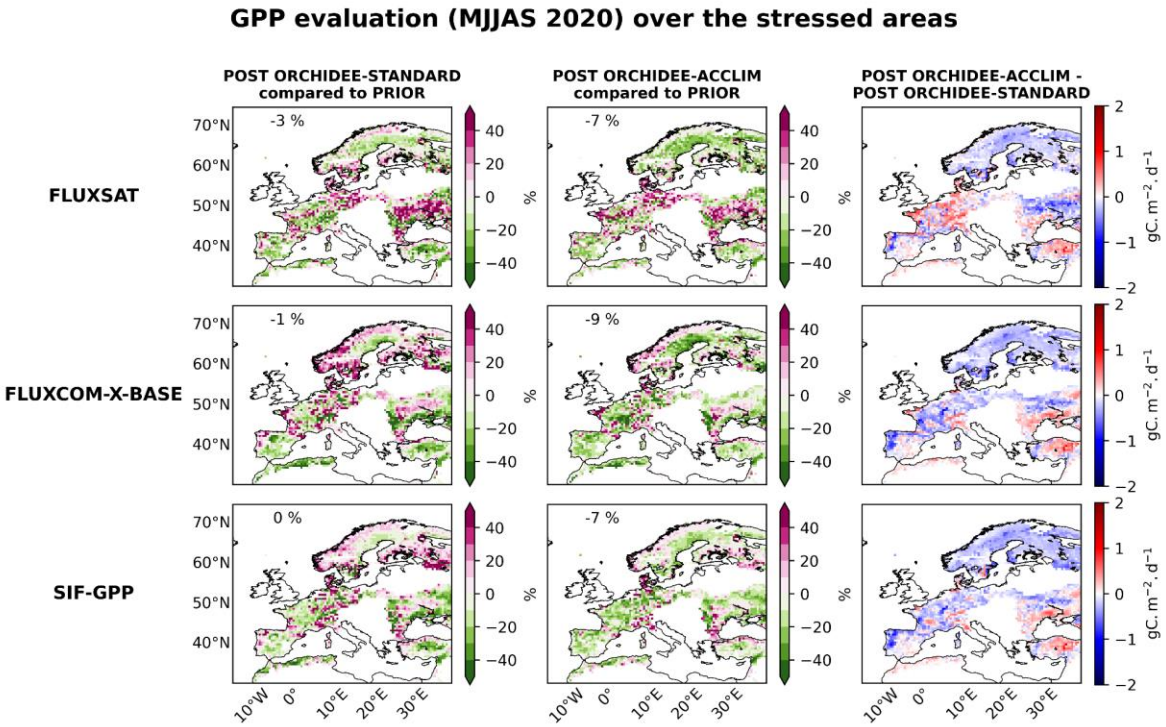

Figure 6: Maps of RMSD change (%) between the RMSD computed between the data-driven GPP and post-optimization simulated ones (for POST ORCHIDEE-STANDARD on the left column or POST ORCHIDEE-ACCLIM in the middle column), and the RMSD obtained with ORCHIDEE-PRIOR, for each evaluation product (FLUXSAT, FLUXCOM-X-BASE, SIF-GPP). The right column shows the RMSD difference ($gC.m^{-2}.d^{-1}$) between POST ORCHIDEE-ACCLIM and POST ORCHIDEE-STANDARD for each evaluation product. The selected grid cells correspond to the stressed areas (section 2.2.3), between May and September 2020. Data are 8-day averages with a 0.5° spatial resolution. For the first two columns, grid cells in green depict an improvement in the posterior simulated GPP compared to ORCHIDEE-PRIOR (reduction in RMSD), while grid cells in pink show a degradation of the simulated GPP after optimization. For these two columns, the median RMSD change (%) is provided in the top left corner.

Figure 6 presents the evaluation of the simulated GPP for POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM compared to ORCHIDEE-PRIOR against the three GPP evaluation products over

stressed areas (based on the SPEI drought index, see section 2.2.3) between May and September 2020. POST ORCHIDEE-ACCLIM achieves a more significant reduction in GPP RMSD compared to ORCHIDEE-PRIOR than POST ORCHIDEE-STANDARD. The median improvement in GPP RMSD for POST ORCHIDEE-ACCLIM ranges between 7% and 9%, over all grid cells considered and depending on the evaluation product, whereas POST ORCHIDEE-STANDARD shows an improvement of no more than 3% in 2020. Higher performance of POST ORCHIDEE-ACCLIM over POST ORCHIDEE-STANDARD is also observed in 2018 and 2019 (Figures S5 and S6). Across 2018, 2019, and 2020, considering all evaluation products, POST ORCHIDEE-STANDARD results in an equivalent number of degraded and improved grid cells, while POST ORCHIDEE-ACCLIM improves the simulated GPP for 61% of the grid cells in stressed areas (Table S7). However, regions such as the Pyrenees, the Alps, and Northern Norway exhibit GPP degradation in both optimized versions. Then, despite the general trend of lower GPP RMSD values with POST ORCHIDEE-ACCLIM compared to POST ORCHIDEE-STANDARD across most grid cells, exceptions include regions like the Northwestern part of France and Southeastern Europe.

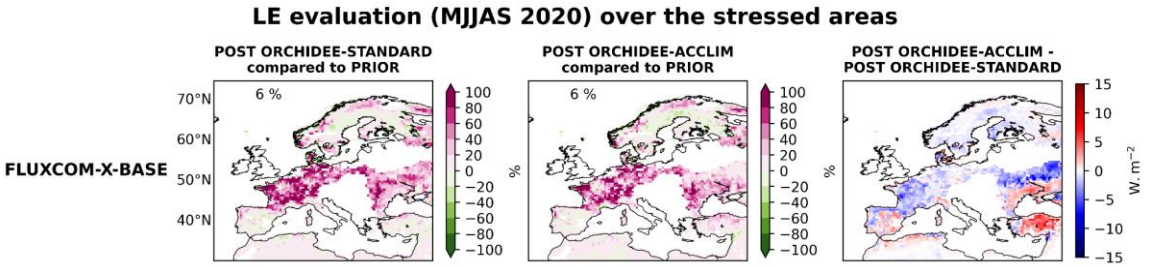

Figure 7: Maps of RMSD change (%) between the RMSD computed between FLUXCOM-X-BASE and post-optimization simulated LE (for POST ORCHIDEE-STANDARD on the left column or POST ORCHIDEE-ACCLIM in the middle column), and the RMSD computed between FLUXCOM-X-BASE and ORCHIDEE-PRIOR LE. The right column shows the RMSD difference ($W.m^{-2}$) between POST ORCHIDEE-ACCLIM and POST ORCHIDEE-STANDARD for each evaluation product. The selected grid cells correspond to the stressed areas (section 2.2.3), between May and September 2020. Data are 8-day averages with a 0.5° spatial resolution. For the first two columns, grid cells in green depict an improvement in the posterior simulated LE compared to ORCHIDEE-PRIOR (reduction in RMSD), while grid cells in pink show a degradation of the simulated LE after optimization. For these two columns, the median RMSD change (%) is provided in the top left corner.

Figure 7 presents a similar evaluation against FLUXCOM-X-BASE LE. Both optimized models exhibit an equivalent median degradation of 6% in simulated LE compared to ORCHIDEE-PRIOR. This degradation is observed across nearly all stressed areas, particularly in Western and Eastern Europe and Northern Norway, with a slight improvement mainly in Sweden. Between the two model versions, POST ORCHIDEE-ACCLIM demonstrates higher performance than POST ORCHIDEE-STANDARD, showing a greater reduction in RMSD across most grid cells, except in Southeastern Europe and Spain. Similar patterns were observed in 2018 and 2019, with POST ORCHIDEE-ACCLIM showing a slightly lower median LE degradation than POST ORCHIDEE-STANDARD compared to ORCHIDEE-PRIOR (Figures S7 and S8 and Table S8). Note that the general degradation in the simulated LE after the two optimizations contrasts with the overall improvement in the simulated GPP for POST ORCHIDEE-ACCLIM (Figure 6).

3.3. Additional validation against LAI over European ecoregions

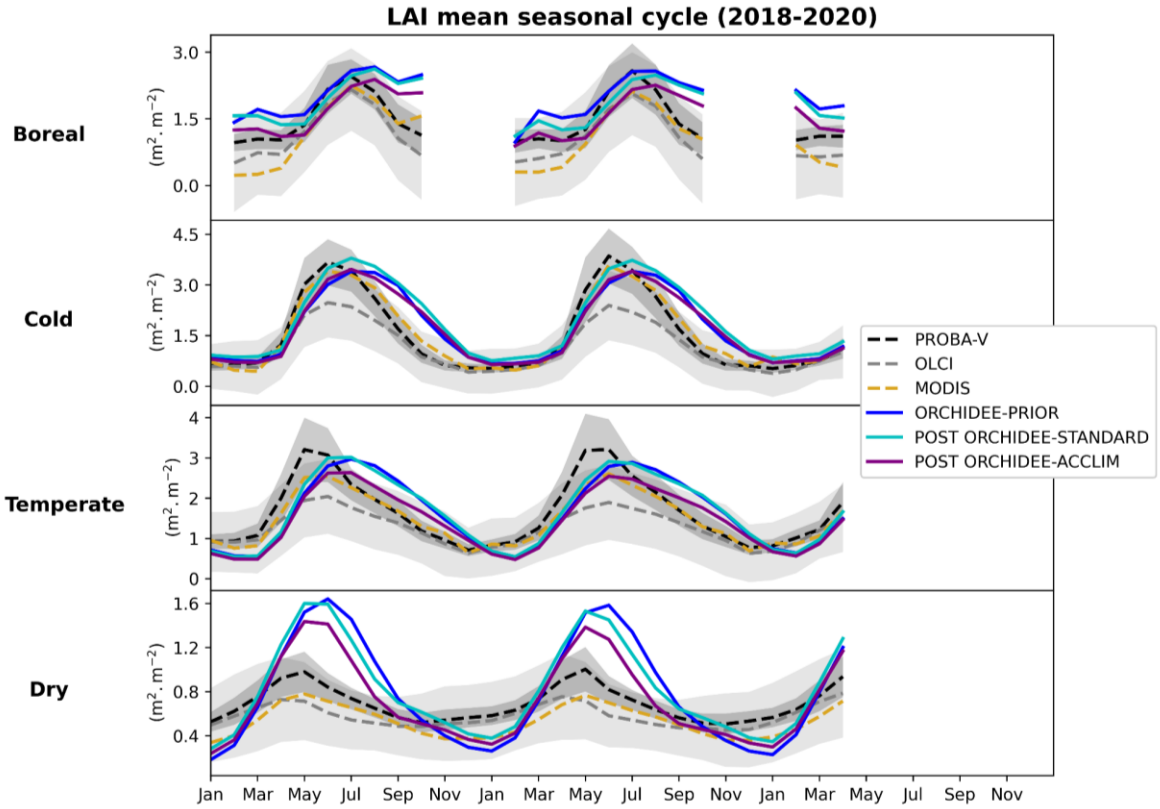

Figure 8: Mean seasonal cycle of LAI ($m^2.m^{-2}$) for ORCHIDEE simulations (solid lines) compared to three evaluation products (MODIS, PROBA-V, OLCI) (dotted lines) over the four ecoregions from 2018 to 2020. Data are monthly averages with a 0.5° spatial resolution. For all evaluation products and ORCHIDEE simulations, grid cells where the evaluation products have missing data have been masked. The gray shaded areas represent the uncertainty in the PROBA-V and OLCI LAI products, for which uncertainty is provided (see section 2.2.2).

Figure 8 shows the mean seasonal cycle of LAI simulated by ORCHIDEE-PRIOR and the two optimized versions, evaluated against three different evaluation products for each ecoregion. Across all ecoregions, POST ORCHIDEE-ACCLIM simulates a lower LAI than POST ORCHIDEE-STANDARD. Both optimized versions reduce the simulated LAI in the boreal and dry ecoregions compared to ORCHIDEE-PRIOR, with POST ORCHIDEE-ACCLIM also reducing LAI in the temperate ecoregion. In contrast, POST ORCHIDEE-STANDARD increases LAI in the cold ecoregion compared to ORCHIDEE-PRIOR.

A notable spread in LAI estimates is observed between the three evaluation products, highlighting significant uncertainties. Despite this, the simulated peak of LAI for the boreal, cold, and temperate ecoregions falls within the range of the evaluation products. However, in the dry ecoregion, the simulated LAI is significantly overestimated, even though POST ORCHIDEE-ACCLIM reduces this overestimation.

All simulated LAI seasonal cycles, except in the dry ecoregion, exhibit a delay of 1 to 2 months in the summer LAI maximum compared to the evaluation products, all of them showing an agreement in the seasonal LAI phase. This results in a systematic overestimation of LAI during mid-

summer and autumn. In the dry ecoregion, while the LAI maximum in ORCHIDEE-PRIOR was also delayed by a month compared to the evaluation product, the optimization succeeded in correcting this phase issue for both POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM.

The two optimized versions show the strongest reduction in LAI RMSD across the 3 evaluation products in the boreal ecoregion, with 25 to 36% for POST ORCHIDEE-ACCLIM and 13 to 16% for POST ORCHIDEE-STANDARD (Table S9). Then, while POST ORCHIDEE-ACCLIM improves the RMSD by 13 to 18% in the dry ecoregion, and by 8 to 17% in the temperate ecoregion, RMSD changes with POST ORCHIDEE-STANDARD range from an increase of 3% to a reduction of 4% in these two ecoregions. The lowest LAI RMSD improvement is found for the cold ecoregion. However, while POST ORCHIDEE-ACCLIM improves it by 6 to 10%, POST ORCHIDEE-STANDARD leads to a degradation by up to 16%.

Note that the significant improvement in the simulated LAI with POST ORCHIDEE-ACCLIM could be attributed to changes in parameter values affecting carbon allocation and pool growth, such as the specific leaf area (*SLA*) (Table S3). Additionally, the improvement in simulated GPP may also play a role, due to the coupling of the carbon assimilation and biomass allocation schemes in ORCHIDEE.

3.4. Projections over Europe

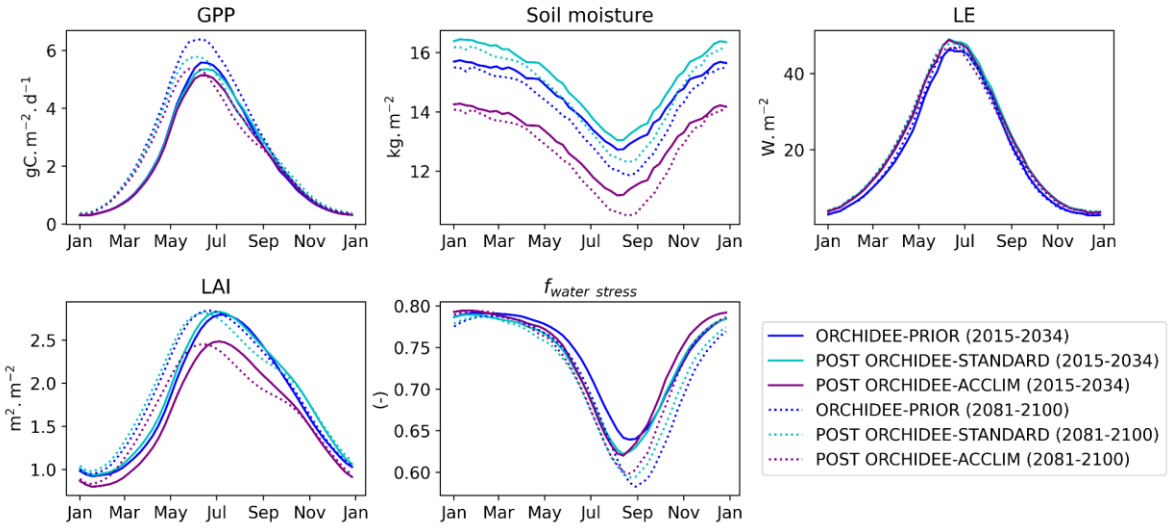

Figure 9: Mean seasonal cycles of simulated GPP (gC.m^{-2}.d^{-1}), LE (W.m^{-2}), LAI (m^2.m^{-2}), $f_{water\ stress}$ (-), and surface soil moisture (kg.m^{-2}) over Europe, projected under the SSP370 scenario for the period 2015-2034 and 2081-2100. Data are 8-day averages with a 2° spatial resolution.

Figure 9 shows the mean seasonal cycles of simulated GPP, LE, LAI, $f_{water\ stress}$, and surface soil moisture over Europe for 20 years at the beginning (2015-2034) and at the end of the century (2081-2100) under the SSP370 scenario. An increase in the seasonal amplitude of GPP is observed across all model versions by the end of the century compared to 2015-2034, with increases of 14% for ORCHIDEE PRIOR, 8% for POST ORCHIDEE-STANDARD, and 4% for POST ORCHIDEE-ACCLIM. The most pronounced increase occurs in the boreal ecoregion (Figure S9). In addition, the mean seasonal cycle of GPP shows an earlier seasonal maximum by the end of the century, by approximately 15 days for the two optimized versions. A similar shift is also found in the seasonal

cycle of LAI. These changes are particularly notable in the dry and temperate ecoregions (Figures S9 and S10).

Conversely, LE mean seasonal cycle, driven by the seasonal cycle of plant transpiration (not shown), shows no significant changes in its seasonal phase or amplitude between 2015-2034 and 2081-2100, nor between the different model versions.

Then, all model versions also project a decrease in surface soil moisture at the end of the century compared to 2015-2034, particularly during the peak of minimum soil moisture in August. Therefore, all versions predict stronger soil moisture stress (i.e., lower $f_{water\ stress}$). However, while the three models show similar increases in soil moisture stress from spring until early August, POST ORCHIDEE-ACCLIM reduces the simulated stress levels afterwards compared to ORCHIDEE PRIOR and POST ORCHIDEE-STANDARD, resulting in the lowest stress level after the summer peak of minimum soil moisture. Notably, POST ORCHIDEE-ACCLIM achieves this despite simulating the lowest surface soil moisture.

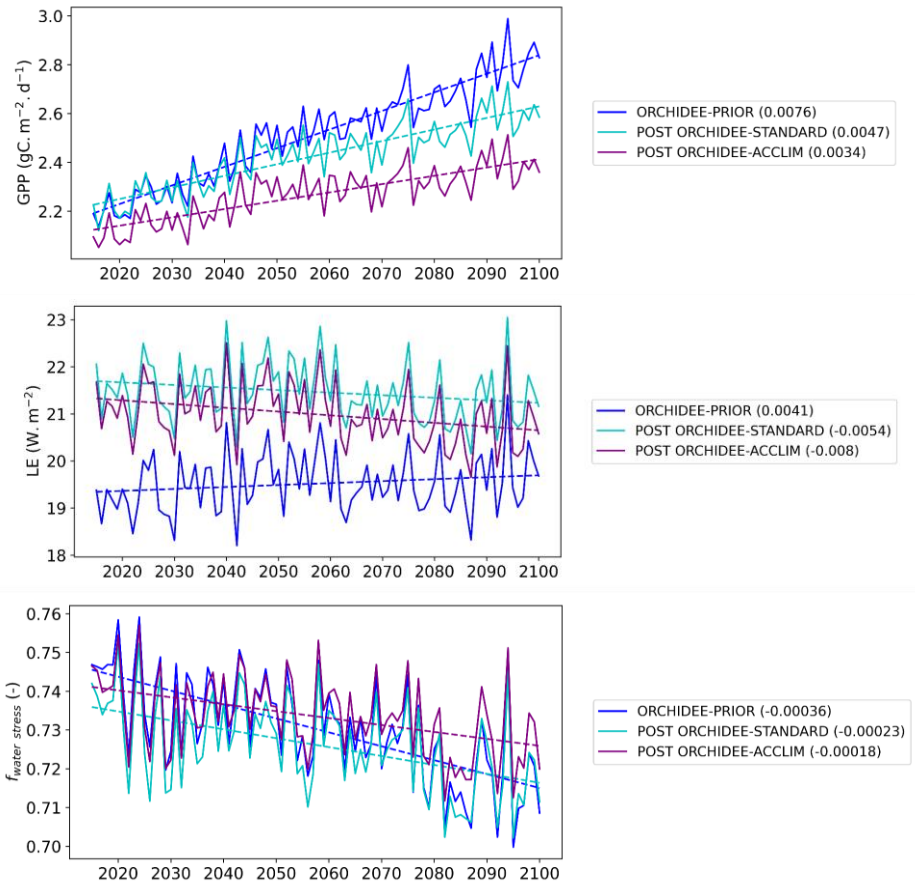

Figure 10: Projected variations of GPP (gC.m^{-2}.d^{-1}) (top), LE (W.m^{-2}) (middle), and $f_{water\ stress}$ (-) (bottom) over Europe from 2015 to 2100 under the SSP370 scenario. The mean annual values are represented in solid lines, and the corresponding linear regressions are in dotted lines with the regression coefficient values (slope) given in the legend for each model version.

Figure 10 illustrates the temporal variations and trends in mean annual GPP, LE, and $f_{water\ stress}$ across Europe from 2015 to 2100. Consistent with the seasonal cycle findings in Figure 9, GPP is projected to increase until 2100, with the fastest rate observed in the ORCHIDEE PRIOR model

(0.0076 $gC.m^{-2}.d^{-1}.y^{-1}$) and the slowest in the POST ORCHIDEE-ACCLIM model (0.0034 $gC.m^{-2}.d^{-1}.y^{-1}$). This increase in GPP can be attributed to the CO_2 fertilization effect. Indeed, under the SSP370 scenario, atmospheric CO_2 levels are projected to rise by an average of 5.46 $ppm.y^{-1}$ between 2015 and 2100, resulting in more than a doubling of the atmospheric CO_2 concentration by 2100 compared to current levels (Figure S11). LAI also shows a slightly positive trend (Figure S12), with comparisons between model versions similar to those for GPP: the ORCHIDEE PRIOR model exhibits the highest increase rate (0.0017 $m^2.m^{-2}.y^{-1}$), while the POST ORCHIDEE-ACCLIM model has the slowest rate (0.0010 $m^2.m^{-2}.y^{-1}$). In contrast, the projected LE shows negligible trends between 2015 and 2100.

Then, all model versions depict an increasing trend in soil moisture stress until the end of the century. However, the POST ORCHIDEE-ACCLIM version projects the smallest trend. Indeed, the soil moisture stress trend in POST ORCHIDEE-ACCLIM is 22% lower than in POST ORCHIDEE-STANDARD and 50% lower than in ORCHIDEE-PRIOR. Therefore, despite the optimized versions initially showing higher soil moisture stress than the prior version at the beginning of the century, POST ORCHIDEE-ACCLIM simulates the lowest soil moisture stress among the three versions from 2040 onwards.

4. Discussion

4.1. Accounting for physiological acclimation mechanisms in vegetation response to drought events in LSMs

This study evidences the importance of better accounting for vegetation's potential to acclimate to environmental drivers, such as VPD, in LSMs during drought events. This DA approach, considering years characterized by drought conditions in Europe, has led to a new definition of the speed of stomatal closure during soil moisture stress, incorporating vegetation acclimation to long-term VPD conditions (eq. 3). The acclimatation parameterization is consistent with the literature: Grossiord et al. (2017) evidenced stomatal acclimation to elevated VPD in some tree species in semi-arid regions during soil drying periods. Other experimental studies have also demonstrated that stomatal conductance can acclimate to long-term VPD exposure (i.e. weeks to months) (Nejad & van Meeteren 2008; Sermons et al., 2012), with high VPD conditions enhancing stomatal conductance and plant transpiration (Marchin et al., 2016). This acclimation response contrasts with the expected decrease in stomatal conductance with increasing VPD, emphasizing the need to differentiate between the short-term response to VPD and the effects of long-term exposure on stomatal behavior in LSMs. As a result, some plant traits defined as parameters in LSMs could be allowed to vary over time to account for vegetation's acclimation potential to environmental conditions.

Several mechanisms behind stomatal acclimation to long term VPD have been suggested. Marchin et al. (2016) proposed that the phytohormone abscisic acid (ABA), which plays a crucial role in stomatal closure, is involved. Stomatal acclimation could be explained by changes in the sensitivity of ABA biosynthesis to VPD or in the responsiveness of stomata to ABA. This acclimation can allow plants to maintain carbon uptake under high VPD conditions despite increased water loss (Wang & Kellomaki, 1997; Kutsch et al., 2001; Herbst et al., 2008). Additionally, stomatal acclimation could help cooling the leaf through transpiration (Nobel, 1974) and maintaining leaf temperature within

an optimal range (Mahan & Upchurch, 1988). This cooling effect of transpiration occurs because the energy that would otherwise increase the leaf's temperature is used in converting liquid water into water vapor (Grossiord et al., 2020).

This study focuses on stomatal acclimation to long-term VPD under soil moisture stress conditions. However, further investigation is needed to understand vegetation's potential to acclimate to long-term soil moisture conditions in response to drought events. Indeed, the effects of these two drivers are difficult to disentangle as VPD and soil moisture are coupled over extended periods (Novick et al., 2024). In LSMs, we could hypothesize that parameters such as the soil moisture threshold above which there is no water stress and the one below which stomata close could acclimate to long-term soil moisture conditions or to the frequency and intensity of drought events. These parameters determine when vegetation begins to respond to a drought event and reach maximum stress, which could be refined as functions incorporating acclimation mechanisms, such as to account for ecological stress memory (Gessler et al., 2020). However, in this study using ORCHIDEE, no relationship was found between θ_{WP} (soil moisture at wilting point) or $p_{max} \cdot \theta_{FC}$ (soil moisture above which there is no stress) and long-term environmental drivers (not shown) when each year (2017 or 2018) was optimized independently. Currently, these parameters depend solely on soil texture type, but future work could refine them to investigate potential acclimation to specific climate and/or soil variables.

4.2. Link between the acclimation function and the isohydricity concept

The $\alpha_{water\ stress}$ coefficient, which defines the speed of stomatal closure during drought events (eq. 2), can be linked to the concept of isohydricity, describing plant water stress strategies (Hochberg et al., 2018). Isohydric species close their stomata when soil water potential drops to preserve leaf water potential, risking carbon starvation, while anisohydric species keep their stomata open, risking hydraulic failure (Martínez-Vilalta et al., 2014). In this study, high $\alpha_{water\ stress}$ values indicate a rapid reduction in stomatal conductance when soil moisture decreases below a threshold marking the onset of a drought event ($p_{max} \cdot \theta_{FC}$), indicative of an isohydric strategy (Figure 4). Conversely, low $\alpha_{water\ stress}$ values suggest that stomata remain open longer as soil moisture declines, reflecting an anisohydric strategy. Following the first DA phase, C_3Crop showed low optimized $\alpha_{water\ stress}$ values (Figure 3), aligning with findings by Li et al. (2017) that crops exhibit more anisohydric behavior based on vegetation optical depth data. In contrast, boreal biomes (BorENF and C_3BorGRA) demonstrated high optimized $\alpha_{water\ stress}$ values, indicating a more isohydric behavior.

In addition to the biome type dependency in stomatal response to droughts, the acclimation function introduces a dependency of $\alpha_{water\ stress}$ on long-term VPD, allowing vegetation responses to drought to vary within the same PFT depending on the local environmental conditions. Novick et al. (2019) emphasized that the degree of isohydricity alone is insufficient to assess stomatal behavior in response to soil moisture decrease, as VPD also influences the relationship between leaf and soil water potentials. Moreover, several studies have shown that plants of the same species can switch between isohydric and anisohydric behavior depending on environmental conditions (Guo et al., 2020), including their growing conditions (Hochberg et al., 2018; Feng et al., 2019). Therefore, defining $\alpha_{water\ stress}$ as a constant dependent solely on biome type would fail to

account for the impact of environmental conditions on the speed of stomatal closure during droughts.

4.3. Representativeness of the observation data and DA limitations

DA performance is limited by uncertainties in the assimilated variables. For instance, uncertainties in *in situ* GPP estimates arise from the partitioning method (Tramontana et al., 2020). Similarly, uncertainty exists in the *in situ* LE data obtained by eddy covariance, which have been corrected using the Bowen ratio method to ensure energy balance closure, as is the case in ORCHIDEE (section 2.2.1). The simulated GPP improves following DA when evaluated both at the site (Figure 5) and the regional (Figures 6, S5, and S6) scales during soil moisture stress. However, while the site scale evaluation shows an improvement in the median LE RMSD by 8% to 9% across all PFTs after DA compared to ORCHIDEE-PRIOR (Figure 5, Table S5), there is a median degradation between 6% and 12% over the stressed areas when evaluated against FLUXCOM-X-BASE LE at the regional scale (Figures 7, S7, and S8). FLUXCOM-X-BASE's upscaling of *in situ* fluxes integrates LE data from the ICOS Warm Winter 2020 database, but no energy balance correction was applied for this global product (Nelson et al., 2024). Therefore, different assumptions between the LE data used for assimilation and evaluation at the site scale and those used for regional scale evaluation could explain the diverging DA performances between these scales.
Uncertainties in the data-driven products considered for the regional scale evaluation also limit the assessment of the simulated variables in ORCHIDEE. For GPP, using three different global evaluation products (FLUXCOM-X-BASE, FLUXSAT, and SIF-GPP) allows accounting for uncertainties related to various estimation methods and underlying assumptions. This approach also identifies discrepancies between different estimates in GPP seasonal cycles over stressed areas in Europe. While FLUXCOM-X-BASE and SIF-GPP show an asymmetric GPP seasonal cycle around the maximum in late June, FLUXSAT presents a symmetrical GPP seasonal cycle around the peak in early July (Figure S13). Therefore, FLUXSAT appears unable to accurately capture the GPP summer drawdown over the stressed areas, making it less suitable for evaluating the simulated GPP during drought years. Notably, across the three evaluation products, the lowest improvement in GPP using ORCHIDEE-ACCLIM is obtained when using FLUXSAT for assessment over the stressed areas (Figures 6, S5, and S6). Regarding the SIF-GPP product, previous studies have highlighted the high performance of SIF to track the impacts of drought stress on plant physiology at large scales, surpassing traditional vegetation indices such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) (Cao et al., 2021; Song et al., 2018; Sun et al., 2015). While DA improves the simulated GPP and LE over the whole period considered in the assimilation, the main evaluation in this study focuses on conditions with significant soil moisture stress and uses two years of independent data (not included in the assimilation). This focus explains some observed degradations in the simulated variables compared to ORCHIDEE-PRIOR for the soil moisture stress periods. The model may therefore lack the genericity needed to apply the optimized parameter values to independent years, potentially highlighting overfitting during the assimilation period. Then, the contrasting results between tree and herbaceous PFTs at the site scale (Figure 5, Table S5) could be related to the specificities of herbaceous structure, phenology, and management. Thus, structural model errors related to herbaceous representation could be reduced by using specific ORCHIDEE versions that explicitly represent these specificities. For

example, the ORCHIDEE-GM version incorporates a grassland management module that includes cutting, grazing, and fertilization (Chang et al., 2013, 2015). For croplands, the ORCHIDEE-CROP version includes a module for better accounting of crop phenology and management practices (Wu et al., 2016).

Finally, specific regions with no improvement in the simulated fluxes following DA can highlight limitations in the DA approach, related to the representativeness of the sites selected for the assimilation. Notably, in Northern Norway, the simulated GPP is degraded compared to ORCHIDEE-PRIOR in both optimized versions across all evaluation products (Figure 6). Boreal broadleaf summergreen forests and C_3BorGRA are prevalent in this region, but the boreal broadleaf summergreen forest PFT was not optimized since no representative ICOS Warm Winter 2020 site was selected in this study. Similarly, a degradation of the simulated GPP compared to ORCHIDEE-PRIOR is observed over the Alps and the Pyrenees, possibly due to the unique characteristics of these high-altitude zones, which complicate accurate simulation of vegetation functioning, or due to the meteorological forcing having too coarse a resolution to accurately represent the climate over these areas.

4.4. Additional challenges in representing vegetation response to drought events in LSMs

In addition to stomatal regulation mechanisms, vegetation response to droughts also involves non-stomatal limitations that can coordinate with stomatal limitations (Flexas et al., 2008; Gago et al., 2016) and should be considered in LSMs. Several studies have reported a decrease in photosynthetic activity under soil moisture stress (Keenan et al., 2010; Gourlez de la Motte et al., 2020), which could be related to a down-regulation of the Rubisco enzyme (Flexas et al., 2004; Sugiura et al., 2020), a decrease in carbon demand (Fatichi et al., 2014), or a reduced mesophyll conductance (Flexas et al., 2012). In ORCHIDEE, the same soil moisture stress response is applied to the stomatal conductance, mesophyll conductance, and photosynthetic rate (section 2.1.2), assuming full coordination between the three. A future refinement of the model could be considered by distinguishing between specific responses. Optimization approaches describing non-stomatal limitations have been proposed (Dewar et al., 2018), however, their implementation in LSMs is challenging due to limited understanding of non-stomatal limitation dynamics (Yang et al., 2019).

Vegetation phenological and anatomical changes under prolonged soil moisture stress are also essential to accurately represent in LSMs. Droughts can induce leaf senescence, reducing leaf area (Hochberg et al., 2017; Schuldt et al., 2020), which limits plant water loss but also carbon uptake required to rebuild canopy structure after drought stress. Sustained droughts can also lead to hydraulic failure (Salmon et al., 2015) and tree mortality (Blackman et al., 2016).

Then, in addition to physiological acclimation such as the stomatal acclimation function to long term VPD conditions described in this study, anatomical acclimation mechanisms have also been reported in response to droughts. Anatomical adjustments influencing stomatal conductance, such as changes in stomatal size and density, were found in some plant species in response to long term high VPD exposure (Du et al., 2020). Acclimation of the plant hydraulic system to high VPD conditions has also been reported, such as through increasing secondary wall thickness (Hacke et al., 2001; Pittermann et al., 2006) or larger roots (Du et al., 2020). Plants can also respond

to drought by developing deeper roots to access soil moisture when topsoil layers are depleted (Chaves et al., 2002).
Finally, an integrated assessment of vegetation response to soil moisture stress would require an accurate representation of drought associated risks in LSMs, such as fires and insect outbreaks (Williams et al., 2016; Allen et al., 2015).

4.5. Projected impact of acclimation

With the projected VPD increasing trend of 0.04 $hPa.y^{-1}$ under the SSP370 scenario between 2015 and 2100 over Europe (Figure S14), the acclimation function in the POST ORCHIDEE-ACCLIM version is expected to enhance vegetation resistance to soil moisture decrease compared to the other model versions. Indeed, POST ORCHIDEE-ACCLIM simulates a lower average soil moisture stress over Europe at the end of the century compared to the two other model versions (Figures 9 and 10). This reduced soil moisture stress would typically be expected to increase vegetation CO_2 and H_2O exchanges, thereby enhancing GPP and LE. However, this anticipated increase during soil moisture stress periods is moderated by the optimization of parameters that influence the mean seasonal cycle of GPP and LE, not just during drought events. For example, the optimization reduces *SLA* for 5 out of the 6 PFTs represented in this study (BorENF, TempDBF, TempENF, C_3BorGRA, C_3TempGRA) in POST ORCHIDEE-ACCLIM, while *SLA* decreases for 2 PFTs only (BorENF, C_3BorGRA) in POST ORCHIDEE-STANDARD (not shown). In ORCHIDEE, LAI is computed from the simulated leaf biomass pool multiplied by *SLA*. Therefore, the decrease in *SLA* can contribute to the lowest LAI and GPP in POST ORCHIDEE-ACCLIM.
In addition to the focus on acclimation to long-term VPD in vegetation response to soil moisture stress, the projections underscore the significance of another factor to which photosynthesis can acclimate. Indeed, the comparison of the GPP seasonal cycles between the beginning and the end of the century indicates a projected shift in the seasonal phase towards earlier in the year by century's end, while no such change is found for LE (Figure 9). This shift could be linked to changes in Vcmax, the maximum rate of carboxylation, which is crucial for photosynthesis. Vcmax responds to temperature by increasing up to an optimal point, beyond which it decreases. The mean air temperature in Europe is projected to rise by 0.07 $°C.y^{-1}$ between 2015 and 2100 (Figure S14). Consequently, higher air temperatures can cause Vcmax to get closer to its optimum temperature earlier in the spring (Figure S15), then declining due to soil moisture limitations (Figure 9). Accounting for Vcmax acclimation to plant growth temperature is critical for GPP projections. The ORCHIDEE model incorporates this acclimation based on Kattge & Knorr (2007). Using equation 2 from their study, the optimum temperature for Vcmax is projected to increase at a rate of 0.036 $°C.y^{-1}$ for C_3 plants (Figure S16).
Finally, note that the projection ability to simulate future GPP and LE dynamics is limited by the absence of coupling between the land surface and the atmosphere, as ORCHIDEE is not run in coupled mode with an atmospheric circulation model. Running the model in coupled mode would improve the accounting for land-atmosphere interactions and feedback. For instance, a reduction in LE affects the energy and water budgets, altering the atmospheric state, which can in turn influence stomatal conductance.

5. Conclusion

To enhance the accuracy of projected vegetation responses to climate change, this study highlights the importance of incorporating acclimation mechanisms in LSMs. This work specifically addresses drought events, revealing that refining vegetation physiological response to soil moisture stress can be achieved by considering the influence of long-term VPD conditions on the speed of stomatal closure.

Under soil moisture stress conditions, integrating long-term VPD acclimation in the ORCHIDEE LSM improved the simulated GPP by 18% on sites when evaluated against *in situ* observations following DA. In contrast, a response based solely on vegetation type did not improve the simulated GPP following DA. At the regional scale in Europe, areas experiencing soil moisture stress saw a 7% to 9% improvement in simulated GPP in 2020 with the acclimation-based response, compared to a maximum of 3% improvement with the biome-type dependency response. Additionally, only the parameter optimization with the model version including long-term VPD acclimation improved the simulated LAI over Europe. However, the two model versions performed similarly in simulating LE at both site and regional scales.

Given the projected global increase in VPD alongside rising air temperatures, future projections under the SSP370 scenario until 2100 were conducted. As anticipated, the model accounting for vegetation acclimation to long-term VPD conditions during droughts projected the slowest increase in soil moisture stress levels. These projections also indicated an increase in GPP over Europe by the end of the century, with an earlier seasonal phase.

In conclusion, this study compared an LSM version that solely accounts for a biome-dependent vegetation response to drought with a version incorporating vegetation potential to acclimate to long-term environmental changes, such as VPD. The differing simulated responses between these model versions highlight the necessity of integrating acclimation mechanisms in LSMs for more accurate simulations and climate projections.

This new vegetation response to soil moisture stress that accounts for acclimation to long-term VPD could be further evaluated using tree ring width data, which provides insights into GPP (Babst et al., 2014; Tei et al., 2019). This assessment would determine if incorporating long-term VPD acclimation improves the correlation between relative changes in tree ring width and simulated GPP.

In future research, acclimation mechanisms that influence stomatal responses to soil moisture stress could be explored in LSMs by incorporating a continuous representation of water potential from soil to leaves through roots and trunk (referred to as hydraulic architecture). This approach more effectively accounts for plant structure, such as height and water storage.

Acknowledgements

The authors thank Nicolas Vuichard for formatting and sharing the meteorological forcing data for the ICOS Warm Winter 2020 database. We are grateful to Vladislav Bastrikov for providing the ORCHIDAS code for data assimilation and the forcing files used for the projections. We also acknowledge Nina Raoult for providing the code to conduct sensitivity analyses. We are very

grateful to Stefan Kern for providing the PROBA-V and OLCI LAI datasets extracted over Europe. We also thank Luis Guanter for providing the SIF-GPP product from the Sen4GPP project.

SPOT/PROBA-V: The product was generated by the land service of Copernicus, the Earth Observation programme of the European Commission. The research leading to the current version of the LAI product has received funding from various European Commission Research and Technical Development programmes. The product is based on SPOT/VGT 1km data (© CNES) / PROBA-V 1km data (© ESA and distributed by VITO) [when using the coarse-resolution data: provided on 0.5 degree x 0.5 degree plate carrée grid] [last access date: August 26, 2020] in netCDF file format by the Integrated Climate Data Center (ICDC) University of Hamburg, Hamburg, Germany.
PROBA-V/S3-OLCI: The product was generated by the global component of the Land Service of Copernicus, the Earth Observation program of the European Commission. The research leading to the current version of the product has received funding from various European Commission Research and Technical Development programs. The product is based on PROBA-V and Sentinel-3 data (© ESA) [when using the coarse-resolution data: provided on 0.5 degree x 0.5 degree plate carrée grid] [last access: May 3 2023] in netCDF file format by the Integrated Climate Data Center (ICDC) University of Hamburg, Hamburg, Germany.

Open Research (availability statement)

The ICOS data used for site simulation, data assimilation, and evaluation can be downloaded from the ICOS Carbon Portal at https://doi.org/10.18160/2G60-ZHAK, licensed under the Creative Commons Attribution 4.0 International licence (CC BY 4.0). The products used for regional evaluation are available from the NASA Goddard Space Flight Center at https://avdc.gsfc.nasa.gov for FLUXSAT, from the ICOS Carbon Portal at https://doi.org/10.18160/5NZG-JMJE for FLUXCOM-X-BASE, and at https://zenodo.org/records/8339044 for SIF-GPP. The LAI dataset for the MODIS product can be downloaded at https://opendap.4tu.nl/thredds/catalog/data2/fig/21858717/1/catalog.html, while the PROBA-V and OLCI products are available at https://www.cen.uni-hamburg.de/en/icdc/data/land/lai-spot-probav.html#kontakt. The SPEI database is freely available at https://digital.csic.es/handle/10261/332007. Then, the Köppen-Geiger classification maps can be downloaded at https://www.gloh2o.org/koppen/.
Figures were made using Python programming language version 3 (Van Rossum and Drake, 2009), available at https://www.python.org. Figure 1 uses the "shadedrelief" map as background (http://www.shadedrelief.com). The sigmoid function describing $\alpha_{water\ stress}$ in Figure 3 was fitted using the curve_fit function from the scipy.optimize module, and the associated coefficient of determination (R^2) was computed with the r2_score function from the sklearn.metrics module.

References

Allen, C. D., Breshears, D. D., & McDowell, N. G. (2015). On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere*, *6*(8). https://doi.org/10.1890/ES15-00203.1

Babst, F., Bouriaud, O., Papale, D., Gielen, B., Janssens, I. A., Nikinmaa, E., Ibrom, A., Wu, J., Bernhofer, C., Köstner, B., Grünwald, T., Seufert, G., Ciais, P., & Frank, D. (2014). Above-ground woody carbon sequestration measured from tree rings is coherent with net ecosystem productivity at five eddy-covariance sites. *New Phytologist*, *201*(4), 1289–1303. https://doi.org/10.1111/nph.12589

Bacour, C., Macbean, N., Chevallier, F., Léonard, S., Koffi, E. N., & Peylin, P. (2023). Assimilation of multiple datasets results in large differences in regional-to global-scale NEE and GPP budgets simulated by a terrestrial biosphere model. *Biogeosciences*, *20*(6), 1089–1111. https://doi.org/10.5194/bg-20-1089-2023

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., & Smets, B. (2013). GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sensing of Environment*, *137*, 299–309. https://doi.org/10.1016/j.rse.2012.12.027

Bastos, A., Ciais, P., Friedlingstein, P., Sitch, S., Pongratz, J., Fan, L., Wigneron, J. P., Weber, U., Reichstein, M., Fu, Z., Anthoni, P., Arneth, A., Haverd, V., Jain, A. K., Joetzjer, E., Knauer, J., Lienert, S., Loughran, T., McGuire, P. C., … Zaehle, S. (2020). Direct and seasonal legacy effects of the 2018 heat wave and drought on European ecosystem productivity. *Science Advances*, *6*(24). https://doi.org/10.1126/sciadv.aba2724

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, H. E., McVicar, T. R., Vergopolan, N., Berg, A., Lutsko, N. J., Dufour, A., Zeng, Z., Jiang, X., van Dijk, A. I. J. M., & Miralles, D. G. (2023). High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. *Scientific Data*, *10*(1). https://doi.org/10.1038/s41597-023-02549-6

Beguería, S., Serrano, S. M. V., Reig-Gracia, F., & Garcés, B. L. (2023). *SPEIbase v.2.9 [Dataset]*. https://doi.org/https://doi.org/10.20350/digitalCSIC/15470

Blackman, C. J., Pfautsch, S., Choat, B., Delzon, S., Gleason, S. M., & Duursma, R. A. (2016). Toward an index of desiccation time to tree mortality under drought. *Plant Cell and Environment*, *39*(10), 2342–2345. https://doi.org/10.1111/pce.12758

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., … Vuichard, N. (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

Bounoua, L., Collatz, G. J., Sellers, P. J., Randall, D. A., Dazlich, D. A., Los, S O, Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., & Jensen, T. G. (1999). Interactions between Vegetation and Climate: Radiative and Physiological Effects of Doubled Atmospheric CO_2.

Bounoua, L., Hall, F. G., Sellers, P. J., Kumar, A., Collatz, G. J., Tucker, C. J., & Imhoff, M. L. (2010). Quantifying the negative feedback of vegetation to greenhouse warming: A modeling approach. *Geophysical Research Letters*, *37*(23). https://doi.org/10.1029/2010GL045338

Bowen, I. S. (1926). The Ratio of Heat Losses by Conduction and by Evaporation from any Water Surface. *Physical Review*, *27*(6), 779–787. https://doi.org/10.1103/PhysRev.27.779

Breshears, D. D., Adams, H. D., Eamus, D., Mcdowell, N. G., Law, D. J., Will, R. E., Williams, A. P., & Zou, C. B. (2013). The critical amplifying role of increasing atmospheric moisture demand on tree mortality and associated regional die-off. In *Frontiers in Plant Science* (Vol. 4, Issue AUG). Frontiers Research Foundation. https://doi.org/10.3389/fpls.2013.00266

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*, *137*, 310–329. https://doi.org/10.1016/j.rse.2013.02.030

Cao, J., An, Q., Zhang, X., Xu, S., Si, T., & Niyogi, D. (2021). Is satellite Sun-Induced Chlorophyll Fluorescence more indicative than vegetation indices under drought condition? *Science of the Total Environment*, *792*. https://doi.org/10.1016/j.scitotenv.2021.148396

Chang, J. F., Viovy, N., Vuichard, N., Ciais, P., Wang, T., Cozic, A., Lardy, R., Graux, A. I., Klumpp, K., Martin, R., & Soussana, J. F. (2013). Incorporating grassland management in ORCHIDEE: Model description and evaluation at 11 eddy-covariance sites in Europe. *Geoscientific Model Development*, *6*(6), 2165–2181. https://doi.org/10.5194/gmd-6-2165-2013

Chang, J., Viovy, N., Vuichard, N., Ciais, P., Campioli, M., Klumpp, K., Martin, R., Leip, A., & Soussana, J. F. (2015). Modeled changes in potential grassland productivity and in grass-fed ruminant livestock density in Europe over 1961-2010. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0127554

Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P. P., Osório, M. L., Carvalho, I., Faria, T., & Pinheiro, C. (2002). How plants cope with water stress in the field. Photosynthesis and growth. *Annals of Botany*, *89*(SPEC. ISS.), 907–916. https://doi.org/10.1093/aob/mcf105

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J. L., Ghattas, J., Grandpeix, J. Y., Idelkadi, A., Mellul, L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., … Zhao, Y. (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

Conradt, T., Engelhardt, H., Menz, C., Vicente-Serrano, S. M., Farizo, B. A., Peña-Angulo, D., Domínguez-Castro, F., Eklundh, L., Jin, H., Boincean, B., Murphy, C., & López-Moreno, J. I. (2023). Cross-sectoral impacts of the 2018–2019 Central European drought and climate resilience in the German part of the Elbe River basin. *Regional Environmental Change*, *23*(1). https://doi.org/10.1007/s10113-023-02032-3

Cowan, I. R., & Farquhar, G. D. (1977). Stomatal function in relation to leaf metabolism and environment. *Symposia of the Society for Experimental Biology*, 31, 471–505

de La Motte, L. G., Beauclaire, Q., Heinesch, B., Cuntz, M., Foltýnová, L., Šigut, L., Kowalska, N., Manca, G., Ballarin, I. G., Vincke, C., Roland, M., Ibrom, A., Lousteau, D., Siebicke, L., Neiryink, J., & Longdoz, B. (2020). Non-stomatal processes reduce gross primary productivity in temperate forest ecosystems during severe edaphic drought: Edaphic drought in forest ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1810). https://doi.org/10.1098/rstb.2019.0527

de Pue, J., Barrios, J. M., Liu, L., Ciais, P., Arboleda, A., Hamdi, R., Balzarolo, M., Maignan, F., & Gellens-Meulenberghs, F. (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

de Rosnay, P., Polcher, J., Bruen, M., & Laval, K. (2002). Impact of a physically based soil water flow and soil-plant interaction representation for modeling large-scale land surface processes. *Journal of Geophysical Research: Atmospheres*, *107*(11). https://doi.org/10.1029/2001jd000634

Dewar, R., Mauranen, A., Mäkelä, A., Hölttä, T., Medlyn, B., & Vesala, T. (2018). New insights into the covariation of stomatal, mesophyll and hydraulic conductances from optimization models incorporating non-stomatal limitations to photosynthesis. *New Phytologist*, *217*(2), 571–585. https://doi.org/10.1111/nph.14848

D'orgeval, T., Polcher, J., & de Rosnay, P. (2008). Hydrology and Earth System Sciences Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes. In *Earth Syst. Sci. Discuss* (Vol. 12). www.hydrol-earth-syst-sci.net/12/1387/2008/

Dreyer, E., Le Roux, X., Montpied, P., Alain Daudet, F., & Masson, F. (2001). Temperature response of leaf photosynthetic capacity in seedlings from seven temperate tree species. *Tree Physiology, Volume 21*, Issue 4, March 2001, Pages 223–232, https://doi.org/10.1093/treephys/21.4.223

Du, Q., Jiao, X., Song, X., Zhang, J., Bai, P., Ding, J., & Li, J. (2020). The Response of Water Dynamics to Long-Term High Vapor Pressure Deficit Is Mediated by Anatomical Adaptations in Plants. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.00758

Dusenge, M. E., Duarte, A. G., & Way, D. A. (2019). Plant carbon metabolism and climate change: elevated CO2 and temperature impacts on photosynthesis, photorespiration and respiration. In *New Phytologist* (Vol. 221, Issue 1, pp. 32–49). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.15283

Duveiller, G., Guanter, L., Nelson, J., Hamdi, Z., Gensheimer, J., Dash, J., Ogutu, B., & Bandopadhyay, S. (2023). *Dataset for intercomparison of Sen4GPP GPP products at fluxtower sites*. Zenodo. https://doi.org/10.5281/zenodo.8339044

Ershadi, A., McCabe, M.F., Evans, J.P., & Wood, E.F. (2015). Impact of model structure and parameterization on Penman Monteith type evaporation models, *J. Hydrol.,* 525, 521–535

Fatichi, S., Leuzinger, S., & Körner, C. (2014). Moving beyond photosynthesis: From carbon source to sink-driven vegetation modeling. In *New Phytologist* (Vol. 201, Issue 4, pp. 1086–1095). https://doi.org/10.1111/nph.12614

Feng, X., Ackerly, D. D., Dawson, T. E., Manzoni, S., McLaughlin, B., Skelton, R. P., Vico, G., Weitz, A. P., & Thompson, S. E. (2019). Beyond isohydricity: The role of environmental variability in determining plant drought responses. *Plant Cell and Environment*, *42*(4), 1104–1111. https://doi.org/10.1111/pce.13486

Flexas, J., Barbour, M. M., Brendel, O., Cabrera, H. M., Carriquí, M., Díaz-Espejo, A., Douthe, C., Dreyer, E., Ferrio, J. P., Gago, J., Gallé, A., Galmés, J., Kodama, N., Medrano, H., Niinemets, Ü., Peguero-Pina, J. J., Pou, A., Ribas-Carbó, M., Tomás, M., … Warren, C. R. (2012). Mesophyll diffusion conductance to CO_2: An unappreciated central player in photosynthesis. In *Plant Science* (Vols. 193–194, pp. 70–84). https://doi.org/10.1016/j.plantsci.2012.05.009

Flexas, J., Bota, J., Loreto, F., Cornic, G., & Sharkey, T. D. (2004). Diffusive and metabolic limitations to photosynthesis under drought and salinity in C3 plants. In *Plant Biology* (Vol. 6, Issue 3, pp. 269–279). https://doi.org/10.1055/s-2004-820867

Flexas, J., Ribas-Carbó, M., Diaz-Espejo, A., Galmés, J., & Medrano, H. (2008). Mesophyll conductance to CO_2: Current knowledge and future prospects. In *Plant, Cell and Environment* (Vol. 31, Issue 5, pp. 602–621). https://doi.org/10.1111/j.1365-3040.2007.01757.x

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., … Zaehle, S. (2020). Global Carbon Budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. https://doi.org/10.5194/essd-12-3269-2020

Gago, J., Daloso, D. de M., Figueroa, C. M., Flexas, J., Fernie, A. R., & Nikoloski, Z. (2016). Relationships of leaf net photosynthesis, stomatal conductance, and mesophyll conductance to primary metabolism: A multispecies meta-analysis approach. *Plant Physiology*, *171*(1), 265–279. https://doi.org/10.1104/pp.15.01660

Gessler, A., Bottero, A., Marshall, J., & Arend, M. (2020). The way back: recovery of trees from drought and its implication for acclimation. In *New Phytologist* (Vol. 228, Issue 6, pp. 1704–1709). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.16703

Goldberg, D. E., & Holland, J. H. (1988). Genetic Algorithms and Machine Learning. *Machine Learning*, *3*(2), 95–99. https://doi.org/10.1023/A:1022602019183

Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T. W., Sperry, J. S., & McDowell, N. G. (2020). Plant responses to rising vapor pressure deficit. In *New Phytologist* (Vol. 226, Issue 6, pp. 1550–1566). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.16485

Grossiord, C., Sevanto, S., Borrego, I., Chan, A. M., Collins, A. D., Dickman, L. T., Hudson, P. J., McBranch, N., Michaletz, S. T., Pockman, W. T., Ryan, M., Vilagrosa, A., & McDowell, N. G. (2017). Tree water dynamics in a drying and warming world. *Plant Cell and Environment*, *40*(9), 1861–1873. https://doi.org/10.1111/pce.12991

Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T. A., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J., & Zhang, Y. (2021). The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission. *Earth System Science Data*, *13*(11), 5423–5440. https://doi.org/10.5194/essd-13-5423-2021

Guo, J. S., Hultine, K. R., Koch, G. W., Kropp, H., & Ogle, K. (2020). Temporal shifts in iso/anisohydry revealed from daily observations of plant water potential in a dominant desert shrub. *New Phytologist*, *225*(2), 713–726. https://doi.org/10.1111/nph.16196

Hacke, U. G., Sperry, J. S., Pockman, W. T., Davis, S. D., & McCulloh, K. A. (2001). Trends in wood density and structure are linked to prevention of xylem implosion by negative pressure. *Oecologia*, *126*(4), 457–461. https://doi.org/10.1007/s004420100628

Haupt, R. L., & Haupt, S. E. (2004). Practical genetic algorithms. *New Jersey: John Wiley & Sons.*

Herbst, M., Rosier, P., Morecroft, M., & Gowing, D. (2008).Comparative measurements of transpiration and canopy conductance in two mixed deciduous woodlands differing in structure

and species composition. *Tree Physiology, Volume 28*, Issue 6, June 2008, Pages 959–970, https://doi.org/10.1093/treephys/28.6.959

Hochberg, U., Rockwell, F. E., Holbrook, N. M., & Cochard, H. (2018). Iso/Anisohydry: A Plant-Environment Interaction Rather Than a Simple Hydraulic Trait. *Trends in Plant Science, 23*(2), 112–120. https://doi.org/10.1016/j.tplants.2017.11.002

Hochberg, U., Windt, C. W., Ponomarenko, A., Zhang, Y. J., Gersony, J., Rockwell, F. E., & Holbrook, N. M. (2017). Stomatal closure, basal leaf embolism, and shedding protect the hydraulic integrity of grape stems. *Plant Physiology*, *174*(2), 764–775. https://doi.org/10.1104/pp.16.01816

Ishida, A., Uemura, A., Koike, N., Matsumoto, Y., & Hoe, A. L. (1999). Interactive effects of leaf age and self-shading on leaf structure, photosynthetic capacity and chlorophyll fluorescence in the rain forest tree, Dryobalanops aromatica. *Tree Physiology*, *19*(11), 741–747. https://doi.org/10.1093/treephys/19.11.741

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

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

Keenan, T., Sabate, S., & Gracia, C. (2010). Soil water stress and coupled photosynthesis-conductance models: Bridging the gap between conflicting reports on the relative roles of stomatal, mesophyll conductance and biochemical limitations to photosynthesis. *Agricultural and Forest Meteorology*, *150*(3), 443–453. https://doi.org/10.1016/j.agrformet.2010.01.008

Kozlowski, T. T., & Pallardy, S. G. (2002). Acclimation and Adaptive Responses of Woody Plants to Environmental Stresses. In *The Botanical Review* (Vol. 68, Issue 2).

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. In *Global Biogeochemical Cycles* (Vol. 19, Issue 1, pp. 1–33). https://doi.org/10.1029/2003GB002199

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., & Richardson, A. D. (2012). Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, *9*(10), 3757–3776. https://doi.org/10.5194/bg-9-3757-2012

Kutsch, W. L., Herbst, M., Vanselow, R., Hummelshøj, P., Jensen, N. O., & Kappen, L. (2001). Stomatal acclimation influences water and carbon fluxes of a beech canopy in northern Germany Basic and Applied Ecology. In *Basic Appl. Ecol* (Vol. 2). http://www.urbanfischer.de/journals/baecol

Lardy, R., Bellocchi, G., & Soussana, J. F. (2011). A new method to determine soil organic carbon equilibrium. *Environmental Modelling and 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., Stoy, P., & Wohlfahrt, G. (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

Li, Y., Guan, K., Gentine, P., Konings, A. G., Meinzer, F. C., Kimball, J. S., Xu, X., Anderegg, W. R. L., McDowell, N. G., Martinez-Vilalta, J., Long, D. G., & Good, S. P. (2017). Estimating Global Ecosystem Isohydry/Anisohydry Using Active and Passive Microwave Satellite Data. *Journal of Geophysical Research: Biogeosciences*, *122*(12), 3306–3321. https://doi.org/10.1002/2017JG003958

Lin, W., Yuan, H., Dong, W., Zhang, S., Liu, S., Wei, N., Lu, X., Wei, Z., Hu, Y., & Dai, Y. (2023). Reprocessed MODIS Version 6.1 Leaf Area Index Dataset and Its Evaluation for Land Surface and Climate Modeling. *Remote Sensing*, *15*(7). https://doi.org/10.3390/rs15071780

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Contoux, C., Cozic, A., Cugnet, D., Dufresne, J. L., Éthé, C., Foujols, M. A., Ghattas, J., Hauglustaine, D., Hu, R. M., Kageyama, M., Khodri, M., … Boucher, O. (2020). Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model. *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2019MS001940

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., Maignan, F., Ottlé, C., Peaucelle, M., Santaren, D., & Peylin, P. (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. In *Global Biogeochemical Cycles* (Vol. 36, Issue 7). John Wiley and Sons Inc. https://doi.org/10.1029/2021GB007177

Mahan, J. R., & Upchurch, D. R. (1988). Maintenance of constant leaf temperature by plants-i. hypothesis-limited homeothermy. *In Environmental and Experimental Botany (Vol. 28, Issue 4)*

Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., Krishnan, P., Bastrikov, V., & MacBean, N. (2021). Optimizing Carbon Cycle Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Mean Net CO_2 Flux and its Inter-Annual Variability. *Journal of Geophysical Research: Biogeosciences*, *126*(10). https://doi.org/10.1029/2021JG006400

Marchin, R. M., Broadhead, A. A., Bostic, L. E., Dunn, R. R., & Hoffmann, W. A. (2016). Stomatal acclimation to vapour pressure deficit doubles transpiration of small tree seedlings with warming. *Plant Cell and Environment*, *39*(10), 2221–2234. https://doi.org/10.1111/pce.12790

Martínez-Vilalta, J., Poyatos, R., Aguadé, D., Retana, J., & Mencuccini, M. (2014). A new look at water transport regulation in plants. *New Phytologist*, *204*(1), 105–115. https://doi.org/10.1111/nph.12912

Massman, W. J. (1999). A model study of kB_H^{-1} for vegetated surfaces using "localized near-field" Lagrangian theory. *J. Hydrol. 223, 27–43.*

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., & Yepez, E. A. (2008). Mechanisms of plant survival and mortality during drought: Why do some plants survive while others succumb to drought? In *New Phytologist* (Vol. 178, Issue 4, pp. 719–739). https://doi.org/10.1111/j.1469-8137.2008.02436.x

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., Crous, K. Y., de Angelis, P., Freeman, M., & Wingate, L. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Meridja, S. (2011). Approche biophysique des processus de développement et de croissance des couverts végétaux : Interaction avec le stress hydrique et optimisation des pratiques culturales en climat méditerranéen. *Sciences agricoles. AgroParisTech*

Monteith, J. L. (1995). Accommodation between transpiring vegetation and the convective boundary layer. *In Journal of Hydrology ELSEVIER (Vol. 3)*

Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. *In Technometrics (Vol. 33, Issue 2)*

Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Mahmoud Hamdi, Z., … Jung, M. (2024). *X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X*. https://doi.org/10.5194/egusphere-2024-165

Nobel, P. S. (1974). An Introduction to Biophysical Plant Physiology. *W. H. Freeman and Company, San Francisco*

Novick, K. A., Ficklin, D. L., Grossiord, C., Konings, A. G., Martínez-Vilalta, J., Sadok, W., Trugman, A. T., Williams, A. P., Wright, A. J., Abatzoglou, J. T., Dannenberg, M. P., Gentine, P., Guan, K., Johnston, M. R., Lowman, L. E. L., Moore, D. J. P., & McDowell, N. G. (2024). The impacts of rising vapour pressure deficit in natural and managed ecosystems. In *Plant Cell and Environment*. John Wiley and Sons Inc. https://doi.org/10.1111/pce.14846

Novick, K. A., Konings, A. G., & Gentine, P. (2019). Beyond soil water potential: An expanded view on isohydricity including land–atmosphere interactions and phenology. *Plant Cell and Environment*, *42*(6), 1802–1815. https://doi.org/10.1111/pce.13517

Tardieu, F., & Simonneau, T. (1998). Variability among species of stomatal control under fluctuating soil water status and evaporative demand: modelling isohydric and anisohydric behaviours. *In Journal of Experimental Botany (Vol. 49)*

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Oren, R., Sperry, J. S., Katul, G. G., Pataki, D. E., Ewers, B. E., Phillips, N., & Schäfer, K. V. R. (1999). Survey and synthesis of intra- and interspecific variation in stomatal sensitivity to vapour pressure deficit. *Plant, Cell and Environment*, *22*(12), 1515–1526. https://doi.org/10.1046/j.1365-3040.1999.00513.x

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Ribeca, A., van Ingen, C., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., … Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1). https://doi.org/10.1038/s41597-020-0534-3

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

Pittermann, J., Sperry, J. S., Wheeler, J. K., Hacke, U. G., & Sikkema, E. H. (2006). Mechanical reinforcement of tracheids compromises the hydraulic efficiency of conifer xylem. *Plant, Cell and Environment*, *29*(8), 1618–1628. https://doi.org/10.1111/j.1365-3040.2006.01539.x

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., & Wright, I. J. (2014). Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology. *Ecology Letters*, *17*(1), 82–91. https://doi.org/10.1111/ele.12211

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S., & Voss, F. (2011). How well do large-scale models reproduce regional hydrological extremes: In Europe? *Journal of Hydrometeorology*, *12*(6), 1181–1204. https://doi.org/10.1175/2011JHM1387.1

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

Rezaei Nejad, A., & van Meeteren, U. (2008). Dynamics of adaptation of stomatal behaviour to moderate or high relative air humidity in Tradescantia virginiana. *Journal of Experimental Botany*, *59*(2), 289–301. https://doi.org/10.1093/jxb/erm308

Salmon, Y., Torres-Ruiz, J. M., Poyatos, R., Martinez-Vilalta, J., Meir, P., Cochard, H., & Mencuccini, M. (2015). Balancing the risks of hydraulic failure and carbon starvation: A twig scale analysis in declining Scots pine. *Plant Cell and Environment*, *38*(12), 2575–2588. https://doi.org/10.1111/pce.12572

Santaren, D., Peylin, P., Bacour, C., Ciais, P., & Longdoz, B. (2014). Ecosystem model optimization using in situ flux observations: Benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances. *Biogeosciences*, *11*(24), 7137–7158. https://doi.org/10.5194/bg-11-7137-2014

Schmied, G., Hilmers, T., Mellert, K. H., Uhl, E., Buness, V., Ambs, D., Steckel, M., Biber, P., Šeho, M., Hoffmann, Y. D., & Pretzsch, H. (2023). Nutrient regime modulates drought response patterns of three temperate tree species. *Science of the Total Environment*, *868*. https://doi.org/10.1016/j.scitotenv.2023.161601

Schuldt, B., Buras, A., Arend, M., Vitasse, Y., Beierkuhnlein, C., Damm, A., Gharun, M., Grams, T. E. E., Hauck, M., Hajek, P., Hartmann, H., Hiltbrunner, E., Hoch, G., Holloway-Phillips, M., Körner, C., Larysch, E., Lübbe, T., Nelson, D. B., Rammig, A., … Kahmen, A. (2020). A first assessment of the impact of the extreme 2018 summer drought on Central European forests. *Basic and Applied Ecology*, *45*, 86–103. https://doi.org/10.1016/j.baae.2020.04.003

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., … Zerroukat, M. (2019). UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513–4558. https://doi.org/10.1029/2019MS001739

Sellers, P. J., Bounoua, L., Collatz, G. J., Randall, D. A., Dazlich, D. A., Los, S. O., Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., & Jensen, T. G. (1996). Comparison of Radiative and Physiological Effects of Doubled Atmospheric CO_2 on Climate. *Science*, *271*(5254), 1402–1406. https://doi.org/10.1126/science.271.5254.1402

Sellers, P. J., Heiser, M. D., & Hall, F. G. (1992). Relations Between Surface Conductance and Spectral Vegetation Indices at Intermediate (100 m^2 to 15 km^2) Length Scales. *In Journal Of Geophysical Research (Vol. 97, Issue D17)*

Sermons, S. M., Seversike, T. M., Sinclair, T. R., Fiscus, E. L., & Rufty, T. W. (2012). Temperature influences the ability of tall fescue to control transpiration in response to atmospheric vapour pressure deficit. *Functional Plant Biology*, *39*(12), 979–986. https://doi.org/10.1071/FP12172

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., … Myneni, R. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. https://doi.org/10.5194/bg-12-653-2015

Song, L., Guanter, L., Guan, K., You, L., Huete, A., Ju, W., & Zhang, Y. (2018). Satellite sun-induced chlorophyll fluorescence detects early response of winter wheat to heat stress in the Indian Indo-Gangetic Plains. *Global Change Biology*, *24*(9), 4023–4037. https://doi.org/10.1111/gcb.14302

Stovall, A. E. L., Shugart, H., & Yang, X. (2019). Tree height explains mortality risk during an intense drought. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12380-6

Su, Z., Schmugge, T., Kustas, W. P., & Massman, W. J. (2001). An Evaluation of Two Models for Estimation of the Roughness Height for Heat Transfer between the Land Surface and the Atmosphere. *J. Appl. Meteor. Climatol.*, 40, 1933–1951, https://doi.org/10.1175/1520-0450(2001)040<1933:AEOTMF>2.0.CO;2

Sugiura, D., Terashima, I., & Evans, J. R. (2020). A decrease in mesophyll conductance by Cell-Wall thickening contributes to photosynthetic downregulation. *Plant Physiology*, *183*(4), 1600–1611. https://doi.org/10.1104/pp.20.00328

Sun, Y., Fu, R., Dickinson, R., Joiner, J., Frankenberg, C., Gu, L., Xia, Y., & Fernando, N. (2015). Drought onset mechanisms revealed by satellite solar-induced chlorophyll fluorescence: Insights from two contrasting extreme events. *Journal of Geophysical Research: Biogeosciences*, *120*(11), 2427–2440. https://doi.org/10.1002/2015JG003150

Tallaksen, L. M., & Stahl, K. (2014). Spatial and temporal patterns of large-scale droughts in Europe: Model dispersion and performance. *Geophysical Research Letters*, *41*(2), 429–434. https://doi.org/10.1002/2013GL058573

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

Tei, S., Sugimoto, A., Kotani, A., Ohta, T., Morozumi, T., Saito, S., Hashiguchi, S., & Maximov, T. (2019). Strong and stable relationships between tree-ring parameters and forest-level carbon fluxes in a Siberian larch forest. *Polar Science*, *21*, 146–157. https://doi.org/10.1016/j.polar.2019.02.001

Tramontana, G., Migliavacca, M., Jung, M., Reichstein, M., Keenan, T. F., Camps-Valls, G., Ogee, J., Verrelst, J., & Papale, D. (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

Urban, J., Ingwers, M. W., McGuire, M. A., & Teskey, R. O. (2017). Increase in leaf temperature opens stomata and decouples net photosynthesis from stomatal conductance in Pinus taeda and Populus deltoides x nigra. *Journal of Experimental Botany*, *68*(7), 1757–1767. https://doi.org/10.1093/jxb/erx052

Van Rossum, G., & Drake, F. L. (2009). Python 3 reference manual. *CreateSpace*

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Volk, J., Huntington, J., Allen, R., Melton, F., Anderson, M., & Kilic, A. (2021). flux-data-qaqc: A Python Package for Energy Balance Closure and Post-Processing of Eddy Flux Data. *Journal of Open Source Software*, *6*(66), 3418. https://doi.org/10.21105/joss.03418

Wang, Q., Zeng, J., Qi, J., Zhang, X., Zeng, Y., Shui, W., Xu, Z., Zhang, R., Wu, X., & Cong, J. (2021). A multi-scale daily SPEI dataset for drought characterization at observation stations over mainland China from 1961 to 2018. *Earth System Science Data*, *13*(2), 331–341. https://doi.org/10.5194/essd-13-331-2021

Wang, K.-Y., & Kellomäki, S. (1997). Stomatal conductance and transpiration in shoots of Scots pine after 4-year exposure to elevated CO_2 and temperature. *Canadian Journal of Botany*, *75*(4), 552–561. https://doi.org/10.1139/b97-061

Warm Winter 2020 Team, & ICOS Ecosystem Thematic Centre. (2022). Warm Winter 2020 ecosystem eddy covariance flux product for 73 stations in FLUXNET-Archive format—release 2022-1 (Version 1.0). *ICOS Carbon Portal*. https://doi.org/10.18160/2G60-ZHAK

Way, D. A., Oren, R., & Kroner, Y. (2015). The space-time continuum: The effects of elevated CO_2 and temperature on trees and the importance of scaling. In *Plant Cell and Environment* (Vol. 38, Issue 6, pp. 991–1007). Blackwell Publishing Ltd. https://doi.org/10.1111/pce.12527

Williams, C. A., Gu, H., MacLean, R., Masek, J. G., & Collatz, G. J. (2016). Disturbance and the carbon balance of US forests: A quantitative review of impacts from harvests, fires, insects, and droughts. In *Global and Planetary Change* (Vol. 143, pp. 66–80). Elsevier B.V. https://doi.org/10.1016/j.gloplacha.2016.06.002

Wu, X., Vuichard, N., Ciais, P., Viovy, N., de Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., di Tommasi, P., Moors, E. J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., … Ripoche, D. (2016). ORCHIDEE-CROP (v0), a new process-based agro-land surface model: Model description and evaluation over Europe. *Geoscientific Model Development*, *9*(2), 857–873. https://doi.org/10.5194/gmd-9-857-2016

Yang, J., Duursma, R. A., de Kauwe, M. G., Kumarathunge, D., Jiang, M., Mahmud, K., Gimeno, T. E., Crous, K. Y., Ellsworth, D. S., Peters, J., Choat, B., Eamus, D., & Medlyn, B. E. (2019). Incorporating non-stomatal limitation improves the performance of leaf and canopy models at high vapour pressure deficit. *Tree Physiology*, *39*(12), 1961–1974. https://doi.org/10.1093/treephys/tpz103

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

Yuan, H., Dai, Y., Xiao, Z., Ji, D., & Shangguan, W. (2011). Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. *Remote Sensing of Environment*, *115*(5), 1171–1187. https://doi.org/10.1016/j.rse.2011.01.001

Zhang, Y., Bastos, A., Maignan, F., Goll, D., Boucher, O., Li, L., Cescatti, A., Vuichard, N., Chen, X., Ammann, C., Altaf Arain, M., Andrew Black, T., Chojnicki, B., Kato, T., Mammarella, I., Montagnani, L., Roupsard, O., Sanz, M. J., Siebicke, L., … Ciais, P. (2020). Modeling the impacts of diffuse light fraction on photosynthesis in ORCHIDEE (v5453) land surface model. *Geoscientific Model Development*, *13*(11), 5401–5423. https://doi.org/10.5194/gmd-13-5401-2020

Supplementary

Text S1: Description of the main processes related to the vegetation response to stress implemented in ORCHIDEE.

VPD effect on stomatal conductance:

ORCHIDEE simulates stomatal conductance g_s (m s^{-1}) following Yin & Struik (2009), with an empirical function describing the effect of VPD f_{VPD} (unitless):

$$g_s = g_0 + \frac{A+R_d}{C_i - C_{i*}} \cdot f_{VPD} \quad \text{(equation S1)}$$

with g_0 (m s^{-1}) the residual stomatal conductance when irradiance approaches zero, A the CO_2 assimilation (µmol m^{-2} s^{-1}) corresponding to the minimum between the electron transport limited rate and the Rubisco-limited rate, and R_d the day respiration (µmol m^{-2} s^{-1}). C_i represents the intercellular partial pressure (µmol m^{-2} s^{-1}) and C_{i*} is the base compensation point when R_d is zero (µmol m^{-2} s^{-1}). f_{VPD} is defined using two empirical coefficients $A_{VPD}$and B_{VPD} (unitless) as:

$$f_{VPD} = \frac{1}{1/(A_{VPD} - B_{VPD} \cdot VDP) - 1} \quad \text{(equation S2)}$$

Root water uptake:

ORCHIDEE simulates a dynamic water uptake by roots in which the root density profile is updated at each timestep depending on the soil water content available for transpiration in each soil layer. Therefore, the $root_{density,l}$ value that weights $f_{water\ stress,l}$ matches the soil water content that can be taken up by root ($\theta_l - \theta_{WP,l}$). Roots can be allocated from the second soil layer to the bottom layer at 2 m deep for trees, while the maximum root depth considered for grasslands and crops is 1 m.

Resistance to soil evaporation:

In ORCHIDEE, a resistance to soil evaporation enables limiting soil evaporation when soil moisture decreases in the top 4 soil layers (~2.15 cm deep), corresponding to litter layers. This resistance to soil evaporation r_{soil} (s m^{-1}) is defined as an empirical exponential function of litter moisture following Sellers et al. (1992):

$$r_{soil} = exp\left(r1_{soil} - r2_{soil} \cdot \frac{\theta_{litter}}{\theta_{SAT,litter}}\right) \quad \text{(equation S4)}$$

with θ_{litter} and $\theta_{SAT,litter}$ the liquid soil moisture content in the litter, and liquid soil moisture content at saturation in the litter (kg m^{-2}), respectively. Then, $r1_{soil}$ and $r2_{soil}$ are two coefficients (unitless) initially estimated in Sellers et al. (1992) for a 5 cm deep soil. Therefore, the values of these coefficients need to be adapted to the litter depth considered in ORCHIDEE.

Table S1: List of ICOS Warm Winter 2020 sites with a fraction of dominant PFT of at least 50% at 0.0125°. Sites only used for evaluation are indicated with *.

Site name, country (ID)	Coordinates (lat, lon)	Dominant PFT fraction at 0.0125° (%) (ESA-CCI)	Soil texture (FAO-USDA)
Temperate Needleleaf Evergreen (TempENF)			
Gludsted Plantage, Denmark (DK-Gds)*	56.0737°, 9.3341°	75.0	Loamy sand
Oberbärenburg, Germany (DE-Obe)	50.7867°, 13.7213°	75.0	Sandy loam
Lavarone, Italy (IT-Lav)	45.9562°, 11.2813°	71.9	Sandy loam
Renon, Italy (IT-Ren)	46.5869°, 11.4337°	50.0	Sandy loam
Norunda, Sweden (SE-Nor)	60.0865°, 17.4795°	75.0	Sandy loam
Hyltemossa, Sweden (SE-Htm)	56.0976°, 13.4190°	68.0	Sandy loam
Bily Kriz forest, Czech Republic (CZ-BK1)	49.5021°, 18.5369°	75.0	Loam
Grillenburg ,Germany (DE-Gri)	50.9500°, 13.5126°	63.3	Loam
Rajec, Czech Republic (CZ-RAJ)	49.4437°, 16.6965°	75.0	Loam
Hetzdorf, Germany (DE-Hzd)	50.9638°, 13.4898°	75.0	Loam
Wustebach, Germany (DE-RuW)	50.5049°, 6.3310°	75.0	Loam
Lettosuo, Finland (FI-Let)	60.6418°, 23.9595°	65.8	Loam
San Rossore 2, Italy (IT-SR2)	43.7320°, 10.2909°	75.0	Clay loam
Font-Blanche, France (FR-FBn)	43.2408°, 5.67865°	66.7	Clay loam
Temperate Broadleaf Summergreen (TempDBF)			
Fyodorovskoye dry spruce, Russia (RU-Fy2)	56.4476°, 32.9019°	60.3	Sandy loam
Lanzhot, Czech Republic (CZ-Lnz)	48.6815°, 16.9463°	54.4	Loam
Stitna, Czech Republic (CZ- Stn)	49.0360°, 17.9699°	61.2	Loam

Hohes Holz, Germany (DE-HoH)	52.0866°, 11.2224°	70.0	Loam
Laegern, Switzerland (CH-Lae)	47.4783°, 8.3644°	54.0	Loam
Soroe, Denmark (DK-Sor)	55.4859°, 11.6446°	50.8	Loam
Hainich, Germany (DE-Hai)	51.0792°, 10.4522°	70.0	Clay loam
Boreal Needleleaf Evergreen (BorENF)			
Kenttarova, Finland (FI-Ken)*	67.98721°, 24.2430°	80.0	Sandy loam
Varrio, Finland (FI-Var)	67.7549°, 29.61°	80.0	Sandy loam
Rosinedal-3, Sweden (SE-Ros)	64.1725°, 19.738°	79.8	Sandy loam
Svartberget, Sweden (SE-Svb)	64.2561°, 19.7745°	80.0	Sandy loam
Hyytiala, Finland (FI-Hyy)	61.8474°, 24.2948°	80.0	Loam
Alp Weissenstein, Switzerland (CH-Aws)	46.5832°, 9.79042°	53.3	Sandy clay loam
Davos, Switzerland (CH-Dav)	46.8153°, 9.85591°	56.6	Sandy clay loam
Temperate Natural Grassland (C3) (C_3TempGRA)			
Chamau, Switzerland (CH-Cha)	47.2102°, 8.4104°	60.6	Loam
Anklam, Germany (DE-Akm)	53.8662°, 13.6834°	65.6	Loam
Rollesbroich, Germany (DE-RuR)	50.6219°, 6.3041°	90.6	Loam
Clara Raised Bog, Ireland (IE-Cra)*	53.3231°, -7.6418°	69.4	Loam
Trebon, Czech Republic (CZ-wet)	49.0247°, 14.7704°	81.6	Sandy clay loam
Crops (C3) (C_3Crop)			
Oensingen crop, Switzerland (CH-Oe2)	47.2864°, 7.7338°	75.4	Loam
Klingenberg, Germany (DE-Kli)	50.8931°, 13.5224°	51.7	Loam
Selhausen Juelich, Germany (DE-RuS)	50.8659°, 6.4471°	79.3	Loam
Monte Bondone, Italy (IT-MBo)	46.0147°, 11.0458°	60.7	Loam

Dorinne, Belgium (BE-Dor)	50.3119°, 4.9681°	67.5	Loam
Majadas del Tietar North, Spain (ES-LM1)	39.9427°, -5.7787°	66.6	Loam
Majadas del Tietar South, Spain (ES-LM2)	39.9346°, -5.7759°	50.0	Loam
Lamasquere, France (FR-Lam)	43.4964°, 1.2379°	64.8	Loam
Lison, Italy (IT-Lsn)	45.7405°, 12.7503°	61.9	Loam
Albuera, Spain (ES-Abr)	38.7018°, -6.7859°	63.9	Sandy clay loam
Gebesee, Germany (DE-Geb)	51.0997°, 10.9146°	82.4	Clay loam
Qvidja, Finland (FI-Qvd)*	60.2952°, 22.3916°	64.8	Clay loam
Boreal Natural Grassland (C3) (C_3BorGRA)			
Torgnon, Italy (IT-Tor)	45.8444°, 7.5781°	51.2	Sandy clay loam

Table S2: SPEI classes used for drought severity characterization (Wang et al., 2021).

SPEI classes	Characterization
SPEI ≥ 2	Extremely wet
1.5 ≤ SPEI < 2.	Severely wet
1. ≤ SPEI < 1.5	Moderately wet
0.5 < SPEI < 1.	Mildly wet
-0.5 ≤ SPEI ≤ 0.5	Normal
-1. < SPEI < -0.5	Mild drought
-1.5 < SPEI ≤ -1.	Moderate drought
-2. < SPEI ≤ -1.5	Severe drought
SPEI ≤ -2.	Extreme drought

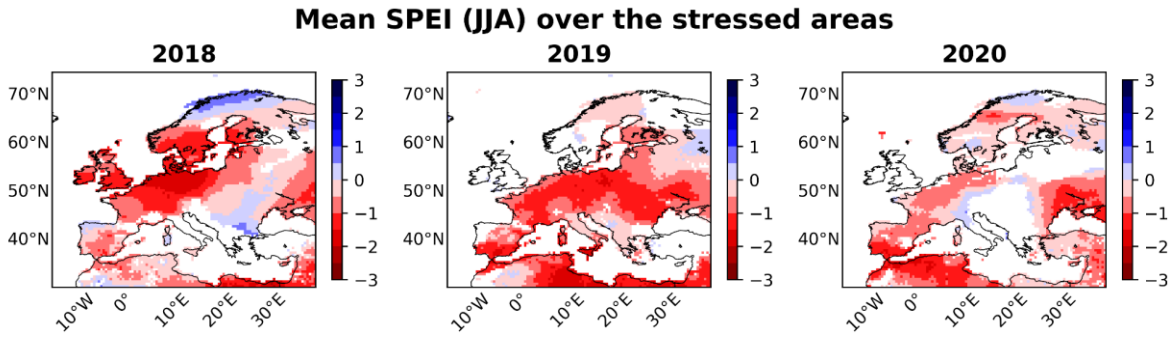

Figure S1: Maps of mean SPEI values (-) between June and August for the years 2018, 2019, and 2020, over the stressed areas. For each year, the stressed areas correspond to the grid cells for which monthly SPEI values reached a value below -1 (mild to extreme droughts) for at least one month in June, July or August.

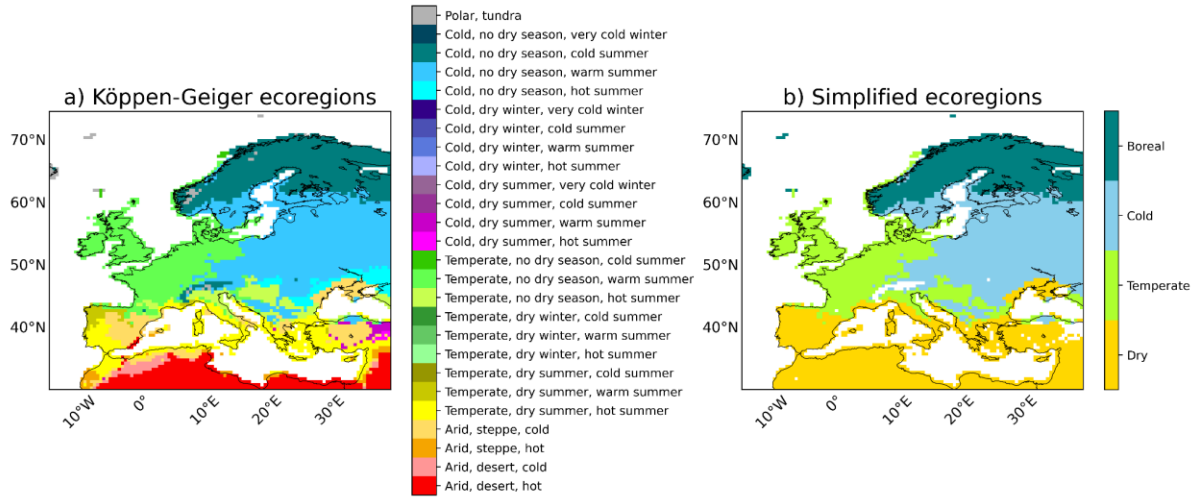

Figure S2: a) Map of European ecoregions according to the original Köppen-Geiger classification (period 1991-2020) at a 0.5° spatial resolution. b) Map of simplified European ecoregions at a 0.5° spatial resolution. In this simplified classification, the following biomes have been grouped together: arid with temperate dry summer biomes into "dry"; temperate no dry season biomes into "temperate"; cold no dry season hot summer with cold no dry season warm summer biomes into "cold"; cold no dry season cold summer with polar tundra biomes into "boreal". Note that the cold dry summer and cold dry winter biomes in a) have been removed from the "dry" ecoregion in b) because of the differences between these cold biomes and the arid and temperate ones included in the "dry" ecoregion. The grid cells of the cold no dry season cold winter biome for latitudes below 50°N in a) (mainly distributed over the Alps) have also been masked in the "boreal" ecoregion in b) because of the specific behavior expected from these high-altitude grid cells compared to the rest of the "boreal" ecoregion.

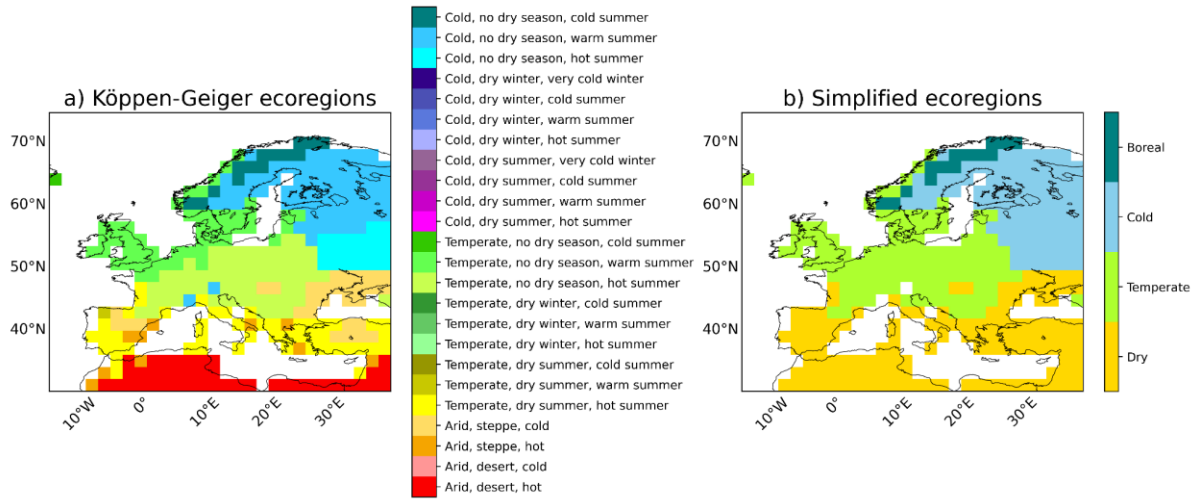

Figure S3: a) Map of future European ecoregions according to the original Köppen-Geiger classification under SSP370 (period 2070-2100) at a 2° spatial resolution. b) Map of simplified future European ecoregions under SSP370 (period 2070-2100) at a 2° spatial resolution. In this simplified classification, the biomes have been grouped as described in Figure S2.

Table S3: Parameters selected for the optimization.

Name	Description (unit)	Dependency	Prior value	Range of variation
Photosynthesis				
$Vcmax_{25}$	Rubisco maximum carboxylation rate at 25°C (µmol m^{-2} s^{-1})	TempENF	35.	[26.; 44.]
		TempDBF	50.	[38.; 63.]
		BorENF	45.	[34.; 56.]
		C_3TempGrass	50.	[38.; 63.]
		C_3Crop	60.	[45.; 75.]
		C_3BorGrass	50.	[38.; 63.]
$acclim_{Vcmax}$	Offset of the linear temperature acclimation relationship for the entropy parameter of the V_{cmax} temperature-dependence function, following Kattge & Knorr (2007) (J K^{-1} mol^{-1})	TempENF	668.39	[501.3; 835.5]
		TempDBF	668.39	[501.3; 835.5]
		BorENF	668.39	[501.3; 835.5]
		C_3TempGrass	668.39	[501.3; 835.5]
		C_3Crop	668.39	[501.3; 835.5]
		C_3BorGrass	668.39	[501.3; 835.5]
$acclim_{Jmax}$	Offset of the linear temperature acclimation relationship for the entropy parameter of the J_{max} temperature-dependence function, following Kattge & Knorr (2007) (J K^{-1} mol^{-1})	TempENF	659.70	[495.; 825.]
		TempDBF	659.70	[495.; 825.]
		BorENF	659.70	[495.; 825.]
		C_3TempGrass	659.70	[495.; 825.]
		C_3Crop	659.70	[495.; 825.]
		C_3BorGrass	659.70	[495.; 825.]
$Leaf_{age\ crit}$	Critical leaf age, used for computing the age-dependence of the maximum	TempENF	910.	[683.; 1138.]
		TempDBF	160.	[120.; 200.]

	carboxylation rate limited by Rubisco activity, and for leaf turnover (Krinner et al., 2005) (days)	BorENF	910.	[683.; 1138.]
		C_3TempGrass	80.	[60.; 100.]
		C_3Crop	90.	[68.; 113.]
		C_3BorGrass	80.	[60.; 100.]
Biomass				
SLA	Specific leaf area (m^2 gC^{-1})	TempENF	0.00926	[0.00695; 0.01157]
		TempDBF	0.026	[0.020; 0.033]
		BorENF	0.00926	[0.00695; 0.01157]
		C_3TempGrass	0.026	[0.020; 0.033]
		C_3Crop	0.026	[0.020; 0.033]
		C_3BorGrass	0.026	[0.020; 0.033]
LAI_{max}	Maximum leaf area index (m^2 m^{-2})	TempENF	5.	[3.75; 6.25]
		TempDBF	5.	[3.75; 6.25]
		BorENF	4.5	[3.4; 5.6]
		C_3TempGrass	2.5	[1.88; 3.13]
		C_3Crop	5.	[3.75; 6.25]
		C_3BorGrass	2.5	[1.88; 3.13]
Stomatal conductance response to VPD				
A_{VPD}	Empirical factor involved in the function describing the effect of VPD on stomatal conductance (-)	TempENF	0.85	[0.64, 1.06]
		TempDBF	0.85	[0.64, 1.06]
		BorENF	0.85	[0.64, 1.06]
		C_3TempGrass	0.85	[0.64, 1.06]
		C_3Crop	0.85	[0.64, 1.06]

		C_3BorGrass	0.85	[0.64, 1.06]
B_{VPD}	Empirical factor involved in the function describing the effect of VPD on stomatal conductance (-)	TempENF	0.14	[0.1, 0.18]
		TempDBF	0.14	[0.1, 0.18]
		BorENF	0.14	[0.1, 0.18]
		C_3TempGrass	0.14	[0.1, 0.18]
		C_3Crop	0.14	[0.1, 0.18]
		C_3BorGrass	0.14	[0.1, 0.18]
Vegetation sensitivity to soil moisture stress				
$\alpha_{water\ stress}$	Empirical factor determining the speed to stomatal closure under soil moisture stress (-)	TempENF	1.	[0.05, 10.]
		TempDBF	1.	[0.05, 10.]
		BorENF	1.	[0.05, 10.]
		C_3TempGrass	1.	[0.05, 10.]
		C_3Crop	1.	[0.05, 10.]
		C_3BorGrass	1.	[0.05, 10.]
p_{max}	Fraction of saturated volumetric soil moisture above which transpiration is maximum ([0,1]) (-)	Loamy sand	0.8	[0.6, 0.9999]
		Sandy loam	0.8	[0.6, 0.9999]
		Loam	0.8	[0.6, 0.9999]
		Sandy clay loam	0.8	[0.6, 0.9999]
		Clay loam	0.8	[0.6, 0.9999]
Soil hydrology				
$n_{van\ Genuchten}$	Root profile Van Genuchten coefficient n (-)	Loamy sand	2.68	[2.01; 3.35]
		Sandy loam	1.89	[1.42; 2.36]
		Loam	1.56	[1.17; 1.95]

		Sandy clay loam	1.48	[1.11; 1.85]
		Clay loam	1.31	[0.98; 1.64]
θ_{SAT}	Volumetric water content at saturation (m^3 m^{-3})	Loamy sand	0.41	[0.31; 0.51]
		Sandy loam	0.41	[0.31; 0.51]
		Loam	0.43	[0.32; 0.54]
		Sandy clay loam	0.39	[0.29; 0.49]
		Clay loam	0.41	[0.31; 0.51]
θ_{FC}	Volumetric water content at field capacity (m^3 m^{-3})	Loamy sand	0.0710	[0.071; 0.0888]
		Sandy loam	0.1218	[0.0914; 0.1523]
		Loam	0.1654	[0.1241; 0.2068]
		Sandy clay loam	0.1695	[0.1390; 0.2119]
		Clay loam	0.2697	[0.2028; 0.309]
θ_{WP}	Volumetric water content at wilting point (m^3 m^{-3})	Loamy sand	0.057	[0.057; 0.071]
		Sandy loam	0.0657	[0.065; 0.0821]
		Loam	0.0884	[0.078; 0.1105]
		Sandy clay loam	0.1112	[0.100; 0.1390]
		Clay loam	0.1496	[0.1122; 0.187]
Resistance to soil evaporation				
$r1_{soil}$	Coefficient 1 for soil resistance (Sellers et al., 1992) (-)	None	8.206	[4.103; 12.309]
$r2_{soil}$	Coefficient 2 for soil resistance (Sellers et al., 1992) (-)	None	4.255	[2.1275; 6.3825]
Aerodynamic conductance				

$r1_{aero}$	Constant 1 used in the formulation of the ratio of friction velocity to the wind speed at the canopy top (Ershadi et al., 2015) (-)	None	0.32	[0.24; 0.40]
$r2_{aero}$	Constant 2 used in the formulation of the ratio of friction velocity to the wind speed at the canopy top (Ershadi et al., 2015) (-)	None	0.264	[0.198; 0.330]
Radiative transfer and leaf absorption				
Clumping	Leaf clumping index (-)	TempENF	0.74	[0.555; 0.925]
		TempDBF	0.70	[0.10; 1.50]
		BorENF	0.55	[0.41; 0.69]
		C_3TempGrass	0.75	[0.10; 1.50]
		C_3Crop	0.75	[0.10; 1.50]
		C_3BorGrass	0.75	[0.10; 1.50]
a_{PSII}	Absorption cross section for photosystem II (-)	None	0.5	[0.375; 0.625]

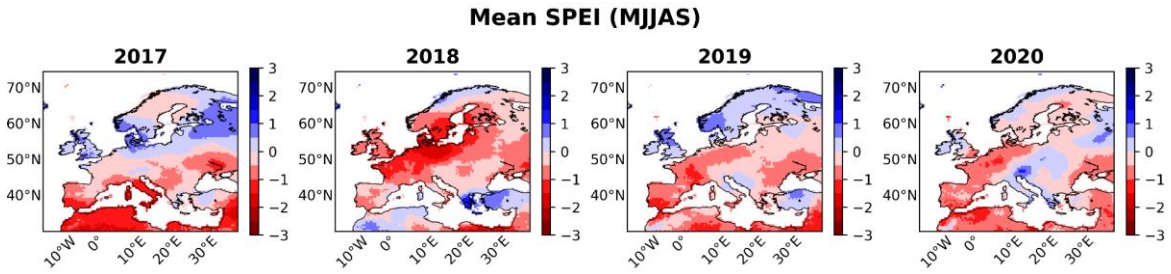

Figure S4: Maps of mean SPEI values (-) between May and September for each year from 2017 to 2020.

Table S4: Mean RMSD value for GPP (gC m^{-2} d^{-1}) and LE (W m^{-2}) per PFT following the first DA steps assimilating GPP and LE data either in 2017 or in 2018. This evaluation is based on RMSD values between the daily observed and simulated variable of the optimized year.

			TempENF	TempDBF	BorENF	C_3TempGrass	C_3Crop	C_3BorGrass	Average over all PFTs
GPP	2017	PRIOR	2.28	2.51	1.90	2.97	2.93	1.14	2.29
		POST	1.95	2.45	1.20	2.75	2.59	0.83	1.96
	2018	PRIOR	2.22	2.31	1.96	2.36	2.81	1.14	2.13
		POST	1.95	1.96	1.26	2.14	2.17	1.05	1.75
LE	2017	PRIOR	23.80	21.76	19.30	31.79	24.97	19.72	23.56
		POST	21.64	19.55	14.43	16.77	22.09	8.49	17.16
	2018	PRIOR	22.20	23.65	26.88	29.48	23.89	12.29	23.06
		POST	20.12	19.60	17.95	15.01	17.64	7.41	16.29

Table S5: Median RMSD value for GPP (gC m^{-2} d^{-1}) and LE (W m^{-2}) per PFT considering all Warm Winter 2020 sites for which $f_{water\ stress}$ reaches a value below 0.85 between May and September in 2018, 2019, or 2020, for "ORCHIDEE-PRIOR" and the two optimized versions: "POST ORCHIDEE-STANDARD" and "POST ORCHIDEE-ACCLIM". This evaluation is based on RMSD values between the daily observed and simulated variable.

		TempENF	TempDBF	BorENF	C_3TempGrass	C_3Crop	C_3BorGrass	Average over all PFTs
GPP	ORCHIDEE-PRIOR	3.03	3.34	3.89	3.87	1.32	0.57	2.67
	POST ORCHIDEE-STANDARD	2.64	4.05	2.10	4.76	1.69	0.98	2.70
	POST ORCHIDEE-ACCLIM	2.57	2.03	1.61	4.57	1.62	0.87	2.21

LE	ORCHIDEE-PRIOR	26.58	31.82	37.06	44.49	18.00	6.12	27.35
	POST ORCHIDEE-STANDARD	23.26	18.89	19.33	69.46	15.53	5.26	25.29
	POST ORCHIDEE-ACCLIM	20.31	19.91	17.03	72.86	14.65	4.10	24.81

Table S6: Median RMSD value for GPP (gC m^{-2} d^{-1}) and LE (W m^{-2}) per PFT considering all Warm Winter 2020 sites from 2018 to 2020, for "ORCHIDEE-PRIOR" and the two optimized versions: "POST ORCHIDEE-STANDARD" and "POST ORCHIDEE-ACCLIM". This evaluation is based on RMSD values between the daily observed and simulated variable.

		TempENF	TempDBF	BorENF	C_3TempGrass	C_3Crop	C_3BorGrass	Average over all PFTs
GPP	ORCHIDEE-PRIOR	2.37	2.31	2.48	3.43	2.92	1.39	2.48
	POST ORCHIDEE-STANDARD	2.41	2.16	1.47	3.34	2.67	1.40	2.24
	POST ORCHIDEE-ACCLIM	3.39	2.17	1.41	3.34	2.68	1.41	2.4
LE	ORCHIDEE-PRIOR	24.76	23.11	27.26	29.47	25.49	14.85	24.16
	POST ORCHIDEE-STANDARD	22.59	20.62	24.35	19.31	20.52	10.37	19.63
	POST ORCHIDEE-ACCLIM	22.73	20.09	25.87	19.30	24.31	10.56	20.48

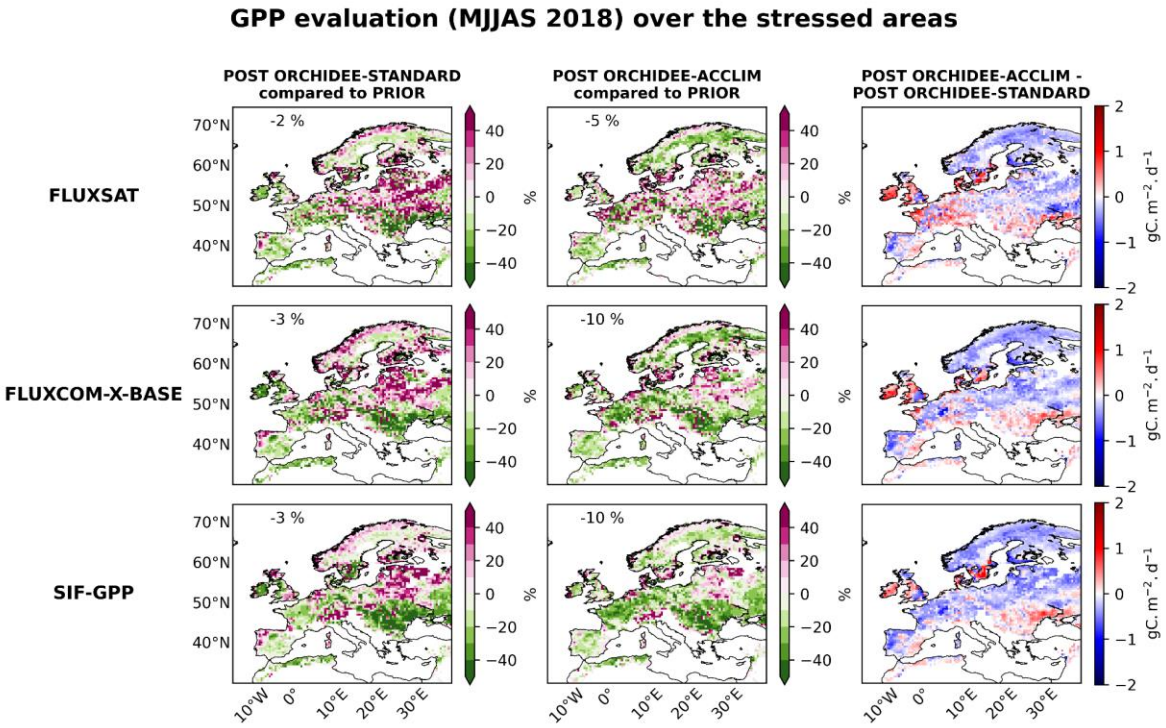

Figure S5: Maps of RMSD change (%) between the RMSD computed between the data-driven and post-optimization simulated GPP (for POST ORCHIDEE-STANDARD on the left column or POST ORCHIDEE-ACCLIM in the middle column), and the RMSD computed between the data-driven and ORCHIDEE-PRIOR GPP, for each evaluation product (FLUXSAT, FLUXCOM-X-BASE, SIF-GPP). The right column shows the RMSD difference (gC m^{-2} d^{-1}) between POST ORCHIDEE-ACCLIM and POST ORCHIDEE-STANDARD for each evaluation product. The selected grid cells correspond to the stressed areas (section 2.2.3), between May and September 2018. Data are 8-day averages with a 0.5° spatial resolution. For the first two columns, grid cells in green depict an improvement in the posterior simulated GPP compared to ORCHIDEE-PRIOR (reduction in RMSD), while grid cells in pink show a degradation of the simulated GPP after optimization. For these two columns, the median RMSD change (%) is provided in the top left corner.

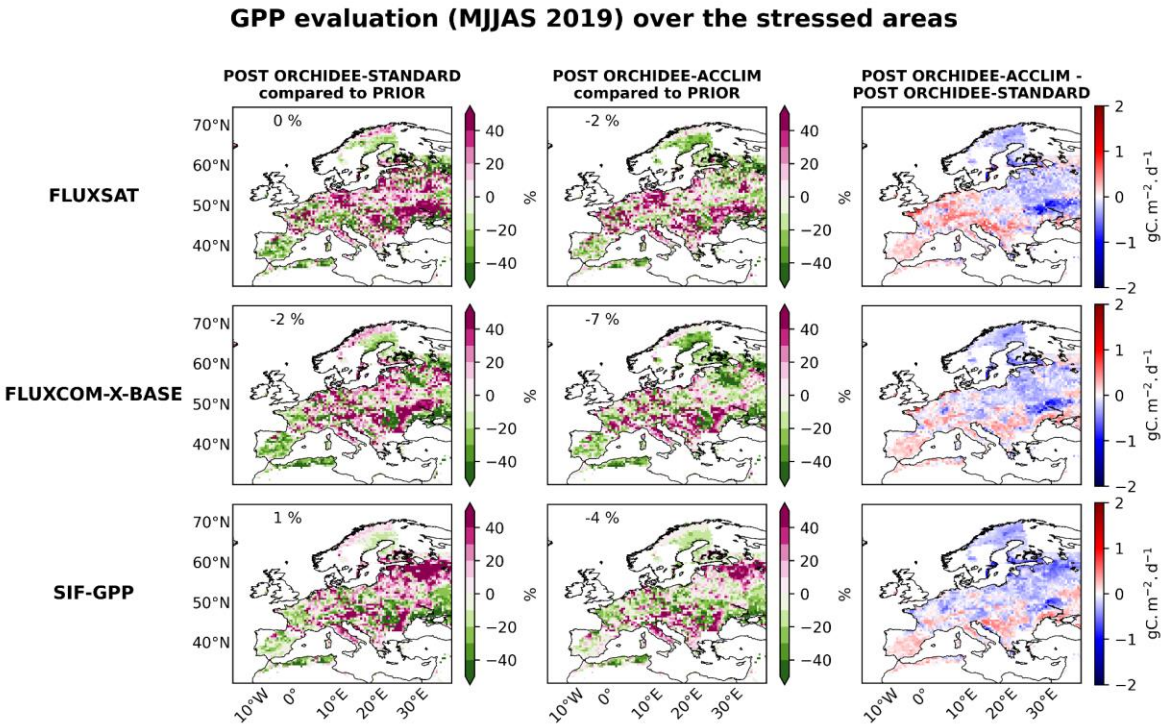

Figure S6: Same as Figure S5 for the year 2019.

Table S7: Percentage of improved grid cells after optimization for the two model versions (POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM) compared to the prior, evaluated using the RMSD for the three GPP evaluation products, over the stressed areas in 2018, 2019, and 2020. The highest improvement between the two optimized model versions are in bold.

	2018			2019			2020			Mean improvement
	FLUXSAT	FLUXCOM-X-BASE	SIF-GPP	FLUXSAT	FLUXCOM-X-BASE	SIF-GPP	FLUXSAT	FLUXCOM-X-BASE	SIF-GPP	
POST ORCHIDEE-STANDARD	53	55	54	49	53	48	55	48	50	52
POST ORCHIDEE-ACCLIM	**57**	**66**	**69**	**54**	**61**	**57**	**60**	**64**	**64**	**61**

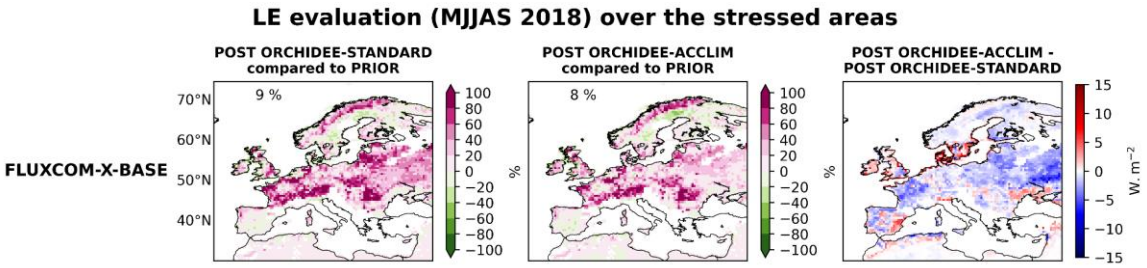

Figure S7: Maps of RMSD change (%) between the RMSD computed between FLUXCOM-X-BASE and post-optimization simulated LE (for POST ORCHIDEE-STANDARD on the left column or POST ORCHIDEE-ACCLIM in the middle column), and the RMSD computed between FLUXCOM-X-BASE and ORCHIDEE-PRIOR LE. The right column shows the RMSD difference (W m^{-2}) between POST ORCHIDEE-ACCLIM and POST ORCHIDEE-STANDARD for each evaluation product. The selected grid cells correspond to the stressed areas (section 2.2.3), between May and September 2020. Data are 8-day averages with a 0.5° spatial resolution. For the first two columns, grid cells in green depict an improvement in the posterior simulated LE compared to ORCHIDEE-PRIOR (reduction in RMSD), while grid cells in pink show a degradation of the simulated LE after optimization. For these two columns, the median RMSD change (%) is provided in the top left corner.

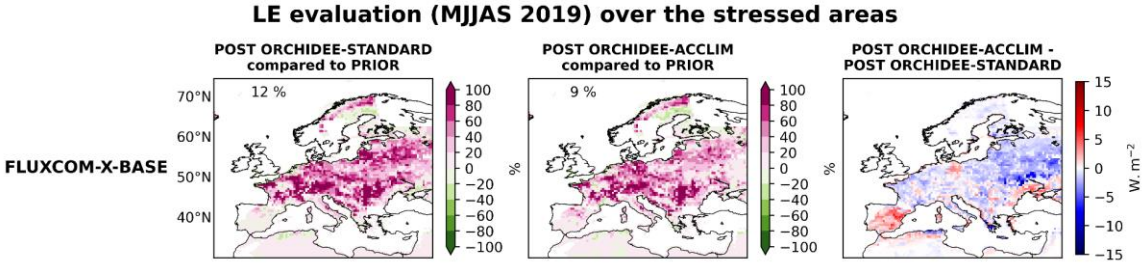

Figure S8: Same as Figure S7 for the year 2019.

Table S8: Percentage of improved grid cells after optimization for the two model versions (POST ORCHIDEE-STANDARD and POST ORCHIDEE-ACCLIM) compared to the prior, evaluated using the RMSD for FLUXCOM-X-BASE LE, over the stressed areas in 2018, 2019, and 2020. The highest improvement between the two optimized model versions are in bold.

	2018	2019	2020	Mean improvement
POST ORCHIDEE-STANDARD	**26**	19	25	23
POST ORCHIDEE-ACCLIM	**26**	**21**	**26**	**24**

Table S9: Mean LAI RMSD (m^2 m^{-2}) values between the 3 ORCHIDEE simulations and the 3 evaluation products (MODIS, PROBA-V, OLCI) for each ecoregion between 2018 and 2020. Data are monthly averages. For all evaluation products, grid cells where the evaluation products have missing data have been masked. The lowest RMSD values between the 3 ORCHIDEE simulations are in bold.

	Boreal			**Cold**			**Temperate**			**Dry**			**Europe**		
	ORCHIDEE-PRIOR	POST ORCHIDEE-STANDARD	POST ORCHIDEE-ACCLIM	ORCHIDEE-PRIOR	POST ORCHIDEE-STANDARD	POST ORCHIDEE-ACCLIM	ORCHIDEE-PRIOR	POST ORCHIDEE-STANDARD	POST ORCHIDEE-ACCLIM	ORCHIDEE-PRIOR	POST ORCHIDEE-STANDARD	POST ORCHIDEE-ACCLIM	ORCHIDEE-PRIOR	POST ORCHIDEE-STANDARD	POST ORCHIDEE-ACCLIM
PROBA-V	0.99	0.86	**0.74**	1.08	1.10	**1.01**	1.06	1.03	**0.98**	0.88	0.85	**0.76**	0.99	0.97	**0.89**
OLCI	1.13	0.99	**0.77**	1.08	1.25	**1.04**	1.02	1.04	**0.85**	0.9	0.87	**0.76**	1.01	1.04	**0.87**
MODIS	1.06	0.89	**0.68**	0.83	0.85	**0.75**	0.77	0.74	**0.66**	0.73	0.71	**0.60**	0.81	0.78	**0.67**

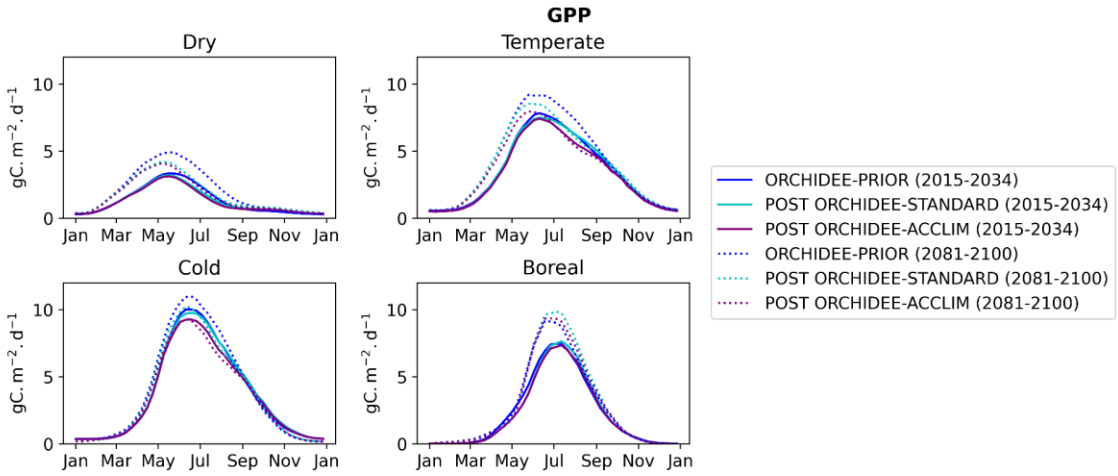

Figure S9: Mean seasonal cycle of simulated GPP ($gC.m^{-2}.d^{-1}$) over each ecoregion in Europe, projected under the SSP370 scenario for the period 2015-2034 and 2081-2100. Data are 8-day averages with a 2° spatial resolution.

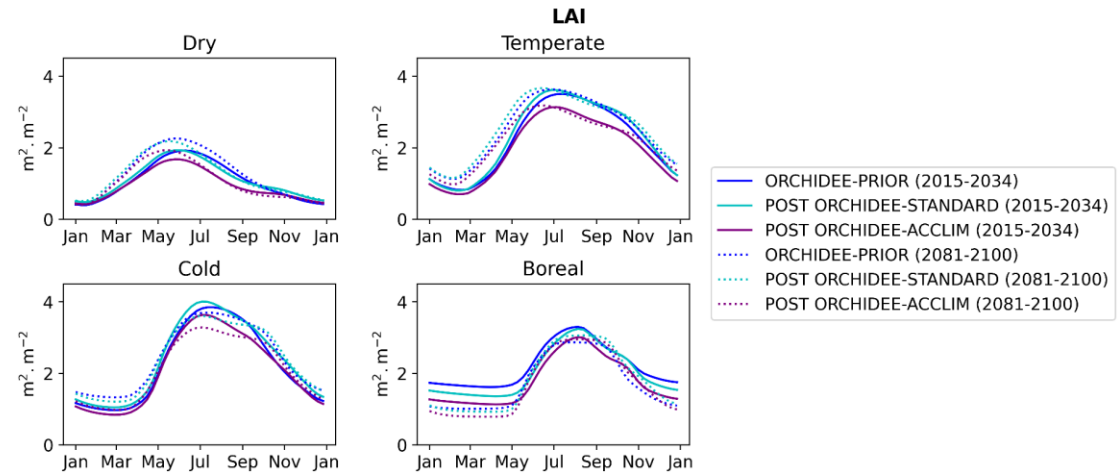

Figure S10: Mean seasonal cycle of simulated LAI ($m^2.m^{-2}$) over each ecoregion in Europe, projected under the SSP370 scenario for the period 2015-2034 and 2081-2100. Data are 8-day averages with a 2° spatial resolution.

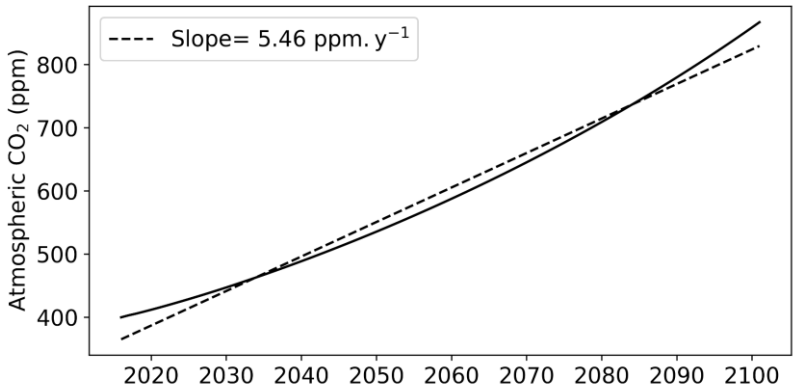

Figure S11: Projected variations of atmospheric CO_2 concentration (ppm) over Europe from 2015 to 2100 under the SSP370 scenario. The mean annual values are represented in solid lines and the corresponding linear regressions are in dotted lines with the regression coefficient values (slope).

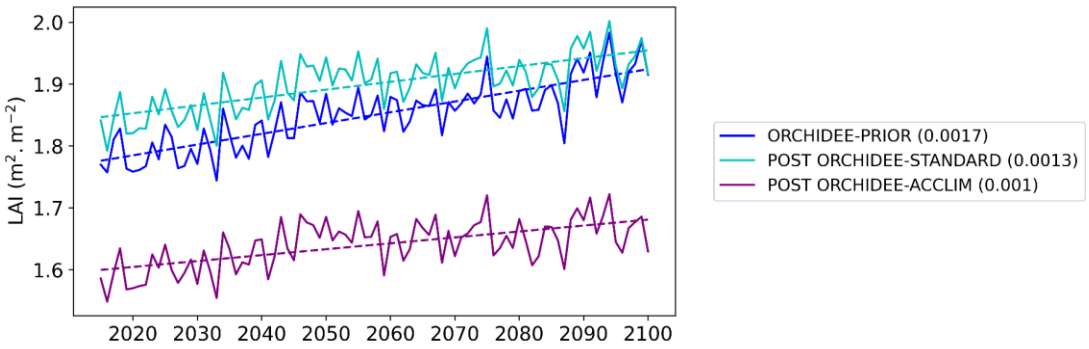

Figure S12: Projected variations of LAI ($m^2.m^{-2}$) over Europe from 2015 to 2100 under the SSP370 scenario. The mean annual values are represented in solid lines, and the corresponding linear regressions are in dotted lines with the regression coefficient values (slope) given in the legend for each model version.

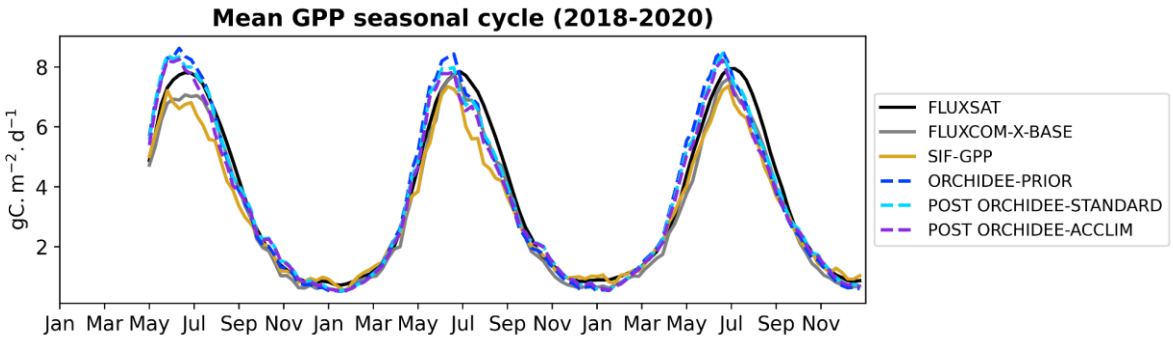

Figure S13: Mean GPP seasonal cycle (gC m^{-2} d^{-1}) for the evaluation products (FLUXSAT, FLUXCOM-X-BASE, SIF-GPP) (solid lines) and ORCHIDEE simulations (dotted lines) over the stressed areas in Europe from 2018 to 2020. Data are 8-day averages at a 0.5° spatial resolution. Grid cells where the GPP evaluation products have missing data have been masked.

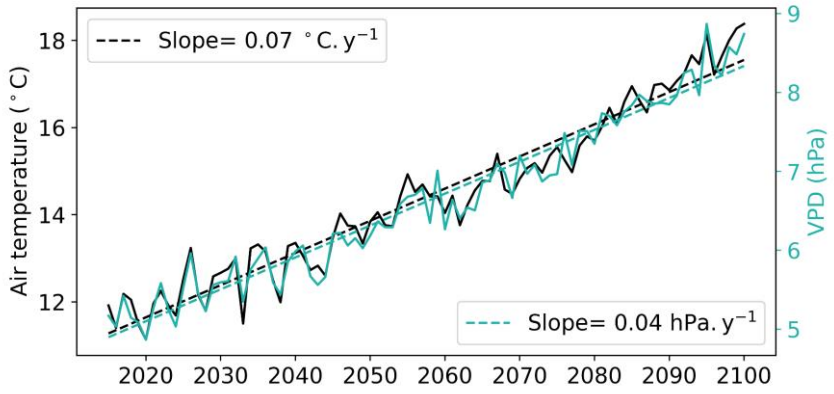

Figure S14: Projected variations of air temperature (°C) and VPD (hPA) over Europe from 2015 to 2100 under the SSP370 scenario. The mean annual values are represented in solid lines and the corresponding linear regressions are in dotted lines with the regression coefficient values (slope).

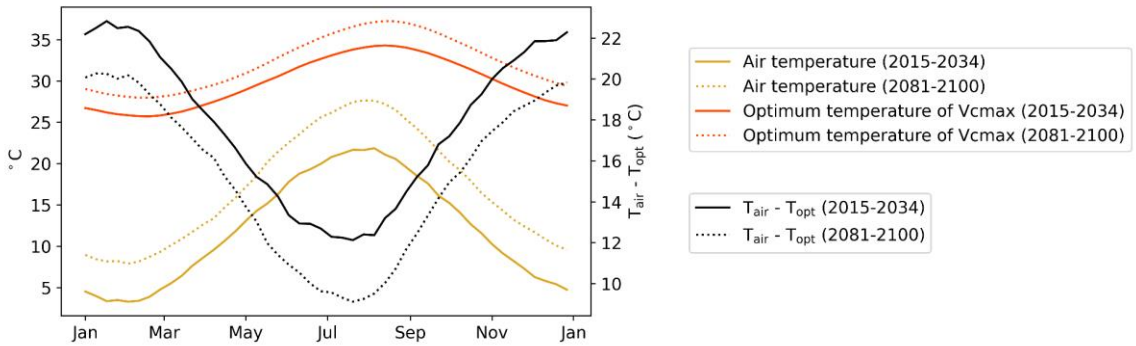

Figure S15: Mean seasonal cycle of air temperature (T_{air} in °C) (yellow), optimum temperature of Vcmax (T_{opt} in °C) (orange), and the difference between T_{air} and T_{opt} (°C) (black) over Europe, projected under the SSP370 scenario for the period 2015-2034 and 2081-2100. Data are 8-day averages with a 2° spatial resolution.

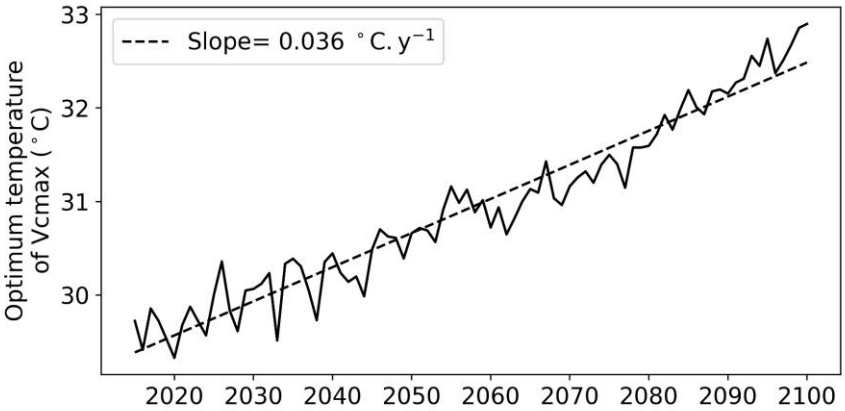

Figure S16: Projected variations of the optimum temperature of Vcmax (°C) for C_3 plants over Europe from 2015 to 2100 under the SSP370 scenario. The mean annual values are represented in solid lines and the corresponding linear regressions are in dotted lines with the regression coefficient values (slope).

6.3 Future prospects for evaluating and integrating vegetation physiological acclimation to atmospheric moisture stress in ORCHIDEE

In the above study, the acclimation function that modulates vegetation response to soil moisture stress was evaluated against eddy covariance GPP and LE data at the site scale, and against satellite-based GPP and LE products at the regional scale across Europe. For future research, it would be valuable to use independent datasets for further evaluation of this function. For example, tree ring width data, which has been shown to correlate with GPP (section 1.3.2), could be utilized by comparing the relative changes in tree ring width with the ones in simulated GPP at various sites that have experienced drought events. This approach could help determine whether incorporating acclimation to long-term VPD in the simulated response to soil moisture stress improves the correlation between relative changes in tree ring width and in GPP in ORCHIDEE. Furthermore, integrating this acclimation function in the ORCHIDEE version that simulates tree ring width variability (Barichivich et al, 2021) could enable further calibration of the acclimation function's parameters through the assimilation of tree ring data. Indeed, Barichivich et al. (2021) demonstrated that tree ring width as well as carbon and oxygen stable isotope data could be used not only to evaluate the growth variability simulated in ORCHIDEE but also to constrain the model's stomatal response to drought stress.

Another independent evaluation of the acclimation function could be conducted using atmospheric CO_2 concentration data from measurement towers, such as those in the ICOS network (https://www.icos-cp.eu/), which includes 46 towers across 16 European countries. Since tall tower atmospheric CO_2 concentration measurements capture influences from broader areas than local phenomena, they can provide insights into surface processes at the regional scale. For this evaluation, it would be necessary to transport all surface CO_2 fluxes contributing to atmospheric CO_2 concentration - including NEE, which is the difference between GPP and ecosystem respiration - in order to simulate atmospheric CO_2 concentrations. This approach would allow us to assess whether the simulated CO_2 concentrations derived from the NEE that incorporate the acclimation function improve the agreement with observed atmospheric CO_2 concentrations. The evaluation could be conducted using the Community Inversion Framework (CIF) (Berchet et al., 2021), a platform that supports running multiple transport models, such as LMDZ (Chevallier et al., 2005), FLEXPART (Pisso et al., 2019), and the Transport Model version 5 (Krol et al., 2005), enabling the quantification and comparison of errors arising from atmospheric transport modeling.

The study presented in the previous section encourages the integration of new mechanisms of vegetation responses to environmental changes in LSMs by incorporating plants' ability to acclimate. The focus of this work is on the physiological response of vegetation to soil moisture stress, particularly stomatal regulation and the coordinated non-stomatal responses, including mesophyll conductance and photosynthetic activity. Although the acclimation function developed in this study was designed for a model version that does not represent plant hydraulic architecture, future research could explore how vegetation acclimation can be integrated in physiological responses to soil moisture changes in a model that includes an explicit

representation of plant hydraulic architecture. Note that a hydraulic architecture has already been integrated in ORCHIDEE-CAN (Yao et al., 2022). However, in that version, the hydraulic architecture determines plant transpiration based on leaf water potential only during soil moisture stress conditions. In contrast, the new hydraulic architecture developed in Julien Alléon's PhD work will enable the computation of plant transpiration based on leaf water potential under any conditions. Additionally, this new representation includes a root absorption model that relies on solving a diffusion equation radially around the roots.

Then, the integration of an hydraulic architecture in ORCHIDEE enables explicit simulation of water storage in the plant, and water movement through the soil, roots, stems, leaves, and atmosphere continuum, based on water potential gradients and specific conductivity parameters for each organ. This model version could offer new insights on potential stomatal acclimation to long term environmental conditions influencing vegetation response to drought events. By representing stomatal conductance as a function of leaf water potential, this version may reveal different sensitivities to environmental drivers compared to the ORCHIDEE version without hydraulic architecture. Additionally, this version provides a framework for exploring potential acclimation functions that could influence simulated hydraulic dysfunction or failure mechanisms, such as the percentage loss of conductance (Adams et al., 2017) or xylem vessel embolism (Cochard and Delzon, 2013). Incorporating these acclimation responses into the hydraulic architecture could improve the simulation of stomatal response to soil moisture stress and drought-induced mortality in LSMs.

Finally, the ORCHIDEE version used in this PhD work does not explicitly represent the nitrogen cycle. The nitrogen cycle interacts with the carbon cycle, influencing GPP and its response to climate change with rising atmospheric CO_2 (Reich et al., 2014; Zaehle et al., 2015). Vuichard et al. (2019) evaluated an ORCHIDEE version that includes carbon-nitrogen interactions and found that, under elevated CO_2, the simulated GPP and plant transpiration were significantly lower than those simulated with a model version lacking an explicit nitrogen cycle representation. Therefore, using a model version that accounts for carbon-nitrogen interactions can lead to different vegetation responses to environmental changes and different associated potential acclimation mechanisms.

7 CONCLUSION AND PERSPECTIVES

7.1 General conclusion

GPP and plant transpiration are essential components of the carbon and water cycles, playing a crucial role in ecosystem functioning and their interactions with the atmosphere. Accurately representing GPP and plant transpiration in LSMs is essential for reliable climate simulations and projections, which are key to informing policy decisions. Although various proxies have been used to independently constrain these processes, there remains significant uncertainty, necessitating the investigation of new proxies to enhance our understanding. The primary objective of this PhD thesis was to provide new constraints on GPP and plant transpiration, particularly by exploring the potential of COS as a proxy for these fluxes, with the aim of improving their representation in the ORCHIDEE LSM. To address the research questions central to this work, I implemented ecosystem COS flux representation in ORCHIDEE and employed DA techniques to enhance the model's process representation, using *in situ* flux data, including ecosystem COS flux observations. While the study of ecosystem COS flux observations has evidenced the potential of COS to provide specific constraints on GPP and plant transpiration, it has also revealed limitations in using COS as a proxy for these fluxes, as well as deficiencies in GPP and plant transpiration representation during drought events. These findings led me to focus on better characterizing the dynamics of GPP and plant transpiration under soil moisture stress conditions, to more accurately account for the diversity in vegetation responses and the capacity of vegetation to acclimate to long-term VPD conditions. The main conclusions to the research questions presented in section 1.4 are summarized below.

In this PhD work, I initially focused on integrating a mechanistic representation of ecosystem COS fluxes within ORCHIDEE by implementing a model for soil COS exchanges to complement the existing vegetation COS uptake model. This effort aligns with the growing interest in mechanistic representations of surface COS fluxes to better characterize the contributions of various COS components to the global atmospheric COS budget and to use COS as a proxy for GPP and plant transpiration. The increasing interest in COS's potential to constrain GPP and plant transpiration in LSMs is evidenced by the rising number of LSMs that have incorporated more complex ecosystem COS flux models, moving beyond the earlier empirical representations. This began with SiB (Berry et al., 2013; Kooijmans et al., 2021), followed by ORCHIDEE (Maignan et al., 2021; Abadie et al., 2022), then BEPS (Zhu et al., 2024a, 2024b), and COS is currently being integrated into the Joint United Kingdom Land Environment Simulator (JULES). The inclusion of ecosystem COS fluxes in multiple LSMs allows for the simulation of vegetation and soil COS flux temporal and spatial dynamics on a global scale while accounting for uncertainties arising from differences in flux representations and parameterizations. In ORCHIDEE, the mechanistic-based representation of soil COS exchanges has enabled me to compare the contributions from anoxic soils - often neglected in the global atmospheric COS budget (Remaud et al., 2022, 2023) - with those from oxic soils, revealing a significant contribution from anoxic soils, producing +96 GgS y^{-1} compared to a net oxic soil sink of -126 GgS y^{-1} over 2009-2016. During this implementation phase, I also refined

the vegetation COS uptake model in ORCHIDEE by incorporating spatially and temporally varying atmospheric COS concentrations to drive simulated COS fluxes, while the vegetation COS model previously used a constant COS concentration. This refinement is crucial for the global atmospheric COS budget, as using a variable COS concentration has reduced the net COS uptake by vegetation and soil by 8% and 10%, respectively, compared to using a constant concentration. As our understanding of vegetation and soil COS fluxes continues to advance, with experimental studies revealing new processes that contribute to ecosystem COS fluxes, such as the role of living roots (Kitz et al., 2024) or the potential of crops to emit COS (Belviso et al., 2022a), the COS models in ORCHIDEE should gradually evolve to incorporate these new processes.

The integration of a mechanistic-based representation of ecosystem COS flux in ORCHIDEE has enabled the use of COS flux as a proxy to provide new constraints on GPP and plant transpiration parameterization. Thanks to the implementation of biospheric COS flux models in LSMs, eddy-covariance ecosystem COS flux observations have begun to be assimilated in LSMs to directly optimize model parameters. In this PhD work, I performed the first assimilation of ecosystem COS flux data in ORCHIDEE using the longest available time series of such measurements from the Hyytiälä boreal evergreen needleleaf forest in Finland. This data assimilation framework revealed that the simulated vegetation COS uptake is more sensitive to stomatal conductance parameters than GPP, underscoring the potential of COS flux to specifically constrain stomatal regulation, which couples GPP and transpiration fluxes. Due to the stronger constraint on stomatal conductance provided by COS flux, this study found that assimilating ecosystem COS fluxes in addition to GPP data improves not only the simulated COS fluxes and GPP at Hyytiälä but also LE, a result not achieved when assimilating GPP data alone. In addition to emphasizing the value of COS fluxes for constraining GPP and LE in LSMs, the assimilation of ecosystem COS flux data also offered a revised estimate of the contribution of boreal evergreen needleleaf forests to the global atmospheric COS budget. This revision suggests a stronger biospheric COS sink in the Northern high latitudes, aligning with the previously identified missing COS sink in this region (Ma et al., 2021; Remaud et al., 2022).

Working with COS fluxes in ORCHIDEE also underscored limitations in using this proxy to improve the representation of GPP and plant transpiration in LSMs. A key limitation is the current scarcity of biospheric COS flux observations at the spatial and temporal scales required to inform LSM processes. Exploiting the full potential of COS would be greatly facilitated by establishing a large network of *in situ* COS flux observations, such as FLUXNET for CO_2, H_2O, and energy fluxes. At present, the limited number of sites with ecosystem COS flux observations prevents the calibration of parameters for all PFTs represented in LSMs, and typically does not capture a full seasonal cycle. Furthermore, conducting biospheric COS flux measurements under conditions that simulate future climates, such as elevated CO_2 (White et al., 2010), could be valuable for informing simulations of GPP and plant transpiration in climate projections. Another important limitation in using COS as a proxy for GPP and plant transpiration is the incomplete understanding of the mechanisms governing COS surface fluxes. The identification of new contributions to ecosystem COS fluxes by experimental studies (Kitz et al., 2024; Belviso et al., 2022a), which are not yet accounted for in current LSM representations, adds complexity to modeling ecosystem COS fluxes and could complicate efforts to constrain GPP and transpiration using COS flux observations alone. One possible solution to address these limitations is adopting a multi-tracer approach.

Through this PhD thesis, I extended my work on biospheric COS fluxes to the study of atmospheric COS concentrations by contributing to several modeling and experimental studies. In these collaborations, ORCHIDEE simulated vegetation COS uptake and soil COS exchanges were used to investigate other biospheric COS fluxes not yet included in the model, such as the potential for crops to emit COS, relying on local COS concentration measurements in two French agroecosystems. Observations of a recent decrease in atmospheric COS concentrations from a tall tower in France led to an assessment of how the atmospheric COS trend affects simulated ecosystem COS fluxes in ORCHIDEE, highlighting the need to account for variable atmospheric COS concentrations in COS flux models implemented in LSMs, as recent studies have shown (Kooijmans et al., 2021; Abadie et al., 2022). Then, comparisons between observed and simulated atmospheric COS concentrations, following the transport of all surface COS flux components, revealed discrepancies in COS anthropogenic emissions over Europe, as estimated by the widely used inventory of Zumkehr et al. (2018) (Ma et al., 2021; Remaud et al., 2022; 2023). The transport studies involving ORCHIDEE simulated biospheric COS fluxes also provided additional evaluation of the simulated ecosystem COS fluxes, identifying potential regional missing COS sinks to address the current imbalance in the atmospheric COS budget. This highlighted areas for future improvement in surface COS flux estimates, such as the underestimation of the simulated biospheric COS uptake in Northern high latitudes. Finally, atmospheric COS and CO_2 concentrations have been used in transport and inversion frameworks to provide insights into two critical components of the global CO_2 cycle: GPP and respiration. These studies illustrate the potential of using COS and CO_2 concentration data to constrain these surface fluxes simulated in LSMs and to serve as an additional evaluation in model intercomparison exercises.

During the study of COS, I highlighted deficiencies in the representation of GPP and plant transpiration under soil moisture stress conditions when examining a drought event at the Hyytiälä site. After identifying the current limitations in how vegetation physiological response to drought is modeled in ORCHIDEE, I focused on improving the simulation of GPP and plant transpiration dynamics under soil moisture stress for the final part of my PhD thesis. This work is centered on Europe, which has experienced significant soil moisture stress in recent years (Bastos et al., 2020), with the ICOS eddy covariance tower network capturing the effects on CO_2 and H_2O fluxes. I assimilated *in situ* GPP and LE data from over 40 sites across six different biomes in Europe, which showed that the speed of stomatal closure in response to decreasing soil moisture can be modeled as a function of the rolling 12-month average of VPD, with PFT-dependent parameters. Integrating such a response in LSMs enables to account for the diversity of drought responses not just between different PFTs but also within the same PFT under varying environmental conditions. I conducted site-scale evaluation under soil moisture stress, which revealed that incorporating long-term VPD acclimation in ORCHIDEE improved the simulated GPP by 18% (in terms of RMSD reduction) after DA, whereas a response based solely on vegetation type showed no improvement. At the European scale, an evaluation with three satellite-based GPP products demonstrated a 7% to 9% improvement in simulated GPP over drought-affected areas in 2020 using the acclimation-based response, compared to a maximum of 3% improvement with the vegetation-type-dependent response. Furthermore, only the model version with long-term VPD acclimation improved the simulated LAI across Europe. However, both model versions

performed similarly in simulating LE at both site and regional scales, with an 8 to 9% improvement at the site level but a slight degradation at the regional scale compared to the prior. Then, I performed projections until 2100 under the SSP370 scenario, considering the projected rise in global VPD, which showed that the model incorporating VPD acclimation resulted in the lowest increasing trend in soil moisture stress, 22% lower than the model relying solely on a vegetation-type-dependent response. This study underscores the importance of including long-term acclimation mechanisms, such as the effect of VPD, in LSMs to improve the accuracy of GPP and plant transpiration simulations beyond drought events. It highlights the broader value of incorporating acclimation processes into vegetation responses to environmental changes in LSMs to improve the reliability of climate projections and potentially mitigate some of the adverse effects of climate change.

7.2 Perspectives

7.2.1 Towards a multiproxy approach

Over the years, numerous methods have been developed to estimate GPP and plant transpiration, including relying on various proxy-based approaches (section 1.3). However, as demonstrated in this PhD thesis with COS, no single proxy perfectly captures GPP or plant transpiration. Each proxy offers distinct advantages, comes with specific limitations and uncertainties, and operates at particular temporal and spatial scales. To address the shortcomings inherent in individual proxies, a multiproxy approach could be adopted. By combining the strengths of multiple proxies, this approach could provide a more robust and comprehensive understanding of ecosystem processes.

A multiproxy approach takes advantage of the complementary nature of different proxies, each providing unique insights into different aspects of plant function. For example, COS informs on stomatal conductance (Seibt et al., 2010; Wehr et al., 2017; Wohlfahrt et al., 2012; Berkelhammer et al., 2020), which couples GPP and plant transpiration. In contrast, SIF relates to the light-dependent phase of photosynthesis (Frankenberg & Berry, 2018) and can provide information on photosynthetic efficiency, responding almost instantaneously to changes in the photosynthetic machinery (Porcar-Castell et al., 2014). Additionally, the photochemical reflectance index (PRI), derived from narrow-band reflectance observations (Gamon et al., 1992), is associated with non-photochemical quenching (NPQ) - a third energy pathway that competes with photochemistry and fluorescence - and has been linked to variations in photosynthetic efficiency (Garbulsky et al., 2011). In contrast, stable isotopes of water, such as ^{18}O, offer insights into plant water use sources and transpiration rates (Sheshshayee et al., 2005). Moreover, sap flow measurements, including those from the SAPFLUXNET database (Poyatos et al., 2021), provide direct information on plant water transport, which is closely related to transpiration rates (Flo et al., 2019). Then, high-frequency vegetation optical depth (VOD) satellite retrievals, such as the X- and Ku-band, are sensitive to changes in upper-canopy water content as well as biomass, and can provide insights into stomatal regulation on a diurnal timescale (Konings et al., 2021).

The complementarity among various proxies can also be leveraged to address different temporal and spatial resolutions. Proxies that can be measured from space, such as SIF, PRI, VOD, or traditional VIs like NDVI and EVI, provide global coverage (Liao et al., 2023; Garbulsky et al., 2011). In contrast, some proxies, such as sap flow and stable isotope measurements, offer detailed insights at the leaf or individual tree level (Dai et al., 2023; Poyatos et al., 2021). COS, on the other hand, provides information from the leaf level with chamber measurements (Seibt et al., 2010; Kooijmans et al., 2019) to the global scale with satellite atmospheric COS concentration retrievals from MIPAS (Glatthor et al., 2015). Proxies available at smaller scales typically offer finer temporal resolution, with sap flow and chamber or eddy covariance COS measurements that can be available at sub-hourly intervals (Poyatos et al., 2021; Seibt et al., 2010; Berkelhammer et al., 2014; Commane et al., 2015; Kooijmans et al., 2019; Wehr et al., 2017), whereas satellite-based proxies generally provide data at daily or coarser resolutions (Garbulsky et al., 2011). Therefore, combining multiple proxies allows for a more detailed and comprehensive spatial and temporal study of GPP and plant transpiration.

Moreover, different proxies exhibit varying sensitivities to environmental factors, making a multiproxy approach more robust under diverse environmental conditions. For example, during a heatwave in a Mediterranean pine forest, Wohlfahrt et al. (2018) used COS flux measurements to show that the early decline in GPP was linked to stomatal limitations, while SIF remained nearly constant and did not reflect the GPP variations during this phase. However, the subsequent drop in SIF during the peak of the heatwave corresponded with biochemical impairment of photosynthesis. Additionally, VOD retrievals can be particularly valuable under drought conditions, as they have been found to capture vegetation physiological changes, informing on plant water-saving strategies through stomatal regulation (Li et al., 2023).

Finally, because different proxies are based on distinct measurement techniques, combining them can reduce uncertainty in GPP and plant transpiration estimates by minimizing the potential bias that could result from relying on a single methodology.

Since GPP and plant transpiration are influenced by a variety of biophysical processes, such as stomatal conductance, photosynthetic efficiency, and water transport, all of which are governed by multiple environmental factors, adopting a multiproxy approach allows these processes to be more comprehensively constrained in LSMs. Given that various proxies respond differently to environmental changes, incorporating multiple proxies could enable LSMs to more accurately simulate GPP and plant transpiration dynamics under varying climate conditions. Additionally, calibrating model parameters through DA that incorporates proxies with varying temporal and spatial resolutions can help bridge the gap between short-term and long-term observations, as well as between small-scale and large-scale observations.

For ORCHIDEE, beyond the COS flux models implemented as presented in this work, a representation of SIF has also been previously integrated (Bacour et al., 2019). The SIF model in ORCHIDEE is currently being revised to differentiate between fluorescence emissions from sunlit and shaded leaves (Maignan et al., in prep), and the representation of non-photochemical quenching (NPQ) is likewise being updated (Leverne et al., in revision). PRI data could also be used to inform the simulated NPQ, which is currently constrained only by measurements representative of three PFTs (Bacour et al., 2019). A next step would be to implement an observation operator for VOD, potentially linking it to simulated stomatal conductance, or leaf water potential (Liu et al.,

2021) in a model version that incorporates an explicit hydraulic architecture. Integrating all these proxy representations in ORCHIDEE paves the way for multiproxy DA, allowing for the optimization of a broader range of parameters than what is achievable through the assimilation of a single proxy to further constrain the various processes that determine GPP and plant transpiration.

A multiproxy approach presents numerous opportunities for DA to constrain GPP and plant transpiration in LSMs. A stepwise approach could be adopted given the high computational cost associated with assimilating multiple datastreams (Peylin et al., 2016). In this approach, the different proxies would be assimilated sequentially, with information from each step informing the next. Specifically, this approach could take advantage of the varying spatial and temporal resolutions of the different observational data. Initial steps could focus on assimilating global-scale satellite data, such as SIF, VOD, or PRI, to capture a wide range of environmental conditions and PFTs by selecting homogeneous pixels around the world. Following this, finer-scale data would be assimilated to provide more precise constraints on the processes determining GPP and plant transpiration, including regional information from atmospheric COS concentrations and *in situ* measurements such as COS flux data and sap flow data. This hierarchical strategy could help assess how each proxy and assimilation step improves the overall model accuracy in simulating GPP and plant transpiration.
Furthermore, in a multiproxy DA approach, considering spatially variable parameters in the representations of GPP and plant transpiration, instead of static values, during DA could help move beyond the traditional reliance on PFT-based parameterization in LSMs. This approach would allow parameters to be optimized to better reflect the influence of local environmental conditions. In particular, such optimization could help determine whether certain parameters representing plant traits should be redefined as functions that more accurately capture vegetation potential to acclimate to varying environmental conditions. This highlights the importance of having access to global-scale proxy observation data, such as satellite-based retrievals, as well as the need to develop extensive *in situ* observation networks to provide data across diverse ecosystems worldwide.

7.2.2 Integrating optimality principles

DA aims to improve the simulation of GPP and plant transpiration in LSMs by optimizing parameters. This approach is increasingly supported by the growing quantity and diversity of remote sensing and field observations (Duncanson et al., 2019; Houborg et al., 2015). However, another challenge for LSMs is determining whether certain parameters should be assigned values that differ only between PFTs or soil texture types for example, or if they should be treated as model parameters that optimize plant traits in response to environmental conditions (Fisher et al., 2020).

Therefore, alternative approaches can be investigated to address parametric uncertainty in LSMs. Notably, optimality approaches focus on optimizing a functional relationship that balances the costs and benefits for a specific plant trait, as modulated by environmental factors. The underlying principle is that evolutionary processes have led to the selection of trait values that maximize

survival in given environments. Consequently, optimality approaches offer parameter-sparse representations of plant processes and facilitate the modeling of vegetation adjustments to environmental changes. Notably, optimization theories can include plant plasticity and acclimation to environmental drivers (Joshi et al., 2022).

For instance, optimality principles have been applied to model stomatal conductance, traditionally based on the idea that stomata operate to maximize CO_2 uptake for photosynthesis while minimizing water loss through transpiration (Prentice et al., 2014; Wolf et al., 2016). However, a significant challenge with this approach lies in accurately formulating the cost of stomatal opening and determining the appropriate constraints to impose (Dewar et al., 2018). To address this, various optimization theories have been developed: some balance carbon gain with the risk of vascular damage from embolism formation (Sperry et al., 2017), while others balance carbon gain against water loss, weighted by canopy hydraulic conductivity (Zhu et al., 2023). These different stomatal optimization models consequently predict varying sensitivities of stomatal conductance to environmental stresses, such as decreasing soil moisture or increasing VPD (Sabot et al., 2022; Zhu et al., 2023).

Overall, these stomatal optimization models offer the advantage of requiring relatively few parameters (Dewar et al., 2009) and avoid the need to explicitly prescribe the direct effects of environmental factors on stomatal conductance. In contrast, most LSMs rely on empirical representations of stomatal behavior (Ball & Berry, 1987; Leuning, 1995) due to the incomplete understanding of stomatal regulation needed to implement a mechanistic representation (Wolf et al., 2016). By integrating optimality principles, LSMs could more accurately capture the rapid and acclimated responses of photosynthesis, plant transpiration, and stomatal conductance, while reducing the number of parameters that contribute to uncertainties in climate simulations and projections. To further validate the effectiveness of these principles, it would be valuable to conduct large-scale assessment of their impact on the simulated GPP and plant transpiration, and intercomparisons between different LSMs. Such studies could reveal whether the incorporation of optimality principles leads to better agreement in simulated GPP and plant transpiration across various models, ultimately improving the predictive capability and reliability of LSMs.

7.2.3 Closing remarks

To conclude, the use of multiple proxies to inform GPP and plant transpiration, even when these processes incorporate optimality principles with relatively few parameters, relies heavily on the availability of observational data for continuous improvement in process representation, parameterization, and evaluation of LSMs. Strengthening collaborations between experimental and modeling communities is crucial to effectively use and develop new observational data. For example, alongside observational data that inform ecosystem functioning and responses to current environmental conditions, experimental efforts like the Free-Air CO_2 Enrichment (FACE) experiments (Ainsworth et al., 2005) are crucial for understanding ecosystem responses to future climate scenarios and for improving climate projections in LSMs. LSMs enable the investigation of the impacts of small-scale processes on larger temporal and spatial scales. However, LSMs can

also help identify critical data needs and guide future experimental studies by generating new hypotheses about ecosystem functioning. Therefore, recognizing and enhancing the continuity between experimental and modeling efforts can significantly improve LSMs ability to simulate climate and vegetation interactions, ultimately contributing to better ecological and climate-related decision-making.

During my PhD, I had the opportunity to collaborate with several experts from the experimental community to combine modeling and experimental efforts. This experience highlighted the importance and advantages of building strong connections between modeling and experimental research, which I would like to continue developing in my future research career.
In addition to these collaborations, my work with the ORCHIDEE LSM over the past few years has motivated me to explore other LSMs to better understand the diversity of approaches and formalisms used to represent climate and vegetation interactions. In particular, I aim to deepen my expertise by working with an LSM that integrates the coupling between the carbon, nitrogen, and phosphorus cycles. In the future, I am also interested in conducting coupled simulations with an atmospheric circulation model to better account for land-atmosphere interactions and climate feedback mechanisms.

8 REFERENCES

Abadie, C., Maignan, F., Remaud, M., Kohonen, K., Sun, W., Kooijmans, L., Vesala, T., Seibt, U., Raoult, N., Bastrikov, V., Belviso, S., & Peylin, P. (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*(7). https://doi.org/10.1029/2023JG007407

Abadie, C., Maignan, F., Remaud, M., Ogee, J., Campbell, J. E., Whelan, M. E., Kitz, F., Spielmann, F. M., Wohlfahrt, G., Wehr, R., Sun, W., Raoult, N., Seibt, U., Hauglustaine, D., Lennartz, S. T., Belviso, S., Montagne, D., & Peylin, P. (2022). Global modelling of soil carbonyl sulfide exchanges. *Biogeosciences*, *19*(9), 2427–2463. https://doi.org/10.5194/bg-19-2427-2022

Adams, H. D., Zeppel, M. J. B., Anderegg, W. R. L., Hartmann, H., Landhäusser, S. M., Tissue, D. T., Huxman, T. E., Hudson, P. J., Franz, T. E., Allen, C. D., Anderegg, L. D. L., Barron-Gafford, G. A., Beerling, D. J., Breshears, D. D., Brodribb, T. J., Bugmann, H., Cobb, R. C., Collins, A. D., Dickman, L. T., … McDowell, N. G. (2017). A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature Ecology & Evolution*, *1*(9), 1285–1291. https://doi.org/10.1038/s41559-017-0248-x

Ainsworth, E. A., & Long, S. P. (2005). What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. In *New Phytologist* (Vol. 165, Issue 2, pp. 351–372). https://doi.org/10.1111/j.1469-8137.2004.01224.x

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Molecular Biology of the Cell* (4th edition). Garland Science.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., & Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. In *Reviews of Geophysics* (Vol. 53, Issue 3, pp. 785–818). Blackwell Publishing Ltd. 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

Babst, F., Bouriaud, O., Papale, D., Gielen, B., Janssens, I. A., Nikinmaa, E., Ibrom, A., Wu, J., Bernhofer, C., Köstner, B., Grünwald, T., Seufert, G., Ciais, P., & Frank, D. (2014). Above-ground woody carbon sequestration measured from tree rings is coherent with net ecosystem productivity at five eddy-covariance sites. *New Phytologist*, *201*(4), 1289–1303. https://doi.org/10.1111/nph.12589

Bacour, C., Macbean, N., Chevallier, F., Léonard, S., Koffi, E. N., & Peylin, P. (2023). Assimilation of multiple datasets results in large differences in regional-to global-scale NEE and GPP budgets simulated by a terrestrial biosphere model. *Biogeosciences*, *20*(6), 1089–1111. https://doi.org/10.5194/bg-20-1089-2023

Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., Peylin, P., Chevallier, F., Vuichard, N., & Bastrikov, V. (2019). Improving Estimates of Gross Primary Productivity by Assimilating Solar-Induced Fluorescence Satellite Retrievals in a Terrestrial Biosphere Model Using a Process-Based SIF Model. *Journal of Geophysical Research: Biogeosciences*, *124*(11), 3281–3306. https://doi.org/10.1029/2019JG005040

Badgley, G., Anderegg, L. D. L., Berry, J. A., & Field, C. B. (2019). Terrestrial gross primary production: Using NIRV to scale from site to globe. *Global Change Biology*, *25*(11), 3731–3740. https://doi.org/10.1111/gcb.14729

Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances.* http://advances.sciencemag.org/cgi/

Bailey, V. L., Bond-Lamberty, B., DeAngelis, K., Grandy, A. S., Hawkes, C. v., Heckman, K., Lajtha, K., Phillips, R. P., Sulman, B. N., Todd-Brown, K. E. O., & Wallenstein, M. D. (2018). Soil carbon cycling proxies: Understanding their critical role in predicting climate change feedbacks. In *Global Change Biology* (Vol. 24, Issue 3, pp. 895–905). Blackwell Publishing Ltd. https://doi.org/10.1111/gcb.13926

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions. In *Progress in Photosynthesis Research* (pp. 221–224). Springer Netherlands. https://doi.org/10.1007/978-94-017-0519-6_48

Barichivich, J., Peylin, P., Launois, T., Daux, V., Risi, C., Jeong, J., & Luyssaert, S. (2021). A triple tree-ring constraint for tree growth and physiology in a global land surface model. *Biogeosciences*, *18*(12), 3781–3803. https://doi.org/10.5194/bg-18-3781-2021

Barraza, V., Restrepo-Coupe, N., Huete, A., Grings, F., Beringer, J., Cleverly, J., & Eamus, D. (2017). Estimation of latent heat flux over savannah vegetation across the North Australian Tropical Transect from multiple sensors and global meteorological data. *Agricultural and Forest Meteorology*, *232*, 689–703. https://doi.org/10.1016/j.agrformet.2016.10.013

Bastos, A., Ciais, P., Friedlingstein, P., Sitch, S., Pongratz, J., Fan, L., Wigneron, J. P., Weber, U., Reichstein, M., Fu, Z., Anthoni, P., Arneth, A., Haverd, V., Jain, A. K., Joetzjer, E., Knauer, J., Lienert, S., Loughran, T., McGuire, P. C., … Zaehle, S. (2020). Direct and seasonal legacy effects of the 2018 heat wave and drought on European ecosystem productivity. *Science Advances*, *6*(24). https://doi.org/10.1126/sciadv.aba2724

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

Belviso, S., Pison, I., Petit, J.-E., Berchet, A., Remaud, M., Simon, L., Ramonet, M., Delmotte, M., Kazan, V., Yver-Kwok, C., & Lopez, M. (2023). The Z-2018 emissions inventory of COS in Europe: A semiquantitative multi-data-streams evaluation. *Atmospheric Environment*, *300*, 119689. https://doi.org/10.1016/j.atmosenv.2023.119689

Belviso, S., Abadie, C., Montagne, D., Hadjar, D., Tropée, D., Vialettes, L., Kazan, V., Delmotte, M., Maignan, F., Remaud, M., Ramonet, M., Lopez, M., Yver-Kwok, C., & Ciais, P. (2022a). Carbonyl sulfide (COS) emissions in two agroecosystems in central France. *PLoS ONE*, *17*(12 December). https://doi.org/10.1371/journal.pone.0278584

Belviso, S., Remaud, M., Abadie, C., Maignan, F., Ramonet, M., & Peylin, P. (2022b). Ongoing Decline in the Atmospheric COS Seasonal Cycle Amplitude over Western Europe: Implications for Surface Fluxes. *Atmosphere*, *13*(5). https://doi.org/10.3390/atmos13050812

Belviso, S., Lebegue, B., Ramonet, M., Kazan, V., Pison, I., Berchet, A., Delmotte, M., Yver-Kwok, C., Montagne, D., & Ciais, P. (2020). A top-down approach of sources and non-photosynthetic sinks of carbonyl sulfide from atmospheric measurements over multiple years in the Paris region (France). *PLoS ONE*, *15*(2). https://doi.org/10.1371/journal.pone.0228419

Berchet, A., Pison, I., Huselstein, C., Narbaud, C., Remaud, M., Belviso, S., Abadie, C., & Maignan, F. (in review). Can we gain knowledge on COS anthropogenic and biogenic emissions from a single atmospheric mixing ratios measurement site? *EGUsphere* [preprint]. https://doi.org/10.5194/egusphere-2024-549

Berchet, A., Sollum, E., Thompson, R. L., Pison, I., Thanwerdas, J., Broquet, G., Chevallier, F., Aalto, T., Berchet, A., Bergamaschi, P., Brunner, D., Engelen, R., Fortems-Cheiney, A., Gerbig, C., Groot Zwaaftink, C. D., Haussaire, J. M., Henne, S., Houweling, S., Karstens, U., … Zhao, Y. (2021). The Community Inversion Framework v1.0: A unified system for atmospheric inversion studies. *Geoscientific Model Development*, *14*(8), 5331–5354. https://doi.org/10.5194/gmd-14-5331-2021

Berkelhammer, M., Alsip, B., Matamala, R., Cook, D., Whelan, M. E., Joo, E., Bernacchi, C., Miller, J., & Meyers, T. (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., Provencal, R., Chen, H., & Yakir, D. (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., & Bjorkman, O. (1980) Photosynthetic Response and Adaptation to Temperature in Higher Plants. *Annual Review of Plant Biology,* 31, 491-543. http://dx.doi.org/10.1146/annurev.pp.31.060180.002423

Berry, J., Wolf, A., Campbell, J. E., Baker, I., Blake, N., Blake, D., Denning, A. S., Kawa, S. R., Montzka, S. A., Seibt, U., Stimler, K., Yakir, D., & Zhu, Z. (2013). A coupled model of the global cycles of carbonyl sulfide and CO_2: A possible new window on the carbon cycle. *Journal of Geophysical Research: Biogeosciences*, *118*(2), 842–852. https://doi.org/10.1002/jgrg.20068

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. In *Advances in Water Resources* (Vol. 16).

Blažević, I., Montaut, S., Burčul, F., Olsen, C. E., Burow, M., Rollin, P., & Agerbirk, N. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry*, *169*, 112100. https://doi.org/10.1016/j.phytochem.2019.112100

Bloem, E., Haneklaus, S., Kesselmeier, J., & Schnug, E. (2012). Sulfur Fertilization and Fungal Infections Affect the Exchange of H_2S and COS from Agricultural Crops. *Journal of Agricultural and Food Chemistry*, *60*(31), 7588–7596. https://doi.org/10.1021/jf301912h

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. In *Science* (Vol. 359, Issue 6375). American Association for the Advancement of Science. https://doi.org/10.1126/science.aam8328

Bonan, G. B., Pollard, D., & Thompson, S. L. (1992). Effects of boreal forest vegetation on global climate. *Nature*, *359*(6397), 716–718. https://doi.org/10.1038/359716a0

Boretti, A., & Florentine, S. (2019). Atmospheric CO_2 concentration and other limiting factors in the growth of C_3 and C_4 plants. *Plants*, *8*(4). https://doi.org/10.3390/plants8040092

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., … Vuichard, N. (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

Bounoua, L., Collatz, G. J., Sellers, P. J., Randall, D. A., Dazlich, D. A., Los, S O, Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., & Jensen, T. G. (1999). Interactions between Vegetation and Climate: Radiative and Physiological Effects of Doubled Atmospheric CO_2.

Bounoua, L., Hall, F. G., Sellers, P. J., Kumar, A., Collatz, G. J., Tucker, C. J., & Imhoff, M. L. (2010). Quantifying the negative feedback of vegetation to greenhouse warming: A modeling approach. *Geophysical Research Letters*, *37*(23). https://doi.org/10.1029/2010GL045338

Briggs, L. J., & Shantz, H. L. (1913). Water Requirement of Plants. *Bureau of Plant Industry Bulletin.*

Burba, G. (2013). Eddy Covariance Method for Scientific, Industrial, Agricultural and Regulatory Applications: A Field Book on Measuring Ecosystem Gas Exchange and Areal Emission Rates. *Li-Cor Biosciences.*

Campbell, J. E., Whelan, M. E., Seibt, U., Smith, S. J., Berry, J. A., & Hilton, T. W. (2015). Atmospheric carbonyl sulfide sources from anthropogenic activity: Implications for carbon cycle constraints. *Geophysical Research Letters*, *42*(8), 3004–3010. https://doi.org/10.1002/2015GL063445

Cano, F. J., López, R., & Warren, C. R. (2014). Implications of the mesophyll conductance to CO_2 for photosynthesis and water-use efficiency during long-term water stress and recovery in two contrasting Eucalyptus species. *Plant, cell & environment*, *37*(11), 2470–2490. https://doi.org/10.1111/pce.12325

Chaves, M. M., Maroco, J. P., & Pereira, J. S. (2003). Understanding plant responses to drought — from genes to the whole plant. *Functional Plant Biology*, *30*(3), 239. https://doi.org/10.1071/FP02076

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

Chen, C., Riley, W. J., Prentice, I. C., & Keenan, T. F. (2022). CO_2 fertilization of terrestrial photosynthesis inferred from site to global scales. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(10), e2115627119. https://doi.org/10.1073/pnas.2115627119

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J. L., Ghattas, J., Grandpeix, J. Y., Idelkadi, A., Mellul, L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., … Zhao, Y. (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

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A., & Ciais, P. (2005). Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *Journal of Geophysical Research Atmospheres*, *110*(24), 1–13. https://doi.org/10.1029/2005JD006390

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). *Biogeosciences*, *20*(13), 2573–2594. https://doi.org/10.5194/bg-20-2573-2023

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., Arain, M. A., Arkebauer, T. J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Bracho, R., Brown, S., … Zona, D. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301–302*. https://doi.org/10.1016/j.agrformet.2021.108350

Ciais, P., Denning, A. S., Tans, P. P., Berry, J. A., Randall, D. A., Collatz, G. J., Sellers, P. J., White, J. W. C., Trolier, M., Meijer, H. A. J., Francey, R. J., Monfray, P., & Heimann, M. (1997). A three-dimensional synthesis study of $\delta^{18}O$ in atmospheric CO_2: 1. Surface fluxes. *Journal of Geophysical Research Atmospheres*, *102*(5), 5857–5872. https://doi.org/10.1029/96jd02360

Cochard, H., Badel, E., Herbette, S., Delzon, S., Choat, B., & Jansen, S. (2013). Methods for measuring plant vulnerability to cavitation: A critical review. In *Journal of Experimental Botany* (Vol. 64, Issue 15, pp. 4779–4791). https://doi.org/10.1093/jxb/ert193

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C_4 Plants. *Functional Plant Biology*, *19*(5), 519–538. https://doi.org/10.1071/PP9920519

Commane, R., Meredith, L. K., Baker, I. T., Berry, J. A., Munger, J. W., Montzka, S. A., Templer, P. H., Juice, S. M., Zahniser, M. S., & Wofsy, S. C. (2015). Seasonal fluxes of carbonyl sulfide in a midlatitude forest. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(46), 14162–14167. https://doi.org/10.1073/pnas.1504131112

Cornwell, W. K., Wright, I. J., Turner, J., Maire, V., Barbour, M. M., Cernusak, L. A., Dawson, T., Ellsworth, D., Farquhar, G. D., Griffiths, H., Keitel, C., Knohl, A., Reich, P. B., Williams, D. G., Bhaskar, R., Cornelissen, J. H. C., Richards, A., Schmidt, S., Valladares, F., … Santiago, L. S. (2018). Climate and soils together regulate photosynthetic carbon isotope discrimination within C_3 plants worldwide. *Global Ecology and Biogeography*, *27*(9), 1056–1067. https://doi.org/10.1111/geb.12764

Costanza, R., de Groot, R., Farberll, S., Grassot, M., Hannon, B., Limburg, K., Naeem, S., O, R. v, Paruelo, J., Raskin, R. G., & Suttonllll, P. (1997). The value of the world's ecosystem services and natural capital.

Cowling, S. A., & Field, C. B. (2003). Environmental control of leaf area production: Implications for vegetation and land-surface modeling. *Global Biogeochemical Cycles*, *17*(1), 7-1-7–14. https://doi.org/10.1029/2002gb001915

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., Betancourt, J. L., Colt, S., Cravens, A. E., Dalton, M. S., Dunham, J. B., Hay, L. E., Hayes, M. J., McEvoy, J., McNutt, C. A., Moritz, M. A., Nislow, K. H., Raheem, N., & Sanford, T. (2017). Defining ecological drought for the twenty-first century. *Bulletin of the American Meteorological Society*, *98*(12), 2543–2550. https://doi.org/10.1175/BAMS-D-16-0292.1

Creese, C., Oberbauer, S., Rundel, P., & Sack, L. (2014). Are fern stomatal responses to different stimuli coordinated? Testing responses to light, vapor pressure deficit, and CO_2 for diverse species grown under contrasting irradiances. *New Phytologist*, *204*(1), 92–104. https://doi.org/10.1111/nph.12922

Cunningham, S. C. (2004). Stomatal sensitivity to vapour pressure deficit of temperate and tropical evergreen rainforest trees of Australia. *Trees - Structure and Function*, *18*(4), 399–407. https://doi.org/10.1007/s00468-004-0318-y

Cuntz, M., Ciais, P., Hoffmann, G., Allison, C. E., Francey, R. J., Knorr, W., Tans, P. P., White, J. W. C., & Levin, I. (2003). A comprehensive global three-dimensional model of $\delta^{18}O$ in atmospheric CO_2: 2. Mapping the atmospheric signal. *Journal of Geophysical Research: Atmospheres*, *108*(17). https://doi.org/10.1029/2002jd003154

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, *3*(1), 52–58. https://doi.org/10.1038/nclimate1633

Dai, J., Zhang, X., Wang, L., Wang, R., Luo, Z., He, X., & Rao, Z. (2023). Water stable isotope characteristics and water use strategies of co-occurring plants in ecological and economic forests in subtropical monsoon regions. *Journal of Hydrology*, *621*, 129565. https://doi.org/10.1016/j.jhydrol.2023.129565

Daily, G. (1997). Introduction: What Are Ecosystem Services? In *Daily, G.C.* (pp. 1–10). Nature's Services: Societal Dependence on Natural Ecosystems.

Damm, A., Elber, J., Erler, A., Gioli, B., Hamdi, K., Hutjes, R., Kosvancova, M., Meroni, M., Miglietta, F., Moersch, A., Moreno, J., Schickling, A., Sonnenschein, R., Udelhoven, T., van der Linden, S., Hostert, P., & Rascher, U. (2010). Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). *Global Change Biology*, *16*(1), 171–186. https://doi.org/10.1111/j.1365-2486.2009.01908.x

de Pue, J., Barrios, J. M., Liu, L., Ciais, P., Arboleda, A., Hamdi, R., Balzarolo, M., Maignan, F., & Gellens-Meulenberghs, F. (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

Delbart, N., Picard, G., le Toan, T., Kergoat, L., Quegan, S., Woodward, I., Dye, D., & Fedotova, V. (2008). Spring phenology in boreal Eurasia over a nearly century time scale. *Global Change Biology*, *14*(3), 603–614. https://doi.org/10.1111/j.1365-2486.2007.01505.x

Dewar, R. C., Franklin, O., Makela, A., Mcmurtrie, R. E., & Valentine, H. T. (2009). Optimal function explains forest responses to global change. *BioScience*, *59*(2), 127–139. https://doi.org/10.1525/bio.2009.59.2.6

Dewar, R., Mauranen, A., Mäkelä, A., Hölttä, T., Medlyn, B., & Vesala, T. (2018). New insights into the covariation of stomatal, mesophyll and hydraulic conductances from optimization models incorporating nonstomatal limitations to photosynthesis. *New Phytologist*, *217*(2), 571–585. https://doi.org/10.1111/nph.14848

Diefendorf, A. F., Mueller, K. E., Wing, S. L., Koch, P. L., & Freeman, K. H. (2010). Global patterns in leaf ^{13}C discrimination and implications for studies of past and future climate. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5738–5743. https://doi.org/10.1073/pnas.0910513107

D'Odorico, P., Gonsamo, A., Gough, C. M., Bohrer, G., Morison, J., Wilkinson, M., Hanson, P. J., Gianelle, D., Fuentes, J. D., & Buchmann, N. (2015). The match and mismatch between photosynthesis and land surface phenology of deciduous forests. *Agricultural and Forest Meteorology*, *214–215*, 25–38. https://doi.org/10.1016/j.agrformet.2015.07.005

Doughty, R., Wang, Y., Johnson, J., Parazoo, N., Magney, T., Pierrat, Z., Xiao, X., Guanter, L., Köhler, P., Frankenberg, C., Somkuti, P., Ma, S., Qin, Y., Crowell, S., & Moore, B. (2024). A novel data-driven global model of photosynthesis using solar-induced chlorophyll fluorescence. https://doi.org/10.22541/essoar.168167172.20799710/v1

Douville, H., Raghavan, K., Renwick, J., Allan, R. P., Arias, P. A., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T. Y., Gergis, J., Jiang, D., Khan, A., Mba, W. P., Rosenfeld, D., Tierney, J., & and Zolina, O. (2023). Water Cycle Changes. In *Climate Change 2021 – The Physical Science Basis* (pp. 1055–1210). Cambridge University Press. https://doi.org/10.1017/9781009157896.010

Drake, J. E., Power, S. A., Duursma, R. A., Medlyn, B. E., Aspinwall, M. J., Choat, B., Creek, D., Eamus, D., Maier, C., Pfautsch, S., Smith, R. A., Tjoelker, M. G., & Tissue, D. T. (2017). Stomatal and non-stomatal limitations of photosynthesis for four tree species under drought: A comparison of model formulations. *Agricultural and Forest Meteorology*, *247*, 454–466. https://doi.org/10.1016/j.agrformet.2017.08.026

Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., Crowther, T. W., Falkowski, M., Kellner, J. R., Labrière, N., Lucas, R., MacBean, N., McRoberts, R. E., Meyer, V., Næsset, E., Nickeson, J. E., … Williams, M. (2019). The Importance of Consistent Global Forest Aboveground Biomass Product Validation. *Surveys in Geophysics*, *40*(4), 979–999. https://doi.org/10.1007/s10712-019-09538-8

Duveiller, G., Guanter, L., Nelson, J., Hamdi, Z., Gensheimer, J., Dash, J., Ogutu, B., & Bandopadhyay, S. (2023). Dataset for intercomparison of Sen4GPP GPP products at fluxtower sites. Zenodo. https://doi.org/10.5281/zenodo.8339044

Farquhar, G. D., & Lloyd, J. (1993). Carbon and Oxygen Isotope Effects in the Exchange of Carbon Dioxide between Terrestrial Plants and the Atmosphere. In J. R. Ehleringer, A. E. Hall, & G. D.

Farquhar (Eds.), *Stable Isotopes and Plant Carbon-water Relations* (pp. 47–70). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-08-091801-3.50011-8

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A Biochemical Model of Photosynthetic CO_2 Assimilation in Leaves of C_3 Species. In *Planta* (Vol. 149).

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2018MS001453

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., & Bonan, G. (2015). Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development*, *8*(11), 3593–3619. https://doi.org/10.5194/gmd-8-3593-2015

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., … Zander, R. (2008). MIPAS: an instrument for atmospheric and climate research. In *Atmos. Chem. Phys* (Vol. 8). www.atmos-chem-phys.net/8/2151/2008/

Flexas, J., Barbour, M. M., Brendel, O., Cabrera, H. M., Carriquí, M., Díaz-Espejo, A., Douthe, C., Dreyer, E., Ferrio, J. P., Gago, J., Gallé, A., Galmés, J., Kodama, N., Medrano, H., Niinemets, Ü., Peguero-Pina, J. J., Pou, A., Ribas-Carbó, M., Tomás, M., … Warren, C. R. (2012). Mesophyll diffusion conductance to CO_2: An unappreciated central player in photosynthesis. In *Plant Science* (Vols. 193–194, pp. 70–84). https://doi.org/10.1016/j.plantsci.2012.05.009

Flexas, J., Ribas-Carbó, M., Bota, J., Galmés, J., Henkle, M., Martínez-Cañellas, S., & Medrano, H. (2006). Decreased Rubisco activity during water stress is not induced by decreased relative water content but related to conditions of low stomatal conductance and chloroplast CO_2 concentration. *New Phytologist*, *172*(1), 73–82. https://doi.org/10.1111/j.1469-8137.2006.01794.x

Flo, V., Martinez-Vilalta, J., Steppe, K., Schuldt, B., & Poyatos, R. (2019). A synthesis of bias and uncertainty in sap flow methods. *Agricultural and Forest Meteorology*, *271*, 362–374. https://doi.org/10.1016/j.agrformet.2019.03.012

Frankenberg, C., & Berry, J. (2018). Solar Induced Chlorophyll Fluorescence: Origins, Relation to Photosynthesis and Retrieval. In S. Liang (Ed.), *Comprehensive Remote Sensing* (pp. 143–162). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-409548-9.10632-3

Franks, P. J., Cowan, I. R., & Farquhar, G. D. (1997). The apparent feedforward response of stomata to air vapour pressure deficit: Information revealed by different experimental procedures with two rainforest trees. *Plant, Cell and Environment*, *20*(1), 142–145. https://doi.org/10.1046/j.1365-3040.1997.d01-14.x

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., … Zheng, B. (2022). Global Carbon Budget 2022. *Earth System Science Data*, *14*(11), 4811–4900. https://doi.org/10.5194/essd-14-4811-2022

Gamon, J. A., Peñuelas, J., & Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, *41*(1), 35–44. https://doi.org/10.1016/0034-4257(92)90059-S

Gao, J., Zhao, P., Shen, W., Niu, J., Zhu, L., & Ni, G. (2015). Biophysical limits to responses of water flux to vapor pressure deficit in seven tree species with contrasting land use regimes. *Agricultural and Forest Meteorology*, *200*, 258–269. https://doi.org/10.1016/j.agrformet.2014.10.007

Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., & Filella, I. (2011). The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: A review and meta-analysis. *Remote Sensing of Environment*, *115*(2), 281–297. https://doi.org/10.1016/j.rse.2010.08.023

Geng, C., & Mu, Y. (2006). Carbonyl sulfide and dimethyl sulfide exchange between trees and the atmosphere. *Atmospheric Environment*, *40*(7), 1373–1383. https://doi.org/10.1016/j.atmosenv.2005.10.023

Gimeno, T. E., Ogée, J., Royles, J., Gibon, Y., West, J. B., Burlett, R., Jones, S. P., Sauze, J., Wohl, S., Benard, C., Genty, B., & Wingate, L. (2017). Bryophyte gas-exchange dynamics along varying hydration status reveal a significant carbonyl sulphide (COS) sink in the dark and COS source in the light. *New Phytologist*, *215*(3), 965–976. https://doi.org/10.1111/nph.14584

Gimeno, T. E., Saavedra, N., Ogée, J., Medlyn, B. E., & Wingate, L. (2019). A novel optimization approach incorporating non-stomatal limitations predicts stomatal behaviour in species from six plant functional types. *Journal of Experimental Botany*, *70*(5), 1639–1651. https://doi.org/10.1093/jxb/erz020

Glatthor, N., Höpfner, M., Baker, I. T., Berry, J., Campbell, J. E., Kawa, S. R., Krysztofiak, G., Leyser, A., Sinnhuber, B. M., Stiller, G. P., Stinecipher, J., & von Clarmann, T. (2015). Tropical sources and sinks of carbonyl sulfide observed from space. *Geophysical Research Letters*, *42*(22), 10082–10090. https://doi.org/10.1002/2015GL066293

Glatthor, N., Höpfner, M., Leyser, A., Stiller, G. P., von Clarmann, T., Grabowski, U., Kellmann, S., Linden, A., Sinnhuber, B. M., Krysztofiak, G., & Walker, K. A. (2017). Global carbonyl sulfide (OCS) measured by MIPAS/Envisat during 2002-2012. *Atmospheric Chemistry and Physics*, *17*(4), 2631–2652. https://doi.org/10.5194/acp-17-2631-2017

Goldberg, D. E., & Holland, J. H. (1988). Genetic Algorithms and Machine Learning. *Machine Learning*, *3*(2), 95–99. https://doi.org/10.1023/A:1022602019183

Goward, S. N., & Huemmrich, K. E. (1992). Vegetation Canopy PAR Absorptance and the Normalized Difference Vegetation Index: An Assessment Using the SAIL Model (Vol. 39).

Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T. W., Sperry, J. S., & McDowell, N. G. (2020). Plant responses to rising vapor pressure deficit. In *New Phytologist* (Vol. 226, Issue 6, pp. 1550–1566). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.16485

Grossiord, C., Sevanto, S., Borrego, I., Chan, A. M., Collins, A. D., Dickman, L. T., Hudson, P. J., McBranch, N., Michaletz, S. T., Pockman, W. T., Ryan, M., Vilagrosa, A., & McDowell, N. G. (2017). Tree water dynamics in a drying and warming world. *Plant Cell and Environment*, *40*(9), 1861–1873. https://doi.org/10.1111/pce.12991

Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T. A., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J., & Zhang, Y. (2021). The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission. *Earth System Science Data*, *13*(11), 5423–5440. https://doi.org/10.5194/essd-13-5423-2021

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J. E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., & Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(14). https://doi.org/10.1073/pnas.1320008111

Hamed Alemohammad, S., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., Miralles, D., Prigent, C., & Gentine, P. (2017). Water, Energy, and Carbon with Artificial Neural Networks (WECANN): A statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. *Biogeosciences*, *14*(18), 4101–4124. https://doi.org/10.5194/bg-14-4101-2017

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes*, *10*, 4–10. https://doi.org/10.1016/j.wace.2015.08.001

Haupt, R. L., & Haupt, S. E. (2004). Practical genetic algorithms. *John Wiley & Sons.*

Hochberg, U., Rockwell, F. E., Holbrook, N. M., & Cochard, H. (2018). Iso/Anisohydry: A Plant-Environment Interaction Rather Than a Simple Hydraulic Trait. *Trends in Plant Science, 23*(2), 112–120. https://doi.org/10.1016/j.tplants.2017.11.002

Houborg, R., Fisher, J. B., & Skidmore, A. K. (2015). Advances in remote sensing of vegetation function and traits. *International Journal of Applied Earth Observation and Geoinformation*, *43*, 1–6. https://doi.org/10.1016/j.jag.2015.06.001

Huete, A. R., Liu, H. Q., Batchily, K, & van Leeuwen, W. (1997). A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. In *REMOTE SENS. ENVIRON* (Vol. 5). Elsevier Science Inc.

Ishida, A., Uemura, A., Koike, N., Matsumoto, Y., & Hoe, A. L. (1999). Interactive effects of leaf age and self-shading on leaf structure, photosynthetic capacity and chlorophyll fluorescence in the rain forest tree, Dryobalanops aromatica. *Tree Physiology*, *19*(11), 741–747. https://doi.org/10.1093/treephys/19.11.741

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature*, *496*(7445), 347–350. https://doi.org/10.1038/nature11983

Jeong, S. J., Ho, C. H., Gim, H. J., & Brown, M. E. (2011). Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology*, *17*(7), 2385–2399. https://doi.org/10.1111/j.1365-2486.2011.02397.x

Joiner, J., Yoshida, Y., Köhler, P., Campbell, P., Frankenberg, C., van der Tol, C., Yang, P., Parazoo, N., Guanter, L., & Sun, Y. (2020). Systematic orbital geometry-dependent variations in satellite solar-induced fluorescence (SIF) retrievals. *Remote Sensing*, *12*(15). https://doi.org/10.3390/RS12152346

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

Joshi, J., Stocker, B. D., Hofhansl, F., Zhou, S., Dieckmann, U., & Prentice, I. C. (2022). Towards a unified theory of plant photosynthesis and hydraulics. *Nature Plants*, *8*(11), 1304–1316. https://doi.org/10.1038/s41477-022-01244-5

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., & Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1). https://doi.org/10.1038/s41597-019-0076-8

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., S Goll, D., Haverd, V., Köhler, P., Ichii, K., K Jain, A., Liu, J., Lombardozzi, D., … Reichstein, M. (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

Kalma, J. D., McVicar, T. R., & McCabe, M. F. (2008). Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. In *Surveys in Geophysics* (Vol. 29, Issues 4–5, pp. 421–469). https://doi.org/10.1007/s10712-008-9037-z

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., Yang, B., & Richardson, A. D. (2014). Net carbon uptake has increased

through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604. https://doi.org/10.1038/nclimate2253

Kettle, A. J., Kuhn, U., von Hobe, M., Kesselmeier, J., & Andreae, M. O. (2002). Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks. *Journal of Geophysical Research Atmospheres*, *107*(22). https://doi.org/10.1029/2002JD002187

Kitz, F., Gerdel, K., Hammerle, A., Laterza, T., Spielmann, F. M., & Wohlfahrt, G. (2017). In situ soil COS exchange of a temperate mountain grassland under simulated drought. *Oecologia*, *183*(3), 851–860. https://doi.org/10.1007/s00442-016-3805-0

Kitz, F., Spielmann, F. M., Hammerle, A., Kolle, O., Migliavacca, M., Moreno, G., Ibrom, A., Krasnov, D., Noe, S. M., & Wohlfahrt, G. (2020). Soil COS Exchange: A Comparison of Three European Ecosystems. *Global Biogeochemical Cycles*, *34*(4). https://doi.org/10.1029/2019GB006202

Kitz, F., Wachter, H., Spielmann, F., Hammerle, A., & Wohlfahrt, G. (2024). Root and rhizosphere contribution to the net soil COS exchange. *Plant and Soil*, *498*(1–2), 325–339. https://doi.org/10.1007/s11104-023-06438-0

Kohn, M. J. (2010). Carbon isotope compositions of terrestrial C3 plants as indicators of (paleo)ecology and (paleo)climate. *107*. https://doi.org/10.1073/pnas.1004933107/-/DCSupplemental

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

Konings, A. G., & Gentine, P. (2017). Global variations in ecosystem-scale isohydricity. *Global Change Biology*, *23*(2), 891–905. https://doi.org/10.1111/gcb.13389

Konings, A. G., Holtzman, N. M., Rao, K., Xu, L., & Saatchi, S. S. (2021). Interannual Variations of Vegetation Optical Depth are Due to Both Water Stress and Biomass Changes. *Geophysical Research Letters*, *48*(16). https://doi.org/10.1029/2021GL095267

Kooijmans, L. M. J., Sun, W., Aalto, J., Erkkilä, K. M., Maseyk, K., Seibt, U., Vesala, T., Mammarella, I., & Chen, H. (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

Kooijmans, L. M. J., Uitslag, N. A. M., Zahniser, M. S., Nelson, D. D., Montzka, S. A., & Chen, H. (2016). Continuous and high-precision atmospheric concentration measurements of COS, CO_2, CO and H_2O using a quantum cascade laser spectrometer (QCLS). *Atmospheric Measurement Techniques*, *9*(11), 5293–5314. https://doi.org/10.5194/amt-9-5293-2016

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. In *Global Biogeochemical Cycles* (Vol. 19, Issue 1, pp. 1–33). https://doi.org/10.1029/2003GB002199

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., & Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. In *Atmos. Chem. Phys* (Vol. 5). www.atmos-chem-phys.org/acp/5/417/SRef-ID:1680-7324/acp/2005-5-417EuropeanGeosciencesUnion

Kumarathunge, D. P., Medlyn, B. E., Drake, J. E., Tjoelker, M. G., Aspinwall, M. J., Battaglia, M., Cano, F. J., Carter, K. R., Cavaleri, M. A., Cernusak, L. A., Chambers, J. Q., Crous, K. Y., de Kauwe, M. G., Dillaway, D. N., Dreyer, E., Ellsworth, D. S., Ghannoum, O., Han, Q., Hikosaka, K., … Way, D. A. (2019). Acclimation and adaptation components of the temperature dependence of plant photosynthesis at the global scale. *New Phytologist*, *222*(2), 768–784. https://doi.org/10.1111/nph.15668

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., & Richardson, A. D. (2012). Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, *9*(10), 3757–3776. https://doi.org/10.5194/bg-9-3757-2012

Lasslop, G., Reichstein, M., Papale, D., Richardson, A., Arneth, A., Barr, A., Stoy, P., & Wohlfahrt, G. (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

Lavergne, A., Hemming, D., Prentice, I. C., Guerrieri, R., Oliver, R. J., & Graven, H. (2022). Global decadal variability of plant carbon isotope discrimination and its link to gross primary production. *Global Change Biology*, *28*(2), 524–541. https://doi.org/10.1111/gcb.15924

Leakey, A. D. B., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., & Ort, D. R. (2009). Elevated CO_2 effects on plant carbon, nitrogen, and water relations: Six important lessons from FACE. *Journal of Experimental Botany*, *60*(10), 2859–2876. https://doi.org/10.1093/jxb/erp096

Lennartz, S. T., Marandino, C. A., von Hobe, M., Cortes, P., Quack, B., Simo, R., Booge, D., Pozzer, A., Steinhoff, T., Arevalo-Martinez, D. L., Kloss, C., Bracher, A., Röttgers, R., Atlas, E., & Krüger, K. (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

Leuning, R. (1995). A critical appraisal of a combined stomatal-photosynthesis model for C $_3$ plants. *Plant, Cell & Environment*, *18*(4), 339–355. https://doi.org/10.1111/j.1365-3040.1995.tb00370.x

Leuning, R., Zhang, Y. Q., Rajaud, A., Cleugh, H., & Tu, K. (2008). A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation. *Water Resources Research*, *44*(10). https://doi.org/10.1029/2007WR006562

Leverne, L., Abadie, C., Bacour, C., Zhang, Y., Black, T.A., Bastrikov, V., Raoult, N., Peylin, P., Krieger-Liszkay, A., & Maignan, F. (in revision). Improving the GPP of boreal evergreen needleleaf forests estimated by a land surface model through a physiologically-based representation of NPQ and co-assimilation of space-borne SIF and *in situ* GPP

Li, J., Shen, L., Zhang, Y., Liu, Y., Wu, J., & Wang, A. (2024). Carbonyl Sulfide (COS) in Terrestrial Ecosystem: What We Know and What We Do Not. *Atmosphere*, *15*(7), 778. https://doi.org/10.3390/atmos15070778

Li, W., Pacheco-Labrador, J., Migliavacca, M., Miralles, D., Hoek van Dijke, A., Reichstein, M., Forkel, M., Zhang, W., Frankenberg, C., Panwar, A., Zhang, Q., Weber, U., Gentine, P., & Orth, R. (2023). Widespread and complex drought effects on vegetation physiology inferred from space. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-40226-9

Li, X., Piao, S., Huntingford, C., Peñuelas, J., Yang, H., Xu, H., Chen, A., Friedlingstein, P., Keenan, T. F., Sitch, S., Wang, X., Zscheischler, J., & Mahecha, M. D. (2023). Global variations in critical drought thresholds that impact vegetation. *National Science Review*, *10*(5). https://doi.org/10.1093/nsr/nwad049

Li, X., & Xiao, J. (2019). A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data. *Remote Sensing*, *11*(5). https://doi.org/10.3390/rs11050517

Liao, Z., Zhou, B., Zhu, J., Jia, H., & Fei, X. (2023). A critical review of methods, principles and progress for estimating the gross primary productivity of terrestrial ecosystems. In *Frontiers in Environmental Science* (Vol. 11). Frontiers Media S.A. https://doi.org/10.3389/fenvs.2023.1093095

Liu, J., Jiang, C., Wu, H., Guo, L., Zhang, H., & Zhao, Y. (2023). Controls on leaf water hydrogen and oxygen isotopes: A local investigation across seasons and altitude. *Hydrology and Earth System Sciences*, *27*(2), 599–612. https://doi.org/10.5194/hess-27-599-2023

Liu, Y., Holtzman, N. M., & Konings, A. G. (2021). Global ecosystem-scale plant hydraulic traits retrieved using model-data fusion. *Hydrology and Earth System Sciences*, *25*(5), 2399–2417. https://doi.org/10.5194/hess-25-2399-2021

Lloyd, J., Farquhar, G. D., Lloyd, J., & Farquhar, G. D. (1994). ^{13}C discrimination during CO_2 assimilation by the terrestrial biosphere. In *Oecologia* (Vol. 99).

Lu, X., Liu, Z., An, S., Miralles, D. G., Maes, W., Liu, Y., & Tang, J. (2018). Potential of solar-induced chlorophyll fluorescence to estimate transpiration in a temperate forest. *Agricultural and Forest Meteorology*, *252*, 75–87. https://doi.org/10.1016/j.agrformet.2018.01.017

Ma, J., Kooijmans, L. M. J., Cho, A., Montzka, S. A., Glatthor, N., Worden, J. R., Kuai, L., Atlas, E. L., & Krol, M. C. (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., Maignan, F., Ottlé, C., Peaucelle, M., Santaren, D., & Peylin, P. (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. In *Global Biogeochemical Cycles* (Vol. 36, Issue 7). John Wiley and Sons Inc. https://doi.org/10.1029/2021GB007177

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

Maes, W. H., Pagán, B. R., Martens, B., Gentine, P., Guanter, L., Steppe, K., Verhoest, N. E. C., Dorigo, W., Li, X., Xiao, J., & Miralles, D. G. (2020). Sun-induced fluorescence closely linked to ecosystem transpiration as evidenced by satellite data and radiative transfer models. *Remote Sensing of Environment*, *249*. https://doi.org/10.1016/j.rse.2020.112030

Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., Krishnan, P., Bastrikov, V., & MacBean, N. (2021). Optimizing Carbon Cycle Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Mean Net CO_2 Flux and its Inter-Annual Variability. *Journal of Geophysical Research: Biogeosciences*, *126*(10). https://doi.org/10.1029/2021JG006400

Maignan, F., Bacour, C., Abadie, C., Leverne, L., Zhang, Y., & Ciais, P. (in prep.). Improving the radiative transfer model for solar-induced chlorophyll fluorescence in a land surface model

Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K. M., Commane, R., Wehr, R., Elliott Campbell, J., Belviso, S., Montzka, S. A., Raoult, N., Seibt, U., Shiga, Y. P., Vuichard, N., Whelan, M. E., & Peylin, P. (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

Malhi, Y., Franklin, J., Seddon, N., Solan, M., Turner, M. G., Field, C. B., & Knowlton, N. (2020). Climate change and ecosystems: Threats, opportunities and solutions. In *Philosophical Transactions of the*

Royal Society B: Biological Sciences (Vol. 375, Issue 1794). Royal Society Publishing. https://doi.org/10.1098/rstb.2019.0104

Marchin, R. M., Broadhead, A. A., Bostic, L. E., Dunn, R. R., & Hoffmann, W. A. (2016). Stomatal acclimation to vapour pressure deficit doubles transpiration of small tree seedlings with warming. *Plant Cell and Environment*, *39*(10), 2221–2234. https://doi.org/10.1111/pce.12790

Martínez-Vilalta, J., Poyatos, R., Aguadé, D., Retana, J., & Mencuccini, M. (2014). A new look at water transport regulation in plants. *New Phytologist*, *204*(1), 105–115. https://doi.org/10.1111/nph.12912

Martin-StPaul, N., Delzon, S., & Cochard, H. (2017). Plant resistance to drought depends on timely stomatal closure. In *Ecology Letters* (Vol. 20, Issue 11, pp. 1437–1447). Blackwell Publishing Ltd. https://doi.org/10.1111/ele.12851

Masaki, Y., Iizuka, R., Kato, H., Kojima, Y., Ogawa, T., Yoshida, M., Matsushita, Y., & Katayama, Y. (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/jsme2.ME20058

Maseyk, K., Berry, J. A., Billesbach, D., Campbell, J. E., Torn, M. S., Zahniser, M., & Seibt, U. (2014). Sources and sinks of carbonyl sulfide in an agricultural field in the Southern Great Plains. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(25), 9064–9069. https://doi.org/10.1073/pnas.1319132111

McCarroll, D., Gagen, M. H., Loader, N. J., Robertson, I., Anchukaitis, K. J., Los, S., Young, G. H. F., Jalkanen, R., Kirchhefer, A., & Waterhouse, J. S. (2009). Correction of tree ring stable carbon isotope chronologies for changes in the carbon dioxide content of the atmosphere. *Geochimica et Cosmochimica Acta*, *73*(6), 1539–1547. https://doi.org/10.1016/j.gca.2008.11.041

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., & Yepez, E. A. (2008). Mechanisms of plant survival and mortality during drought: Why do some plants survive while others succumb to drought? In *New Phytologist* (Vol. 178, Issue 4, pp. 719–739). https://doi.org/10.1111/j.1469-8137.2008.02436.x

Meinzer, F. C., Woodruff, D. R., Marias, D. E., Smith, D. D., McCulloh, K. A., Howard, A. R., & Magedman, A. L. (2016). Mapping 'hydroscapes' along the iso- to anisohydric continuum of stomatal regulation of plant water status. In *Ecology Letters* (Vol. 19, Issue 11, pp. 1343–1352). Blackwell Publishing Ltd. https://doi.org/10.1111/ele.12670

Meridja, S. (2011). Approche biophysique des processus de développement et de croissance des couverts végétaux : Interaction avec le stress hydrique et optimisation des pratiques culturales en climat méditerranéen. *Sciences agricoles. AgroParisTech*

Miao, L., Li, S., Zhang, F., Chen, T., Shan, Y., & Zhang, Y. (2020). Future Drought in the Dry Lands of Asia Under the 1.5 and 2.0°C Warming Scenarios. *Earth's Future*, *8*(6). https://doi.org/10.1029/2019EF001337

Millennium Ecosystem Assessment (Program). (2005). *Ecosystems and human well-being : synthesis*. Island Press.

Miralles, D. G., Holmes, T. R. H., de Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

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 CO_2. *Journal of Geophysical Research Atmospheres*, *112*(9). https://doi.org/10.1029/2006JD007665

Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. In *TECHNOMETRICS* (Vol. 33, Issue 2).

Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Mahmoud Hamdi, Z., … Jung, M. (preprint). X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X. *EGUsphere* [preprint]. https://doi.org/10.5194/egusphere-2024-165

Nishida, K., Nemani, R. R., Running, S. W., & Glassy, J. M. (2003). An operational remote sensing algorithm of land surface evaporation. *Journal of Geophysical Research: Atmospheres*, *108*(9). https://doi.org/10.1029/2002jd002062

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., & Phillips, R. P. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Climate Change*, *6*(11), 1023–1027. https://doi.org/10.1038/nclimate3114

Ogée, 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

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Science (New York, N.Y.)*, *313*(5790), 1068–1072. https://doi.org/10.1126/science.1128845

Park, R., & Epstein, S. (1960). Carbon isotope fractionation during photosynthesis. *Geochimica et Cosmochimica Acta*, *21*(1–2), 110–126. https://doi.org/10.1016/S0016-7037(60)80006-3

Panek, J. A., & Waring, R. H. (1997). Stable carbon isotopes as indicators of limitations to forest growth imposed by climate stress. *Ecological Applications*, *7*(3), 854–863. https://doi.org/https://doi.org/10.1890/1051-0761(1997)007[0854:SCIAIO]2.0.CO;2

Peylin, P., Ciais, P., Denning, A. S., Tans, P. P., Berry, J. A., & White, J. W. C. (1999). A 3-dimensional study of $\delta^{18}O$ in atmospheric CO_2: contribution of different land ecosystems. *Tellus B: Chemical and Physical Meteorology*. https://doi.org/10.3402/tellusb.v51i3.16452

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R. L., Groot Zwaaftink, C. D., Evangeliou, N., Sodemann, H., Haimberger, L., Henne, S., Brunner, D., Burkhart, J. F., Fouilloux, A., Brioude, J., Philipp, A., … Stohl, A. (2019). The Lagrangian particle dispersion model FLEXPART version 10.4. *Geoscientific Model Development*, *12*(12), 4955–4997. https://doi.org/10.5194/gmd-12-4955-2019

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., & Berry, J. A. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. In *Journal of Experimental Botany* (Vol. 65, Issue 15, pp. 4065–4095). Oxford University Press. https://doi.org/10.1093/jxb/eru191

Poyatos, R., Granda, V., Flo, V., Adams, M. A., Adorján, B., Aguadé, D., Aidar, M. P. M., Allen, S., Alvarado-Barrientos, M. S., Anderson-Teixeira, K. J., Aparecido, L. M., Altaf Arain, M., Aranda, I., Asbjornsen, H., Baxter, R., Beamesderfer, E., Berry, Z. C., Berveiller, D., Blakely, B., … Martínez-Vilalta, J. (2021). Global transpiration data from sap flow measurements: The SAPFLUXNET database. *Earth System Science Data*, *13*(6), 2607–2649. https://doi.org/10.5194/essd-13-2607-2021

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., & Wright, I. J. (2014). Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology. *Ecology Letters*, *17*(1), 82–91. https://doi.org/10.1111/ele.12211

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

Rebel, K. T., de Jeu, R. A. M., Ciais, P., Viovy, N., Piao, S. L., Kiely, G., & Dolman, A. J. (2012). A global analysis of soil moisture derived from satellite observations and a land surface model. *Hydrology and Earth System Sciences*, *16*(3), 833–847. https://doi.org/10.5194/hess-16-833-2012

Reich, P. B., Hobbie, S. E., & Lee, T. D. (2014). Plant growth enhancement by elevated CO_2 eliminated by joint water and nitrogen limitation. *Nature Geoscience*, *7*(12), 920–924. https://doi.org/10.1038/ngeo2284

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D.,

Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., … Valentini, R. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. In *Global Change Biology* (Vol. 11, Issue 9, pp. 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., Abadie, C., Bacour, C., Lennartz, S., & Peylin, P. (2022). Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling. *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., Niwa, Y., Rodenbeck, C., Belviso, S., Kooijmans, L., Lennartz, S., Maignan, F., Chevallier, F., Chipperfield, M. P., Pope, R. J., Harrison, J. J., Vimont, I., Wilson, C., & Peylin, P. (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). 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

Running, S. W. (1976). Environmental control of leaf water conductance in conifers. *Canadian Journal of Forest Research*, *6*(1), 104–112. https://doi.org/10.1139/x76-013

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. In *BioScience* (Vol. 54, Issue 6). https://academic.oup.com/bioscience/article/54/6/547/294347

Ryu, Y., Baldocchi, D. D., Kobayashi, H., van Ingen, C., Li, J., Black, T. A., Beringer, J., van Gorsel, E., Knohl, A., Law, B. E., & Roupsard, O. (2011). Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales. *Global Biogeochemical Cycles*, *25*(4). https://doi.org/10.1029/2011GB004053

Sabot, M. E. B., de Kauwe, M. G., Pitman, A. J., Medlyn, B. E., Ellsworth, D. S., Martin-StPaul, N. K., Wu, J., Choat, B., Limousin, J., Mitchell, P. J., Rogers, A., & Serbin, S. P. (2022). One Stomatal Model to Rule Them All? Toward Improved Representation of Carbon and Water Exchange in Global Models. *Journal of Advances in Modeling Earth Systems*, *14*(4). https://doi.org/10.1029/2021MS002761

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 (CO_2). In *Biogeosciences* (Vol. 2).

www.biogeosciences.net/bg/2/125/SRef-ID:1726-4189/bg/2005-2-125EuropeanGeosciencesUnion

Santaren, D., Peylin, P., Bacour, C., Ciais, P., & Longdoz, B. (2014). Ecosystem model optimization using in situ flux observations: Benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances. *Biogeosciences*, *11*(24), 7137–7158. https://doi.org/10.5194/bg-11-7137-2014

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. In *Biogeosciences* (Vol. 7). www.biogeosciences.net/7/333/2010/

Sellers, P. J., Bounoua, L., Collatz, G. J., Randall, D. A., Dazlich, D. A., Los, S. O., Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., & Jensen, T. G. (1996). Comparison of Radiative and Physiological Effects of Doubled Atmospheric CO_2 on Climate. *Science*, *271*(5254), 1402–1406. https://doi.org/10.1126/science.271.5254.1402

Sharpe, P., Wu, H., & Spence, R. (1987). Stomatal mechanics. In *Zeiger E, Farquhar GD, Cowan IR* (Stomatal function).

Sheshshayee, M. S., Bindumadhava, H., Ramesh, R., Prasad, T. G., Lakshminarayana, M. R., & Udayakumar, M. (2005). Oxygen isotope enrichment (Δ18O) as a measure of time-averaged transpiration rate. *Journal of Experimental Botany*, *56*(422), 3033–3039. https://doi.org/10.1093/jxb/eri300

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*(2), 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Snyder, K. A., & Tartowski, S. L. (2006). Multi-scale temporal variation in water availability: Implications for vegetation dynamics in arid and semi-arid ecosystems. *Journal of Arid Environments*, *65*(2), 219–234. https://doi.org/10.1016/j.jaridenv.2005.06.023

Sobol, I. M. (1993). Sensitivity Estimates for Nonlinear Mathematical Models. *Mathematical Modelling and Computational Experiments*, *4*, 407–414.

Sperry, J. S., Venturas, M. D., Anderegg, W. R. L., Mencuccini, M., Mackay, D. S., Wang, Y., & Love, D. M. (2017). Predicting stomatal responses to the environment from the optimization of photosynthetic gain and hydraulic cost. *Plant Cell and Environment*, *40*(6), 816–830. https://doi.org/10.1111/pce.12852

Spielmann, F. M., Hammerle, A., Kitz, F., Gerdel, K., & Wohlfahrt, G. (2020). Seasonal dynamics of the COS and CO_2 exchange of a managed temperate grassland. *Biogeosciences*, *17*(16), 4281–4295. https://doi.org/10.5194/bg-17-4281-2020

Spielmann, F. M., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., Ibrom, A., El-Madany, T. S., Gerdel, K., Moreno, G., Kolle, O., Karl, T., Peressotti, A., & Delle Vedove, G. (2019). Gross Primary Productivity of Four European Ecosystems Constrained by Joint CO_2 and COS Flux Measurements. *Geophysical Research Letters*, *46*(10), 5284–5293. https://doi.org/10.1029/2019GL082006

Stimler, K., Berry, J. A., Montzka, S. A., & Yakir, D. (2011). Association between carbonyl sulfide uptake and 18Δ during gas exchange in C_3 and C_4 leaves. *Plant Physiology*, *157*(1), 509–517. https://doi.org/10.1104/pp.111.176578

Stimler, K., Berry, J. A., & Yakir, D. (2012). Effects of carbonyl sulfide and carbonic anhydrase on stomatal conductance. *Plant Physiology*, *158*(1), 524–530. https://doi.org/10.1104/pp.111.185926

Stimler, K., Montzka, S. A., Berry, J. A., Rudich, Y., & Yakir, D. (2010). Relationships between carbonyl sulfide (COS) and CO_2 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., Simpson, I. J., & Campbell, J. E. (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

Stirbet, A., Lazár, D., Guo, Y., & Govindjee, G. (2020). Photosynthesis: Basics, history and modelling. In *Annals of Botany* (Vol. 126, Issue 4, pp. 511–537). Oxford University Press. https://doi.org/10.1093/aob/mcz171

Stovall, A. E. L., Shugart, H., & Yang, X. (2019). Tree height explains mortality risk during an intense drought. *Nature Communications*, *10*(1), 4385. https://doi.org/10.1038/s41467-019-12380-6

Sun, W., Maseyk, K., Lett, C., & Seibt, U. (2015). A soil diffusion-reaction model for surface COS flux: COSSM v1. *Geoscientific Model Development*, *8*(10), 3055–3070. https://doi.org/10.5194/gmd-8-3055-2015

Sun, W., Maseyk, K., Lett, C., & Seibt, U. (2018). Stomatal control of leaf fluxes of carbonyl sulfide and CO_2 in a Typha freshwater marsh. *Biogeosciences*, *15*(11), 3277–3291. https://doi.org/10.5194/bg-15-3277-2018

Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., Drewry, D. T., Verma, M., Porcar-Castell, A., Griffis, T. J., Gu, L., Magney, T. S., Köhler, P., Evans, B., & Yuen, K. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, *358*(6360). https://doi.org/10.1126/science.aam5747

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). https://doi.org/10.1029/2008GL034332

Swann, A. L. S., Hoffman, F. M., Koven, C. D., & Randerson, J. T. (2016). Plant responses to increasing CO_2 reduce estimates of climate impacts on drought severity. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(36), 10019–10024. https://doi.org/10.1073/pnas.1604581113

Lennartz, S., A Marandino, C., von Hobe, M., O Andreae, M., Aranami, K., Atlas, E., Berkelhammer, M., Bingemer, H., Booge, D., Cutter, G., Cortes, P., Kremser, S., S Law, C., Marriner, A., Simó, R., Quack, B., Xie, H., & Xu, X. (2020). Marine carbonyl sulfide (OCS) and carbon disulfide (CS_2): 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

Tang, R., Li, Z. L., & Tang, B. (2010). An application of the Ts-VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and validation. *Remote Sensing of Environment*, *114*(3), 540–551. https://doi.org/10.1016/j.rse.2009.10.012

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

Tei, S., Sugimoto, A., Kotani, A., Ohta, T., Morozumi, T., Saito, S., Hashiguchi, S., & Maximov, T. (2019). Strong and stable relationships between tree-ring parameters and forest-level carbon fluxes in a Siberian larch forest. *Polar Science*, *21*, 146–157. https://doi.org/10.1016/j.polar.2019.02.001

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., & Papale, D. (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

Tucker, C. J., & Sellers, P. J. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, *7*(11), 1395–1416. https://doi.org/10.1080/01431168608948944

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Zhao, M., Running, S. W., Wofsy, S. C., Urbanski, S., Dunn, A. L., & Munger, J. W. (2003). Scaling Gross Primary Production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation. *Remote Sensing of Environment*, *88*(3), 256–270. https://doi.org/10.1016/j.rse.2003.06.005

Uribe, M. del R., Coe, M. T., Castanho, A. D. A., Macedo, M. N., Valle, D., & Brando, P. M. (2023). Net loss of biomass predicted for tropical biomes in a changing climate. *Nature Climate Change*, *13*(3), 274–281. https://doi.org/10.1038/s41558-023-01600-z

van der Tol, C., Vilfan, N., Dauwe, D., Cendrero-Mateo, M. P., & Yang, P. (2019). The scattering and re-absorption of red and near-infrared chlorophyll fluorescence in the models Fluspect and SCOPE. *Remote Sensing of Environment*, *232*. https://doi.org/10.1016/j.rse.2019.111292

Vesala, T., Kohonen, K. M., Kooijmans, L. M. J., Praplan, A. P., Foltýnová, L., Kolari, P., Kulmala, M., Bäck, J., Nelson, D., Yakir, D., Zahniser, M., & Mammarella, I. (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

Vicente-Serrano, S. M., Miralles, D. G., McDowell, N., Brodribb, T., Domínguez-Castro, F., Leung, R., & Koppa, A. (2022). The uncertain role of rising atmospheric CO_2 on global plant transpiration. In *Earth-Science Reviews* (Vol. 230). Elsevier B.V. https://doi.org/10.1016/j.earscirev.2022.104055

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., & Peylin, P. (2019). Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): multi-scale evaluation of gross primary production. *Geoscientific Model Development*, *12*(11), 4751–4779. https://doi.org/10.5194/gmd-12-4751-2019

Wang, Q., Tenhunen, J., Dinh, N. Q., Reichstein, M., Vesala, T., & Keronen, P. (2004). Similarities in ground- and satellite-based NDVI time series and their relationship to physiological activity of a Scots pine forest in Finland. *Remote Sensing of Environment*, *93*(1–2), 225–237. https://doi.org/10.1016/j.rse.2004.07.006

Wang, S., Zhang, Y., Ju, W., Qiu, B., & Zhang, Z. (2021). Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data. *Science of the Total Environment*, *755*. https://doi.org/10.1016/j.scitotenv.2020.142569

Wang, X., Qiu, B., Li, W., & Zhang, Q. (2019). Impacts of drought and heatwave on the terrestrial ecosystem in China as revealed by satellite solar-induced chlorophyll fluorescence. *Science of the Total Environment*, *693*. https://doi.org/10.1016/j.scitotenv.2019.133627

Wang, Y., Sperry, J. S., Anderegg, W. R. L., Venturas, M. D., & Trugman, A. T. (2020). A theoretical and empirical assessment of stomatal optimization modeling. In *New Phytologist* (Vol. 227, Issue 2, pp. 311–325). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.16572

Wehr, R., Commane, R., Munger, J. W., Barry Mcmanus, J., Nelson, D. D., Zahniser, M. S., Saleska, S. R., & Wofsy, S. C. (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

Wei, S., Yi, C., Fang, W., & Hendrey, G. (2017). A global study of GPP focusing on light-use efficiency in a random forest regression model. *Ecosphere*, *8*(5). https://doi.org/10.1002/ecs2.1724

Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., & Lee, X. (2017). Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophysical Research Letters*, *44*(6), 2792–2801. https://doi.org/10.1002/2016GL072235

Whelan, M. (2022). Traversing scales: parts-per-trillion to gigatons. In *New Phytologist* (Vol. 235, Issue 5, pp. 1686–1688). John Wiley and Sons Inc. https://doi.org/10.1111/nph.18337

Whelan, M. E., Hilton, T. W., Berry, J. A., Berkelhammer, M., Desai, A. R., & Elliott Campbell, J. (2016). Carbonyl sulfide exchange in soils for better estimates of ecosystem carbon uptake. *Atmospheric Chemistry and Physics*, *16*(6), 3711–3726. https://doi.org/10.5194/acp-16-3711-2016

Whelan, M. E., Lennartz, S. T., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L. M. J., Hilton, T. W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K. F., Yakir, D., … Elliott Campbell, J. (2018). Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences*, *15*(12), 3625–3657. https://doi.org/10.5194/bg-15-3625-2018

Whelan, M. E., & Rhew, R. C. (2015). Carbonyl sulfide produced by abiotic thermal and photodegradation of soil organic matter from wheat field substrate. *Journal of Geophysical Research: Biogeosciences*, *120*(1), 54–62. https://doi.org/10.1002/2014JG002661

Whelan, M. E., Shi, M., Sun, W., Vries, L. K., Seibt, U., & Maseyk, K. (2022). Soil Carbonyl Sulfide (OCS) Fluxes in Terrestrial Ecosystems: An Empirical Model. *Journal of Geophysical Research: Biogeosciences*, *127*(9). https://doi.org/10.1029/2022JG006858

White, M. L., Zhou, Y., Russo, R. S., Mao, H., Talbot, R., Varner, R. K., & Sive, B. C. (2010). Atmospheric Chemistry and Physics Carbonyl sulfide exchange in a temperate loblolly pine forest grown under ambient and elevated CO_2. In *Atmos. Chem. Phys* (Vol. 10). www.atmos-chem-phys.net/10/547/2010/

Whitley, R., Taylor, D., Macinnis-Ng, C., Zeppel, M., Yunusa, I., O'Grady, A., Froend, R., Medlyn, B., & Eamus, D. (2013). Developing an empirical model of canopy water flux describing the common response of transpiration to solar radiation and VPD across five contrasting woodlands and forests. *Hydrological Processes*, *27*(8), 1133–1146. https://doi.org/10.1002/hyp.9280

Wohlfahrt, G., Brilli, 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., Hammerle, A., Julitta, T., Spielmann, F. M., & Yakir, D. (2018). Sun-induced fluorescence and gross primary productivity during a heat wave. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32602-z

Wolf, A., Anderegg, W. R. L., & Pacala, S. W. (2016). Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), E7222–E7230. https://doi.org/10.1073/pnas.1615144113

Woodward, F. I., Lomas, M. R., & Betts, R. A. (1998). Vegetation-climate feedbacks in a greenhouse world. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1365), 29–39. https://doi.org/10.1098/rstb.1998.0188

Wu, D., Zhao, X., Liang, S., Zhou, T., Huang, K., Tang, B., & Zhao, W. (2015). Time-lag effects of global vegetation responses to climate change. *Global Change Biology*, *21*(9), 3520–3531. https://doi.org/10.1111/gcb.12945

Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E. A., Zhang, Q., & Moore, B. (2004). Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sensing of Environment*, *89*(4), 519–534. https://doi.org/10.1016/j.rse.2003.11.008

Yamori, W., Hikosaka, K., & Way, D. A. (2014). Temperature response of photosynthesis in C_3, C_4, and CAM plants: Temperature acclimation and temperature adaptation. In *Photosynthesis Research* (Vol. 119, Issues 1–2, pp. 101–117). https://doi.org/10.1007/s11120-013-9874-6

Yang, P., van der Tol, C., Verhoef, W., Damm, A., Schickling, A., Kraska, T., Muller, O., & Rascher, U. (2019). Using reflectance to explain vegetation biochemical and structural effects on sun-induced chlorophyll fluorescence. *Remote Sensing of Environment*, *231*. https://doi.org/10.1016/j.rse.2018.11.039

Yao, Y., Joetzjer, E., Ciais, P., Viovy, N., Cresto Aleina, F., Chave, J., Sack, L., Bartlett, M., Meir, P., Fisher, R., & Luyssaert, S. (2022). Forest fluxes and mortality response to drought: Model description (ORCHIDEE-CAN-NHA r7236) and evaluation at the Caxiuanã drought experiment. *Geoscientific Model Development*, *15*(20), 7809–7833. https://doi.org/10.5194/gmd-15-7809-2022

Yin, X., & Struik, P. C. (2009). C_3 and C_4 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

Zaehle, S., Jones, C. D., Houlton, B., Lamarque, J.-F., & Robertson, E. (2015). Nitrogen Availability Reduces CMIP5 Projections of Twenty-First-Century Land Carbon Uptake. *J. Climate*, *28*, 2494–2511. https://doi.org/10.1175/JCLI-D-13

Zhang, Q., Chen, J. M., Ju, W., Zhang, Y., Li, Z., He, L., Pacheco-Labrador, J., Li, J., Qiu, B., Zhang, X., Qiu, F., Chen, B., Chou, S., Zhang, Z., & Shan, N. (2021). Ground-Based Multiangle Solar-Induced Chlorophyll Fluorescence Observation and Angular Normalization for Assessing Crop Productivity. *Journal of Geophysical Research: Biogeosciences*, *126*(5). https://doi.org/10.1029/2020JG006082

Zhang, Q., Liu, X., Zhou, K., Zhou, Y., Gentine, P., Pan, M., & Katul, G. G. (2024). Solar-induced chlorophyll fluorescence sheds light on global evapotranspiration. *Remote Sensing of Environment*, *305*. https://doi.org/10.1016/j.rse.2024.114061

Zhang, Y., Xiao, X., Guanter, L., Zhou, S., Ciais, P., Joiner, J., Sitch, S., Wu, X., Nabel, J., Dong, J., Kato, E., Jain, A. K., Wiltshire, A., & Stocker, B. D. (2016). Precipitation and carbon-water coupling jointly control the interannual variability of global land gross primary production. *Scientific Reports*, *6*. https://doi.org/10.1038/srep39748

Zhou, L., Tucker, C. J., Kaufmann, R. K., Slayback, D., Shabanov, N. v., & Myneni, R. B. (2001). Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. In *Journal of Geophysical Research Atmospheres* (Vol. 106, Issue D17, pp. 20069–20083). Blackwell Publishing Ltd. https://doi.org/10.1029/2000JD000115

Zhou, Y., Zhang, L., Xiao, J., Chen, S., Kato, T., & Zhou, G. (2014). A Comparison of Satellite-Derived Vegetation Indices for Approximating Gross Primary Productivity of Grasslands. *Rangeland Ecology & Management*, *67*(1), 9–18. https://doi.org/https://doi.org/10.2111/REM-D-13-00059.1

Zhu, H., Wu, M., Jiang, F., Vossbeck, M., Kaminski, T., Xing, X., Wang, J., Ju, W., & Chen, J. M. (2024a). Assimilation of carbonyl sulfide (COS) fluxes within the adjoint-based data assimilation system – Nanjing University Carbon Assimilation System (NUCAS v1.0). *Geoscientific Model Development*, *17*(16), 6337–6363. https://doi.org/10.5194/gmd-17-6337-2024

Zhu, H., Xing, X., Wu, M., Ju, W., & Jiang, F. (2024b). Optimizing the terrestrial ecosystem gross primary productivity using carbonyl sulfide (COS) within a two-leaf modeling framework. *Biogeosciences*, *21*(16), 3735–3760. https://doi.org/10.5194/bg-21-3735-2024

Zhu, P., Zhuang, Q., Ciais, P., Welp, L., Li, W., & Xin, Q. (2017). Elevated atmospheric CO_2 negatively impacts photosynthesis through radiative forcing and physiology-mediated climate feedback. *Geophysical Research Letters*, *44*(4), 1956–1963. https://doi.org/10.1002/2016GL071733

Zhu, R., Hu, T., Zhang, Q., Zeng, X., Zhou, S., Wu, F., Liu, Y., & Wang, Y. (2023). A stomatal optimization model adopting a conservative strategy in response to soil moisture stress. *Journal of Hydrology*, *617*, 128931. https://doi.org/10.1016/j.jhydrol.2022.128931

Zobler, L. (1986). *A World Soil File for Global Climate Modeling*. National Aeronautics and Space Administration, Goddard Space Flight Center, Institute for Space Studies.

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

9 Appendix

9.1 Carbonyl sulfide: comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach

Biogeosciences, 18, 2917–2955, 2021
https://doi.org/10.5194/bg-18-2917-2021

Biogeosciences
Open Access
EGU

Carbonyl sulfide: comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach

Fabienne Maignan[1], Camille Abadie[1], Marine Remaud[1], Linda M. J. Kooijmans[2], Kukka-Maaria Kohonen[3], Róisín Commane[4], Richard Wehr[5], J. Elliott Campbell[6], Sauveur Belviso[1], Stephen A. Montzka[7], Nina Raoult[1], Ulli Seibt[8], Yoichi P. Shiga[9], Nicolas Vuichard[1], Mary E. Whelan[10], and Philippe Peylin[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[2]Meteorology and Air Quality, Wageningen University and Research, Wageningen, the Netherlands
[3]Institute for Atmospheric and Earth System Research (INAR)/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland
[4]Dept. Earth & Environmental Sciences, Lamont-Doherty Earth Observatory of Columbia University, New York, NY 10964, USA
[5]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, USA
[6]Sierra Nevada Research Institute, University of California, Merced, California 95343, USA
[7]NOAA Global Monitoring Laboratory, Boulder, Colorado, USA
[8]Dept of Atmospheric & Oceanic Sciences, University of California Los Angeles, California 90095, USA
[9]Universities Space Research Association, Mountain View, CA, USA
[10]Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901, USA

Correspondence: Fabienne Maignan (fabienne.maignan@lsce.ipsl.fr)

Received: 13 October 2020 – Discussion started: 5 November 2020
Revised: 28 February 2021 – Accepted: 11 March 2021 – Published: 12 May 2021

Abstract. Land surface modellers need measurable proxies to constrain the quantity of carbon dioxide (CO_2) assimilated by continental plants through photosynthesis, known as gross primary production (GPP). Carbonyl sulfide (COS), which is taken up by leaves through their stomates and then hydrolysed by photosynthetic enzymes, is a candidate GPP proxy. A former study with the ORCHIDEE land surface model used a fixed ratio of COS uptake to CO_2 uptake normalised to respective ambient concentrations for each vegetation type (leaf relative uptake, LRU) to compute vegetation COS fluxes from GPP. The LRU approach is known to have limited accuracy since the LRU ratio changes with variables such as photosynthetically active radiation (PAR): while CO_2 uptake slows under low light, COS uptake is not light limited. However, the LRU approach has been popular for COS–GPP proxy studies because of its ease of application and apparent low contribution to uncertainty for regional-scale applications. In this study we refined the COS–GPP relationship and implemented in ORCHIDEE a mechanistic model that describes COS uptake by continental vegetation. We compared the simulated COS fluxes against measured hourly COS fluxes at two sites and studied the model behaviour and links with environmental drivers. We performed simulations at a global scale, and we estimated the global COS uptake by vegetation to be $-756\,\mathrm{Gg\,S\,yr^{-1}}$, in the middle range of former studies (-490 to $-1335\,\mathrm{Gg\,S\,yr^{-1}}$). Based on monthly mean fluxes simulated by the mechanistic approach in ORCHIDEE, we derived new LRU values for the different vegetation types, ranging between 0.92 and 1.72, close to recently published averages for observed values of 1.21 for C_4 and 1.68 for C_3 plants. We transported the COS using the monthly vegetation COS fluxes derived from both the mechanistic and the LRU approaches, and we evaluated the simulated COS concentrations at NOAA sites. Although

Published by Copernicus Publications on behalf of the European Geosciences Union.

the mechanistic approach was more appropriate when comparing to high-temporal-resolution COS flux measurements, both approaches gave similar results when transporting with monthly COS fluxes and evaluating COS concentrations at stations. In our study, uncertainties between these two approaches are of secondary importance compared to the uncertainties in the COS global budget, which are currently a limiting factor to the potential of COS concentrations to constrain GPP simulated by land surface models on the global scale.

1 Introduction

Humanity has to face the urgency of climate change if it hopes to limit adverse future impacts (Allen et al., 2018; IPCC, 2019a, b). In order to make reliable predictions of future climate, scientists have built powerful numerical Earth system models (ESMs), where they continuously integrate gained knowledge on a multitude of climate-related and climate-interacting processes. The carbon cycle is at the heart of the present global warming, caused by anthropogenic CO_2 emissions (Ciais et al., 2013). In the global carbon budget, the land component shows the largest uncertainty (Le Quéré et al., 2018; Bloom et al., 2016). Land surface models (LSMs) struggle to accurately represent the large spatial and temporal variability of the CO_2 gross and net fluxes (Anav et al., 2015). CO_2 is first assimilated through plant photosynthesis, before being respired by the ecosystem. The quantity of assimilated carbon is called gross primary productivity (GPP). All other carbon fluxes and stocks derive from this first gross assimilation flux. To help reduce uncertainties in the estimated GPP, LSMs can benefit from knowledge obtained through local eddy covariance measurements of the net ecosystem–atmosphere CO_2 exchange (Friend et al., 2007; Kuppel et al., 2014).

GPP proxies are also used, such as solar-induced fluorescence (Norton et al., 2019; Bacour et al., 2019), isotopic composition of atmospheric CO_2 ($\delta^{18}O$; Farquhar et al., 1993; Welp et al., 2011; $\delta^{13}C$: Peters et al., 2018), and carbonyl sulfide (COS) atmospheric concentrations (Hilton et al., 2015). Using atmospheric COS measurements as a tracer for terrestrial photosynthesis was first suggested by Sandoval-Soto et al. (2005) and Montzka et al. (2007), and Campbell et al. (2008) provided quantitative evidence using airborne observations of COS and CO_2 concentrations and an atmospheric transport model. COS is an atmospheric trace gas that has a molecular structure very similar to CO_2 and is likewise taken up by plants through stomates. COS is hydrolysed within the leaf, with this reaction being catalysed by the enzyme carbonic anhydrase (CA). This reaction is light-independent (Protoschill-Krebs et al., 1996; Goldan et al., 1998) and, because of the high catalytic efficiency of this enzyme (Ogawa et al., 2013; Ogée et al., 2016; Protoschill-Krebs et al., 1996), COS hydrolysis inside the leaf seems therefore to be limited by COS supply driven by changes in stomatal conductance (Goldan et al., 1988; Sandoval-Soto et al., 2005; Seibt et al., 2010; Stimler et al., 2010). Leaves' uptake of COS and CO_2 is thus very similar, but leaves do not produce COS (Protoschill-Krebs et al., 1996; Notni et al., 2007), whereas they emit CO_2 through respiration. That is why vegetation COS fluxes could be used as a proxy for GPP. It is however to be noted that Gimeno et al. (2017) reported COS emissions by bryophytes during daytime.

The approach generally adopted to constrain GPP with COS relies on the determination of a leaf relative uptake (LRU), which is the ratio of COS to CO_2 uptake normalised by their atmospheric concentrations (Sandoval-Soto et al., 2005):

$$\mathrm{LRU} = \frac{F_{\mathrm{COS}}[\mathrm{CO_2}]_{\mathrm{a}}}{\mathrm{GPP}[\mathrm{COS}]_{\mathrm{a}}}, \quad (1)$$

where F_{COS} is the flux of COS uptake (pmol COS m^{-2} s^{-1}), GPP is the gross flux of CO_2 assimilation (µmol CO_2 m^{-2} s^{-1}), $[\mathrm{COS}]_{\mathrm{a}}$ is the atmospheric COS mixing ratio (pmol COS mol^{-1}, ppt), and $[CO_2]_{\mathrm{a}}$ is the atmospheric CO_2 mixing ratio (µmol CO_2 mol^{-1}, ppm).

LRU can be estimated experimentally and then used as a scaling factor for estimating GPP, if F_{COS}, $[\mathrm{COS}]_{\mathrm{a}}$, and $[\mathrm{CO_2}]_{\mathrm{a}}$ are available. Measurements can be made at the leaf level using branch chambers (Seibt et al., 2010; Kooijmans et al., 2019); LRU can also be estimated at the ecosystem level: eddy-covariance flux towers measure the ecosystem total COS flux (Kohonen et al., 2020), and removing the soil contribution gives access to the vegetation part (Wehr et al., 2017). Soil can absorb and emit COS (Whelan et al., 2016; Kitz et al., 2020), with the magnitude of their flux being generally much lower than that of vegetation fluxes (Berkelhammer et al., 2014; Maseyk et al., 2014; Wehr et al., 2017; Whelan et al., 2018). Epiphytes (lichen, mosses) could also have a significant contribution to the ecosystem COS budget (Kuhn and Kesselmeier, 2000; Rastogi et al., 2018).

However, LRU does not appear constant under some environmental conditions. For example, the fixation of carbon from CO_2 relies on light-dependent reactions, unlike the uptake of COS by the CA enzyme, which is light-independent (Stimler et al., 2011). Because of these different responses of COS and CO_2 uptake in leaves, LRU varies with light conditions and decreases sharply with PAR increase (Stimler et al., 2010; Maseyk et al., 2014; Commane et al., 2015; Wehr et al., 2017; Yang et al., 2018). Consequently, LRU values are smaller at midday or in seasons with high incoming light (Kooijmans et al., 2019). Moreover, COS assimilation continues at night as stomatal conductance to gas transfer does not drop to zero, whereas CO_2 uptake by plants stops, leading to an infinite value of LRU. Note however that stomates mostly close at night, so the COS uptake at night is smaller than the COS uptake during the day. The diel (i.e. 24-hourly) variation in LRU with light may however be only of second-

order importance as GPP is very low at low light, and Yang et al. (2018) found that considering sub-daily variations in LRU when computing daily mean GPP values had no importance. It has also been shown that LRU varies between plant species (Stimler et al., 2011), which is why different LRU values were estimated for different vegetation types (Seibt et al., 2010; Whelan et al., 2018). The variability of LRU with plant type and over a day and season (inferred by changes in light conditions) should therefore be carefully accounted for when COS concentrations or flux measurements are used to estimate GPP at the ecosystem and larger scales. We also have to acknowledge that there are still factors that are not accounted for if discrepancies between GPP and COS-based estimations are larger than their estimated respective uncertainties.

Before being able to use COS observations to constrain the simulated GPP, land surface models (LSMs) first need to have an accurate model to simulate vegetation COS fluxes. In a former study, Launois et al. (2015b) simply defined the COS uptake by vegetation as the CO_2 gross uptake simulated by LSMs, scaled with a constant LRU value for each large vegetation class. The goal of this study is to now simulate the uptake of atmospheric COS by continental vegetation in a more complex and realistic way using a mechanistic approach within an LSM and apply this model to evidence the shortcomings or pertinence of the LRU concept, depending on the studied scales.

i. We used the state-of-the art ORCHIDEE LSM (Krinner et al., 2015), and we implemented in it the vegetation COS uptake model of Berry et al. (2013) to simulate the COS fluxes absorbed at the leaf and canopy levels by the continental vegetation.

ii. We evaluated the simulated COS fluxes against measurements at two forest sites, namely the Harvard Forest, United States (Wehr et al., 2017), and Hyytiälä, Finland (Kooijmans et al., 2019; Kohonen et al., 2020; Sun et al., 2018a). We studied the high-frequency behaviour of the modelled conductances over the season and the dependency of the LRU on the environmental and structural conditions.

iii. We compared the simulated mechanistic COS fluxes at the global scale to former estimates; we studied LRU values estimated from monthly fluxes, which are pertinent for atmospheric studies, and compared them to monthly means of high-frequency LRU values.

iv. The mechanistic and LRU simulated COS fluxes were used with the atmospheric transport model LMDz (Hourdin et al., 2006), to provide atmospheric COS concentrations that were evaluated against measurements at sites of the NOAA network.

2 Models, data, and methodology

2.1 Implementation of plant COS uptake in the ORCHIDEE LSM to simulate COS vegetation fluxes

2.1.1 The ORCHIDEE LSM

ORCHIDEE is an LSM developed mainly at Institut Pierre Simon Laplace (IPSL) that computes the water, carbon, and energy balances at the interface between land surfaces and atmosphere (Krinner et al., 2005). Fast processes including hydrology, photosynthesis, and energy balance are run at a half-hourly time step, while other slower processes such as carbon allocation and mortality are simulated at a daily time step. The sub-grid variability for vegetation is represented using fractions of plant functional types (PFTs), grouping plants with similar morphologies and behaviours growing under similar climatic conditions. Photosynthesis follows the Yin and Struik (2009) approach, bringing improvements to the standard Farquhar et al. (1980) model for C_3 plants, the Collatz et al. (1992) model for C_4 plants, and the Ball et al. (1987) model for the stomatal conductance. A main novelty is the introduction of a mesophyll conductance linking the CO_2 concentration at the carboxylation sites, C_c, to the CO_2 intracellular concentration, C_i. For each PFT, the reference value for the maximum photosynthetic capacity at 25 °C, $V_{max,25}$, is derived from literature survey and observation databases and possibly later calibrated using FLUXNET observations (e.g. Kuppel et al., 2012). To compute the maximum photosynthetic capacity at leaf level, V_{max}, the reference value is multiplied at a daily time step by the relative photosynthetic efficiency of leaves based on the mean leaf age following Ishida et al. (1999) (see Eq. A12 and Fig. A12 in Krinner et al., 2005). Leaves are very efficient when they are young and stay so till they approach their pre-defined leaf lifespan. The temperature dependence of the maximum photosynthetic capacity follows Medlyn et al. (2002) and Kattge and Knorr (2007). A water stress function varying between 0 and 1 depending on soil moisture and root profile (de Rosnay and Polcher, 1998) is applied on maximum photosynthetic capacity and conductances. The canopy is discretised in several layers of growing thickness, the number depending on the actual leaf area index (LAI). All the incoming light is considered to be diffuse, and no distinction is made between sun and shaded leaves. The light is attenuated through the canopy following a simple Beer–Lambert absorption law. The CO_2 assimilation, the stomatal conductance, and the intercellular CO_2 concentration C_i are computed per LAI layer, provided LAI is higher than 0.01 and the monthly mean air temperature is higher than −4 °C. The CO_2 assimilation and the stomatal conductance are further summed up over all layers to compute GPP and the total conductance at canopy level. The scaling to the grid cell is made using means weighted by the plant functional type fractions. Phenology is

fully prognostic with PFT-specific phenological models as described in Botta et al. (2000) and MacBean et al. (2015). ORCHIDEE can be run from the site scale to the global scale, coupled with an atmospheric general circulation model, or in off-line mode forced by meteorological fields. In this study, we prescribed the vegetation distribution for site simulations and used yearly PFT maps derived from the ESA Climate Change Initiative (CCI) land cover products for global simulations (Poulter et al., 2015). The soil type is derived from the Zobler map (Zobler, 1986). To account for the CO_2 fertilisation effect, we considered global means of $[CO_2]_a$ with yearly varying values, as provided by the TRENDY model inter-comparison project (Sitch et al., 2015). The impact of not taking into account the spatial and temporal variations in $[CO_2]_a$ on GPP has been studied in Lee et al. (2020); while this simplification has indeed no impact at a global yearly scale for GPP, this may be less true at site and seasonal scales. We used the recent ORCHIDEE version fine-tuned for the Climate Model Intercomparison Project (CMIP) 6 exercise (Boucher et al., 2020), forced by micro-meteorology fields at FLUXNET sites or by 2° CRUNCEP reanalyses at the global scale (https://rda.ucar.edu/datasets/ds314.3/, last access: 19 April 2021).

2.1.2 The Berry model for plant COS uptake

In the ORCHIDEE LSM we implemented the mechanistic model of plant COS uptake based on Berry et al. (2013). In this model, COS follows a diffusive law from the atmosphere to the leaf interior, where it is consumed by CA in the chloroplasts. The uptake from the atmosphere is assumed to be unidirectional, reflecting the fact that COS is generally not produced by plants. The model distinguishes three conductances along the COS path between the atmosphere and the leaf interior: (1) the boundary layer conductance (g_{B_COS}) to gas transfer between the leaf surface and the atmosphere, (2) the stomatal conductance (g_{S_COS}), and (3) the internal conductance (g_{I_COS}). Internal conductance combines the mesophyll conductance and the CA activity into a single equivalent conductance.

The stomatal and boundary layer conductances are associated with factors describing diffusion of COS relative to that of water vapour (1.94 and 1.56, respectively; Stimler et al., 2010). In the chloroplast, the COS hydrolysis is catalysed by the enzyme CA, following first-order kinetics. COS uptake depends on the amount of CA and its relative location to intercellular air spaces, which brings in the mesophyll conductance. These two factors have been shown to scale with the maximum reaction rate of the Rubisco enzyme, V_{max} ($\mu mol\,m^{-2}\,s^{-1}$) (Badger and Price, 1994; Evans et al., 1994). The mesophyll conductance and the first-rate constant are then regrouped into a single equivalent internal conductance, proportional to V_{max}:

$$g_{I\_COS} = \alpha \times V_{max}. \quad (2)$$

The parameter α takes two values depending on the plant photosynthetic pathway (C_3 or C_4). These values were determined experimentally by Berry et al. (2013), who estimated an $\alpha = 0.0012$ for C_3 and an $\alpha = 0.013$ for C_4 species. We thus have the final equation:

$$F_{COS} = [COS]_a \times g_{T\_COS} = [COS]_a \times \left[\frac{1.0}{g_{B\_COS}} + \frac{1.0}{g_{S\_COS}} + \frac{1.0}{g_{I\_COS}}\right]^{-1} = [COS]_a \times \left[\frac{1.56}{g_{B\_W}} + \frac{1.94}{g_{S\_W}} + \frac{1.0}{g_{I\_COS}}\right]^{-1}, \quad (3)$$

where F_{COS} is the flux of COS uptake ($pmol\,COS\,m^{-2}\,s^{-1}$); $[COS]_a$ is the background atmospheric COS mixing ratio considered here to be a constant (500 ppt); g_{T_COS}, g_{B_COS}, g_{S_COS}, and g_{I_COS} are respectively the total, boundary layer, stomatal, and internal conductances to COS ($mol\,COS\,m^{-2}\,s^{-1}$); and g_{B_W} and g_{S_W} are respectively the boundary layer and stomatal conductances to water vapour ($mol\,H_2O\,m^{-2}\,s^{-1}$). Note that in this work $[COS]_a$ is held constant when computing the COS fluxes, contrary to Berry et al. (2013) and Campbell et al. (2017), where $[COS]_a$ is dynamic and taken from the previous time step's PCTM (Parameterized Chemical Transport Model) value. The uncertainty introduced by this simplification is evaluated in the Discussion section. The vegetation COS flux and related conductances are computed for each LAI layer and then summed up to get total values at the canopy level. Unless specified otherwise, fluxes, conductances and LRU are further presented and discussed at the canopy level.

2.1.3 Minimal conductances

As plant CO_2 uptake only occurs under certain conditions such as with sufficient light, temperature, and water, CO_2 assimilation is not calculated in ORCHIDEE when these conditions are not fulfilled. Therefore, the stomatal conductance to CO_2 that is needed to obtain the stomatal conductance to COS is not always computed in ORCHIDEE. However, some studies have shown incomplete stomatal closure at night (Dawson et al., 2007; Lombardozzi et al., 2017; Kooijmans et al., 2019), leading to nighttime COS plant uptake (Berry et al., 2013; Kooijmans et al., 2017). Therefore, we had to define a minimal stomatal conductance to COS under these particular conditions when there is no CO_2 assimilation. The minimal conductance to CO_2 used in ORCHIDEE is based on the residual stomatal conductance if the irradiance approaches zero, represented as the g_0 offset in the stomatal conductance models (see Eq. 15 for C_3 and Eq. 25 for C_4 plants in Yin and Struik, 2009). In the absence of water stress, g_0 takes a constant value for C_3 ($0.00625\,mol\,CO_2\,m^{-2}\,s^{-1}$)

and C_4 ($0.01875\,\mathrm{mol\,CO_2\,m^{-2}\,s^{-1}}$) plants. This constant is multiplied by a water stress function to compute the minimal conductance. This minimal conductance to CO_2 was then applied under conditions when there is no CO_2 assimilation, multiplied by the ratio to convert the conductance to CO_2 into a conductance to COS. We thus model COS assimilation even at night, for all PFTs, and in winter for evergreen species, depending on water stress conditions.

2.1.4 Simulations protocol

All simulations were preceded by a "spin-up" phase to get to an equilibrium state where the considered carbon pools and fluxes are stable with no residual trends in the absence of any disturbances (climate, land use change, CO_2 atmospheric concentrations) (e.g. Wei et al., 2014). A few decades are enough to equilibrate above-ground biomass and GPP. As we transport not only COS, but also CO_2 (see Sect. 2.4 below), we need a longer spin-up where all carbon pools including those in the soil are stable and the net CO_2 fluxes oscillate around zero. Equilibrating the ecosystem photosynthesis with its respiration takes a long time as the slowest soil carbon pool has a residence time on the order of 1000 years. The ORCHIDEE model has a built-in spin-up procedure to accelerate the convergence towards this equilibrium state, using a pseudo-analytical iterative estimation of the targeted carbon pools, based on Lardy et al. (2011). For global simulations, we first performed a 340-year spin-up phase with non-varying pre-industrial atmospheric CO_2 concentration and vegetation map, cycling over the same 10 years of meteorological forcing files, where the final relative variation in the global slowest soil carbon pool was less than 5 %. Starting from this equilibrium state, a transient state simulation was then run applying climate change, land use change, and increasing CO_2 atmospheric concentrations, and COS and GPP fluxes were calculated from 1860 to 2017. We performed site simulations at the Harvard Forest (United States) and Hyytiälä (Finland) FLUXNET sites (see below). For the two sites, we first performed a spin-up simulation cycling over the available years of the FLUXNET forcing files, for around 340 years, using a constant atmospheric CO_2 concentration corresponding to the first year of the FLUXNET forcing file. We then performed the transient simulations over the available FLUXNET years, for each site, with a varying CO_2 atmospheric concentration.

2.2 Evaluation of vegetation COS fluxes at two FLUXNET sites

Vegetation COS fluxes can be measured using branch chambers or estimated using the difference between measurements of ecosystem and soil fluxes. Such measurements were available at the Hyytiälä (Finland) and Harvard Forest (United States) FLUXNET sites. The Hyytiälä site (61.85° N, 24.29° E) is a boreal evergreen needleleaf forest dominated by Scots pine (*Pinus sylvestris*). Branch measurements of COS fluxes were made in a Scots pine tree from March to July 2017 using gas-exchange chambers (Kooijmans et al., 2019); fluxes were derived from mole fraction changes when the chambers were closed once every hour. Measurements were made with an Aerodyne quantum cascade laser spectrometer (QCLS) and were calibrated against reference standards (Kooijmans et al., 2016). Fluxes from empty chambers were regularly measured to be able to correct for gas exchange by the chamber and tubing material (Kooijmans et al., 2019). We also used the Hyytiälä COS ecosystem fluxes (Kohonen et al., 2020); eddy covariance fluxes were measured during the years 2013–2017 at 23 m height, approximately 6 m above the canopy height. Flux data were processed, quality-screened, and gap-filled according to recommendations by Kohonen et al. (2020). Soil fluxes were also available for the year 2015 (Sun et al., 2018a). We thus derived the COS vegetation fluxes at the canopy scale for that year from the difference between ecosystem and soil fluxes. Soil fluxes were generally low compared to plant uptake.

The Harvard Forest site (42.54° N, 72.17° W) is a temperate deciduous broadleaf forest with mainly red oak (*Quercus rubra*), red maple (*Acer rubrum*), and hemlock (*Tsuga canadensis*). Ecosystem COS eddy flux measurements were carried out from a tower from May to October, in 2012 and 2013, using an Aerodyne QCLS and calibrated using gas cylinders. They were further split into vegetation and soil components, using soil chamber CO_2 measurements and a sub-canopy flux-gradient approach (Wehr et al., 2017).

The simulated COS fluxes were evaluated against measurements using the root-mean-square deviation:

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{n=1}^{N}\left(F_{\mathrm{COS}}^{\mathrm{Obs}}(n) - F_{\mathrm{COS}}^{\mathrm{Mod}}(n)\right)^2}{N}}, \quad (4)$$

where N is the number of considered observations, $F_{\mathrm{COS}}^{\mathrm{Obs}}(n)$ is the nth observed COS flux, and $F_{\mathrm{COS}}^{\mathrm{Mod}}(n)$ is the nth modelled COS flux, and the relative RMSD

$$r\mathrm{RMSD} = \frac{\mathrm{RMSD}}{\overline{F_{\mathrm{COS}}^{\mathrm{Obs}}}}, \quad (5)$$

which is the RMSD divided by the mean value of observations.

We also computed the bias, standard deviations, and correlation coefficient.

$$\text{bias} = \overline{F_{\text{COS}}^{\text{Mod}}} - \overline{F_{\text{COS}}^{\text{Obs}}}$$

$$\text{SD}^{\text{Mod}} = \sqrt{\frac{\sum_{n=1}^{N}\left(F_{\text{COS}}^{\text{Mod}}(n) - \overline{F_{\text{COS}}^{\text{Mod}}}\right)^2}{N}}\,\text{SD}^{\text{Obs}}$$

$$= \sqrt{\frac{\sum_{n=1}^{N}\left(F_{\text{COS}}^{\text{Obs}}(n) - \overline{F_{\text{COS}}^{\text{Obs}}}\right)^2}{N}}$$

$$r = \frac{\sum_{n=1}^{N}\left(F_{\text{COS}}^{\text{Obs}}(n) - \overline{F_{\text{COS}}^{\text{Obs}}}\right)\cdot\left(F_{\text{COS}}^{\text{Mod}}(n) - \overline{F_{\text{COS}}^{\text{Mod}}}\right)}{N\cdot\text{SD}^{\text{Obs}}\cdot\text{SD}^{\text{Mod}}} \quad (6)$$

We used partial correlations to identify the main drivers of the modelled conductances. Given the high non-linearity of the equations linking the conductances to their predictors, we also used random forests (RFs) to simulate ORCHIDEE results, and we applied a permutation technique on these RF models to rank predictors (Breiman, 2001). RFs are well adapted for non-linear problems; they were for example used to rank variables of importance for soil COS fluxes in Spielman et al. (2020).

2.3 Global-scale flux estimates and comparisons with the LRU approach

We compared our estimate for plant COS uptake at global scale to former studies, with a focus on the LRU approach. We also applied the LRU approach to derive new estimates of global plant COS uptake for comparison, using a monthly climatology of our modelled GPP fluxes over the 2000–2009 period, a constant atmospheric concentration of 500 ppt for COS and global yearly values for CO_2 (from 368 ppm for the year 2000 to 386 ppm for the year 2009). We considered two sets of constant PFT-dependent LRU values. The first set (LRU_Seibt) was taken from Seibt et al. (2010), based on the observed LRU values displayed in their Table 3 (intermediate column). The second set (LRU_Whelan) used constant values for C_3 (1.68) and C_4 (1.21) plants where the values are an average over different field and laboratory measurements as assembled by Whelan et al. (2018). Both sets are listed in Table 1.

Reciprocally, we derived LRU values using Eq. (1) applied to the monthly climatology of our modelled COS and GPP fluxes over the 2000–2009 period; these will be further called LRU_MonthlyFluxes values. LRU_MonthlyFluxes values were computed for all strictly positive GPP values. For each PFT, we studied the spatio-temporal distribution of LRU_MonthlyFluxes values among grid cells where the PFT was present. We also compared these LRU_MonthlyFluxes values computed from a climatology of monthly fluxes to the climatology of monthly mean LRU values, directly computed from the original half-hourly LRU values and further called Monthly_LRU. Given the non-linearity of the problem, we expect LRU_MonthlyFluxes to be different from Monthly_LRU values. Considering that the objective of the LRU approach was to estimate COS fluxes from GPP using a constant value per PFT, the optimal LRU value for each PFT was obtained by linearly regressing monthly COS fluxes against monthly GPP fluxes multiplied by the ratio of the mean COS to CO_2 concentrations, with no offset. Thus

$$\text{LRU\_Opt} = \frac{\sum_{n=1}^{N} F_{\text{COS}}^{\text{Mod}}(n)\,\text{GPP}^{\text{Mod}}(n)\,\frac{[\text{COS}(n)]_{\text{a}}}{[\text{CO}_2(n)]_{\text{a}}}}{\sum_{n=1}^{N}\left(F_{\text{COS}}^{\text{Mod}}(n)\right)^2}, \quad (7)$$

with N the number of grid cell month simulated fluxes where the PFT is present in the monthly climatology.

We compared this new set of optimal PFT-dependent LRU values against LRU_Seibt and LRU_Whelan.

We finally used the LRU_Opt values to re-compute the monthly mean COS fluxes from our modelled monthly mean GPP and compared with the mechanistic COS flux calculation. The differences, due to the non-linearity of the COS flux calculation, provide some information on the use of a simplified approach based on mean LRU values.

2.4 Simulations of COS concentrations and evaluation at NOAA air sampling sites

The vegetation COS fluxes, as well as all other sources and sinks of the global COS budget, based on their latest estimates, are transported with an atmospheric transport model, so that we are able to simulate 3D COS atmospheric concentrations and compare them to the NOAA surface measurements.

2.4.1 The atmospheric transport model LMDz

In order to simulate COS and CO_2 concentrations in the atmosphere, we used the version of the atmospheric component LMDz of the Institut Pierre Simon Laplace Coupled Model (IPSL-CM) (Dufresne et al., 2013), which has been contributing to the CMIP6 exercise. To reduce the computation time, we used its off-line mode: precomputed air mass fluxes provided by the full version of LMDz are used to transport the different tracers (Hourdin et al., 2006). This version is further called LMDz6 and is described in Remaud et al. (2018) and references therein for the transport of CO_2. The horizontal winds are nudged towards ECMWF meteorological analyses (ERA-5, https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era5) to realistically account for large-scale advection. The tropospheric OH oxidation of COS is calculated from OH monthly data that are produced from a first simulation done with the INCA tropospheric photochemistry scheme (Folberth et al., 2006; Hauglustaine et al., 2004, 2014). The photolysis reaction of

Table 1. Table of LRU per PFT. First column: median and optimal LRU values calculated from the simulated mechanistic COS and GPP fluxes. Middle columns: calculated from Seibt et al. (2010) for the ORCHIDEE PFT classification. Last column: from Whelan et al. (2018).

PFT		ORCHIDEE		Seibt	Whelan
Long name	Abbreviation	Median	Optimal		
1 – Bare soil	Bare	0.00	0.00	0.00	0.00
2 – Tropical broadleaved evergreen forest	TroBroEver	1.56	1.72	3.09	1.68
3 – Tropical broadleaved raingreen forest	TroBroRain	1.48	1.62	3.38	1.68
4 – Temperate needleleaf evergreen forest	TempNeedleEver	1.17	1.39	1.89	1.68
5 – Temperate broadleaved evergreen forest	TempBroEver	0.86	1.06	3.60	1.68
6 – Temperate broadleaved summergreen forest	TempBroSum	1.06	1.31	3.60	1.68
7 – Boreal needleleaf evergreen forest	BorNeedleEver	0.82	0.95	1.89	1.68
8 – Boreal broadleaved summergreen forest	BorBroSum	0.84	1.03	1.94	1.68
9 – Boreal needleleaf summergreen forest	BorNeedleSum	0.76	0.92	1.89	1.68
10 – Temperate C_3 grass	Temp C_3 grass	1.01	1.18	2.53	1.68
11 – C_4 grass	C4grass	1.38	1.45	2.00	1.21
12 – C_3 agriculture	C_3 crops	1.21	1.37	2.26	1.68
13 – C_4 agriculture	C_4 crops	1.75	1.72	2.00	1.21
14 – Tropical C_3 grass	Trop C_3 grass	1.40	1.52	2.39	1.68
15 – Boreal C_3 grass	Bor C_3 grass	0.87	0.97	2.02	1.68

COS in the stratosphere is not considered: the lifetime of COS in the stratosphere is 64 years (Barkley et al., 2008). The model is set up at a horizontal resolution of $3.8° \times 1.9°$ (96 grid cells in longitude and latitude) with 39 hybrid sigma-pressure levels reaching an altitude up to about 75 km, corresponding to a vertical resolution of about 200–300 m in the planetary boundary layer. The model time step is 30 min, and the output concentrations are 3-hourly averaged.

2.4.2 Atmospheric simulations: sampling methods and data processing

We ran the LMDz6 version of the atmospheric transport model described above for the years 2000 to 2009. The prescribed COS and CO_2 fluxes used as model inputs are presented in Tables 2 and 3. The GPP estimated by ORCHIDEE ($148.1\,\mathrm{Gt\,C\,yr^{-1}}$) is in the high range among the model estimates (Anav et al., 2015), with a corresponding high respiration ($145.7\,\mathrm{Gt\,C\,yr^{-1}}$) to ensure a realistic net ecosystem exchange (Friedlingstein et al., 2019). However, other high GPP estimates can be found in the literature such as Welp et al. (2011) that suggest a range of 150 to 175 based on $\delta^{18}O$ data. Likewise, Joiner et al. (2018) have proposed a new GPP product, based on satellite data and calibrated on FLUXNET sites, with an estimate around $140\,\mathrm{Gt\,C\,yr^{-1}}$ for 2007. The fluxes are given as a lower boundary condition of the atmospheric transport model (LMDz), which then simulates the transport of COS and CO_2 by the atmospheric flow. The atmospheric COS seasonal variations are likely to be dominated by the seasonal exchange with the terrestrial vegetation, while the mean mole fractions result from all sources and sinks of COS, some of which are still largely unknown (e.g. ocean fluxes, Whelan et al., 2018). In this study, we only focus on the seasonal cycle and do not attempt to simulate the annual mean value; we thus started from a null initial state. The atmospheric transport is almost linear with respect to the fluxes: the linearity is a property of the atmospheric transport, though it is violated in LMDz because of the presence of slope limiters in the advection scheme. Overall, since all the other LMDz components are linear, LMDz transport is generally considered linear with fluxes (Hourdin and Talagrand, 2006). Relying on this relationship, we first transported each flux separately, and then we added all the simulated concentrations in the end, for each species.

For all COS and CO_2 observations, the model output was sampled at the nearest grid point and vertical level to each station and was extracted at the exact hour when each flask sample had been taken. For each station, the curve-fitting procedure developed by the NOAA Climate Monitoring and Diagnostic Laboratory (NOAA/CMDL) (Thoning, 1989) was applied to modelled and observed COS and CO_2 time series to extract a smooth detrended seasonal cycle. We first fitted a function including a second-order polynomial term and four harmonic terms, and then we applied a low-pass filter with either 80 or 667 d as short-term and long-term cut-off values, respectively, to the residuals. The detrended seasonal cycle is defined as the smooth curve (full function plus short-term residuals) minus the trend curve (polynomial plus long-term residuals).

Table 2. Prescribed COS surface fluxes used as model input. Mean magnitudes of different types of fluxes are given for the period 2000–2009.

Type of COS flux	Temporal resolution	Total ($Gg\,S\,yr^{-1}$)	Data source
Anthropogenic	Monthly, interannual	337.3	Zumkehr et al. (2018)
Biomass burning	Monthly, interannual	56.3	Stinecipher et al. (2019)
Soil	Monthly, climatological	−409.0	Launois et al. (2015b)
Ocean	Monthly, climatological	444.7	Kettle (2002) for indirect oceanic emissions (via CS2 and DMS oxidation), and Launois et al. (2015a) for direct oceanic emissions. The direct emissions are rescaled to be equal to $200\,Gg\,S\,yr^{-1}$ (*).
Vegetation uptake	Monthly, interannual	See Table 1.	This work, including mechanistic and LRU approaches (Seibt et al., 2010; Whelan et al., 2018)

* A bug has been discovered in the parameterisation of direct COS emissions in the NEMO PISCES ocean model: the hydrolysis rate was 3 times too low, resulting in an artificial build-up of COS in seawaters. As a correction, we divided the total amount of oceanic COS fluxes within a year by 3, assuming that the bug does not affect the spatial pattern of direct emissions of COS.

Table 3. Prescribed CO_2 surface fluxes used as model input. Mean magnitudes of different types of fluxes are given for the period 2000–2009.

Type of CO_2 flux	Temporal resolution	Total ($Gt\,C\,yr^{-1}$)	Data source
Fossil fuel	Monthly, interannual	7.7	ECJRC/PBL EDGAR version 4.2
Biomass burning	Monthly, interannual	1.9	GFED 4.1s
Respiration (including the land use emissions and wood harvest)	Monthly, interannual	145.7	ORCHIDEE
Ocean	Monthly, climatological	−1.3	Landschützer et al. (2015)
GPP	Monthly, interannual	−148.1	ORCHIDEE

2.4.3 COS and CO_2 concentrations at the NOAA/Global Monitoring Laboratory (GML) surface sites

We used the NOAA/GML measurements of both CO_2 and COS at 10 sites located in both hemispheres, listed in Table 4.

The samples have been collected as pair flasks one to five times a month since 2000 and are then analysed in the NOAA/GML's Boulder laboratories with gas chromatography and mass spectrometry detection. The measurements are retained only if the difference between the pair flasks is less than 6.3 ppt for COS. These COS measurements can be downloaded from the ftp site ftp://ftp.cmdl.noaa.gov/hats/carbonyl_sulfide/ (last access: 19 April 2021). The CO_2 atmospheric measurements come from the NOAA's GlobalView Plus Observation Package (ObsPack; Cooperative Global Atmospheric Data Integration Project, 2018).

2.4.4 Evaluation metrics

To evaluate and compare the performances of the mechanistic and LRU approaches at different NOAA surface sites, we used the normalised standard deviation (NSD) and the Pearson correlation coefficient (R). NSD is calculated as the ratio between the standard deviation of the simulated concentrations and the observed concentrations at the NOAA surface sites. NSD and R values closer to 1 indicate a better accuracy of the model.

3 Results

3.1 Site-scale COS fluxes, conductances, and LRU

3.1.1 COS fluxes

Diel cycle

COS assimilation is at a minimum at night (between 20:00 and 04:00 local solar time) for observed and simulated fluxes (Fig. 1a). During night, uptake of modelled COS flux is around $-8\,pmol\,m^{-2}\,s^{-1}$ while field observations vary between -5 and $0\,pmol\,m^{-2}\,s^{-1}$. In the morning, both simulated and observed uptakes increase. However, while the simulation shows a maximum assimilation of $-38\,pmol\,m^{-2}\,s^{-1}$ at noon, the maximum assimilation for observations is reached at 10:00 with a flux of $-49\,pmol\,m^{-2}\,s^{-1}$. Observed fluxes thus have a greater daily amplitude than simulated fluxes and are a little ahead of the simulation, but this shift

Table 4. List of air sampling sites selected for evaluation of COS and CO_2 concentrations.

Site	Short name	Coordinates	Elevation (m a.s.l.)	Comment
South Pole, Antarctica, United States	SPO	90.0° S, 24.8° W	2810	
Cape Grim, Australia	CGO	40.4° S, 144.6° W	164	inlet is 70 m aboveground
Tutuila, American Samoa	SMO	14.2° S, 170.6° W	77	
Cape Kumukahi, United States	KUM	19.5° N, 154.8° W	3	
Mauna Loa, United States	MLO	19.5° N, 155.6° W	3397	
Niwot Ridge, United States	NWR	40.0° N, 105.54° W	3475	
Wisconsin, United States	LEF	45.9° N, 90.3° W	868	inlet is 396 m aboveground on a tall tower
Mace Head, Ireland	MHD	53.3° N, 9.9° W	18	
Utqiaġvik, United States	UTK	71.3° N, 155.6° W	8	
Alert, Canada	ALT	82.5° N, 62.3° W	195	

does not seem significant given the large variability of observations, as represented by the 1 standard deviation in Fig. 1a. RMSD for this mean diel cycle is 8.0 pmol m^{-2} s^{-1}, and relative RMSD is 35 %. The bias is −1.7 pmol m^{-2} s^{-1}, the standard deviations are 17.5 pmol m^{-2} s^{-1} for the observations and 12.8 pmol m^{-2} s^{-1} for the simulated fluxes, and the correlation coefficient is 0.91. A similar study at the Hyytiälä site over July–September in the year 2015 (Fig. B1a) yields a similar underestimation of the amplitude of the mean diel cycle, with an RMSD of 4.0 pmol m^{-2} s^{-1} and a relative RMSD of 36 %; the bias is 2.4 pmol m^{-2} s^{-1}, the standard-deviations are 5.5 pmol m^{-2} s^{-1} for the observations and 2.7 pmol m^{-2} s^{-1} for the simulated fluxes, and the correlation coefficient is 0.93.

Seasonal cycle

The simulated weekly seasonal vegetation COS uptake roughly follows the same trend as the observed one (r = 0.53, Fig. 1b). COS uptake increases in spring when the vegetation growing season starts and decreases in autumn at the end of the forest activity period. Simulated and observed fluxes also take similar values over the 2 years. There are however differences: in 2013 the start of the season is simulated about 2 weeks too late in May instead of late April, and measured fluxes peak in May–June and August–September, while the modelled fluxes peak in July. We notice that the amplitude of observed COS flux variations is larger than the one of modelled fluxes. Kohonen et al. (2020) have quantified the relative uncertainty of weekly-averaged ecosystem COS fluxes at 40 %, which is coherent with the large standard deviation computed for field data (Fig. 1b). RMSD for the seasonal cycle is 7.0 pmol m^{-2} s^{-1}, and the relative RMSD is 41 %. The bias is low (−0.3 pmol m^{-2} s^{-1}), and the standard deviations are similar: 6.6 pmol m^{-2} s^{-1} for the observations and 7.7 pmol m^{-2} s^{-1} for the simulated fluxes. At the Hyytiälä site in the year 2015 (Fig. B1b), the RMSD for the seasonal cycle is 2.4 pmol m^{-2} s^{-1}, and the relative RMSD is 25 %; the bias is low too (0.2 pmol m^{-2} s^{-1}) and the standard deviations are also close: 3.6 pmol m^{-2} s^{-1} for the observations and 3.5 pmol m^{-2} s^{-1} for the simulated fluxes. The correlation coefficient is 0.78.

Nighttime fluxes

Figure 2 compares mean daytime and nighttime observed and modelled vegetation COS fluxes and the percentage of the daytime to the total flux, computed for each month over 2012 and 2013 at the Harvard Forest site. We selected an arbitrary PAR threshold of 50 µmol m^{-2} s^{-1} to split between daytime and nighttime fluxes. We see that the modelled nighttime flux varies across the growing season, with a maximum uptake of −10 pmol m^{-2} s^{-1} reached in July and a lower absorption in the enclosing colder months. This seasonal variation can be explained by the seasonal change in LAI and the conductance dependency on T_{air}, which increases in summer. The observed nighttime fluxes are of the same magnitude but present an opposite seasonal cycle with lower uptake at the summer peak, albeit variations are within the 1 standard deviation represented in Fig. 1a. The modelled nighttime fluxes account from 22 % of the total COS uptake at the peak of the growing season to 45 % in April at the very beginning. The observed ones exhibit slightly lower values, between 14 % and 37 %. At Hyytiälä, the modelled nighttime ratio is also slightly higher (between 30 % and 34 %) than the observed one (between 20 % and 25 %, Fig. B2). These ratios are in line with other studies: Maseyk et al. (2014) reported a ratio of 29 ± 5 % over a wheat field in Oklahoma, and Sun et al. (2018c) reported one of 23 % for the San Joaquin Freshwater Marsh site in California. The results may vary given the definitions adopted for nighttime and daytime periods.

3.1.2 Modelled conductances

To investigate the importance of each conductance in vegetation COS uptake, we compared the three simulated conductances: leaf boundary layer, stomatal, and internal, studying their variability and their drivers at the diel and seasonal scales. The boundary layer conductance to COS is higher

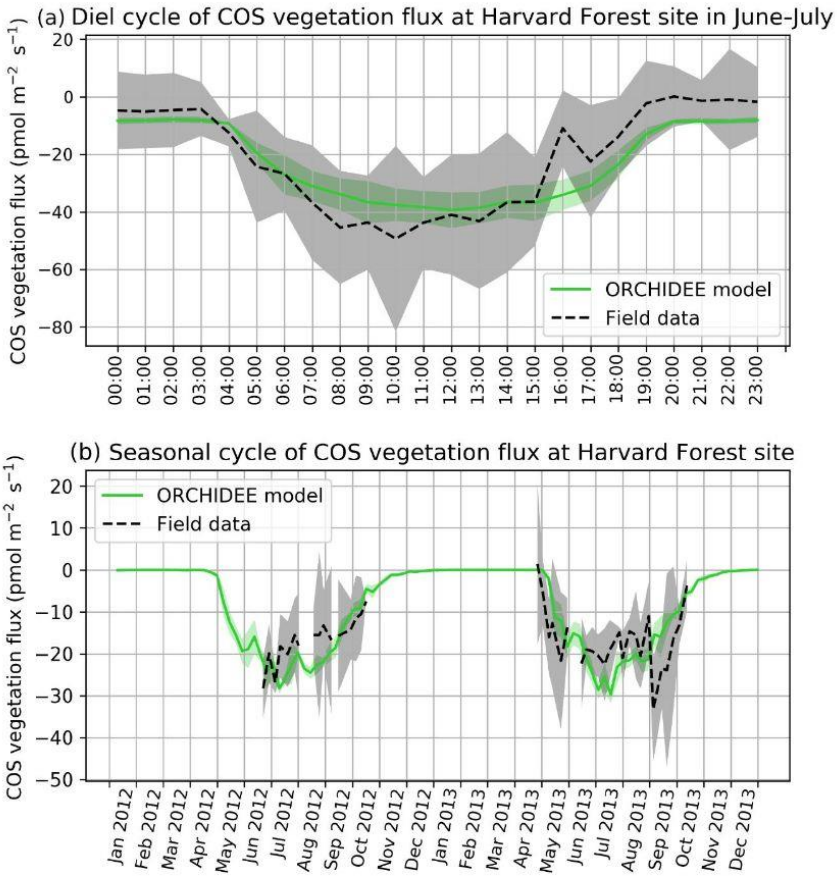

Figure 1. **(a)** Mean diel cycle of observed vegetation COS flux (Wehr et al., 2017) and modelled COS vegetation flux in June and July 2012 and 2013, at Harvard Forest, using an atmospheric convention where an uptake of COS by the ecosystem is negative. The shaded areas above and below each curve represent 1 standard deviation of the considered hourly values over the June–July period. **(b)** Mean seasonal cycle of simulated and observed weekly average vegetation COS flux in 2012 and 2013, at Harvard Forest. The shaded areas above and below each curve represent 1 standard deviation of the daily means within the considered week. We imposed the condition of having observations on at least 2 different days to compute the corresponding weekly mean.

than the two other conductances by a median factor larger than 25 (see Table A1 for more detailed statistics). As a high conductance value is equivalent to a low resistance to COS transfer, we focused only on the stomatal (g_{S_COS}) and internal (g_{I_COS}) conductances, which are the two most limiting factors to plant COS uptake.

Figure 3 presents the mean diel cycles of the simulated total, stomatal, and internal conductances for each season, computed over 2012 at Harvard Forest and 2017 at Hyytiälä. For practicality, we shifted the month of December before the month of January of the same year to compute the winter mean. The seasonal variations are similar at both sites. The conductances, as well as the amplitude of their diurnal cycle, increase from winter to summer and decline in autumn. Harvard Forest is predominantly a deciduous forest, and winter values of the conductances are zero at this site as there are no leaves in that season. Hyytiälä on the other hand is an evergreen pine forest, such that daytime stomatal conductance in winter does not become zero. The stomatal conductance peaks between 09:00 and 13:00, depending on site and season, while the internal conductance peaks later in the afternoon. The total conductance is in general limited by the internal conductance. The stomatal conductance is limiting roughly between 18:00 and 06:00 from spring to autumn at Harvard and only in June–July–August roughly between 21:00 and 09:00 at Hyytiälä.

These results are consistent with the results obtained at branch level by Kooijmans et al. (2019), who found that the COS flux was limited by the internal conductance in the early season and later during daytime, while the effect of the stomatal conductance was larger at night. For the Harvard Forest site, Wehr et al. (2017) computed the stomatal conductance using both a water flux method and a COS flux method and obtained a close agreement between two different methods; the mesophyll conductance is modelled using an experimental temperature response, and the biochemical conductance, representing CA activity, is modelled using a simple parameter (0.055 mol m^{-2} s^{-1}); both scale with LAI to get canopy

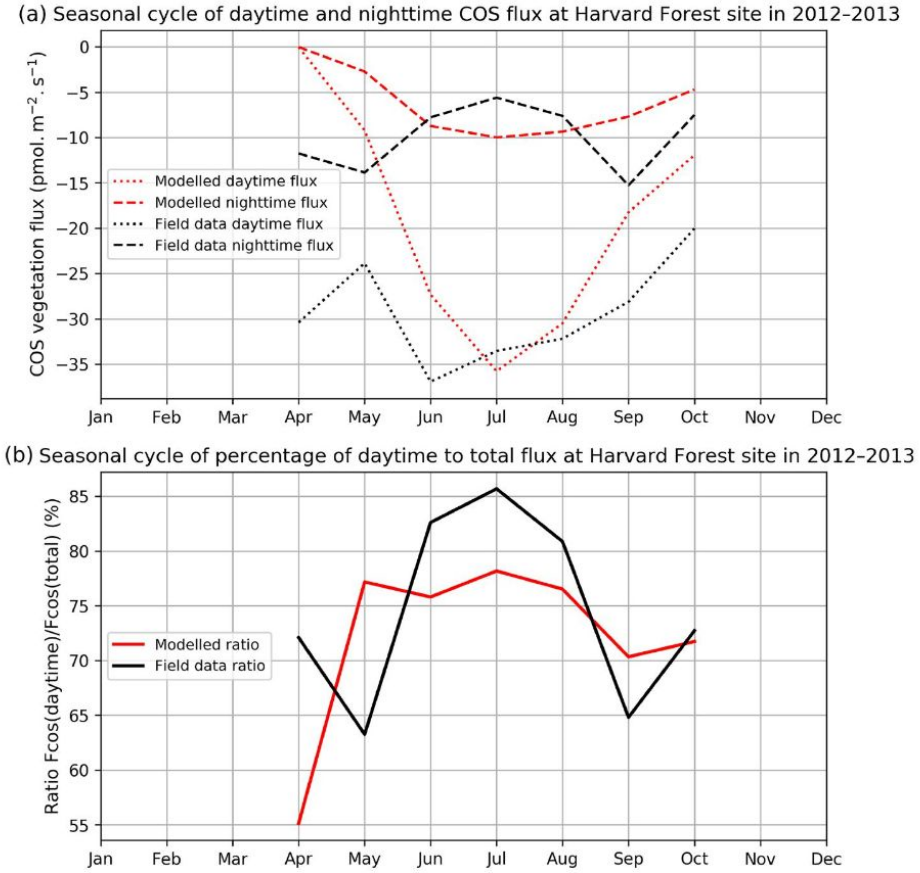

Figure 2. **(a)** Seasonal cycle of daytime (dotted curve) and nighttime (dashed curve) for observed (black) and modelled (red) vegetation COS fluxes. **(b)** Seasonal cycle of percentage of the daytime to the total flux (solid curve), at Harvard Forest in 2012–2013.

estimates. Wehr et al. (2017) found similar maximum values around 0.27 mol m^{-2} s^{-1} during daytime, from May to October, for the stomatal conductance and for the biochemical conductance (their Fig. 4); adding the slightly larger mesophyll conductance (peaking around 1.0 mol m^{-2} s^{-1}) to the biochemical conductance would thus also lead to a more limiting role of the internal conductance (peaking around 0.21 mol m^{-2} s^{-1}) during daytime, albeit not as strong as for the modelled one (peaking around 0.13 mol m^{-2} s^{-1}); the simulated stomatal conductance exhibits minimum and maximum values similar to the observation-based ones but peaks more sharply in the morning.

To better understand the conductance behaviour, we studied the relative importance of their drivers. These include environmental variables directly or indirectly involved in their modelling: air surface temperature (T_{air}), photosynthetically active radiation (PAR), vapour pressure deficit (VPD), and soil moisture (SM), as well as LAI, as leaf-level conductances are summed over LAI layers to provide canopy-level conductances. Partial correlations are computed for all half-hourly values of the variables associated with LRU values between 0 and 8 and are provided in Table A2. We also used half-hourly ORCHIDEE outputs associated with LRU values between 0 and 8 to train random forest models for conductances at the two sites, taking into account the same five predictors. A random predictor was also added to check that the variable importance was correctly estimated. All RF models have an accuracy of at least 96 %. Figures B3 and B4 present the relative ranking of the five predictors for the two conductances and the two sites. The ranking is different between the two methods (partial correlation versus RF), but they agree that at both sites the main driver for the internal conductance is air temperature and the main driver for the stomatal conductance is PAR.

As expected, g_{l_COS} mainly depends on T_{air}. This is explained by the fact that g_{l_COS} is proportional to V_{max}, which represents the Rubisco activity for CO_2; V_{max} is assumed to be a measure for the mesophyll diffusion and for the CA activity for COS, which are the components of the internal conductance (Berry et al., 2013). V_{max} depends on T_{air}, considered here to be a proxy of the leaf temperature (Yin and Struik, 2009). This strong link explains why g_{l_COS} is more limiting in winter, as T_{air} is low with thus lower enzyme activities, and, as soon as T_{air} rises in spring, g_{l_COS} becomes less limiting, especially at night. PAR is the most important variable for the stomatal conductance at the two sites. Due to

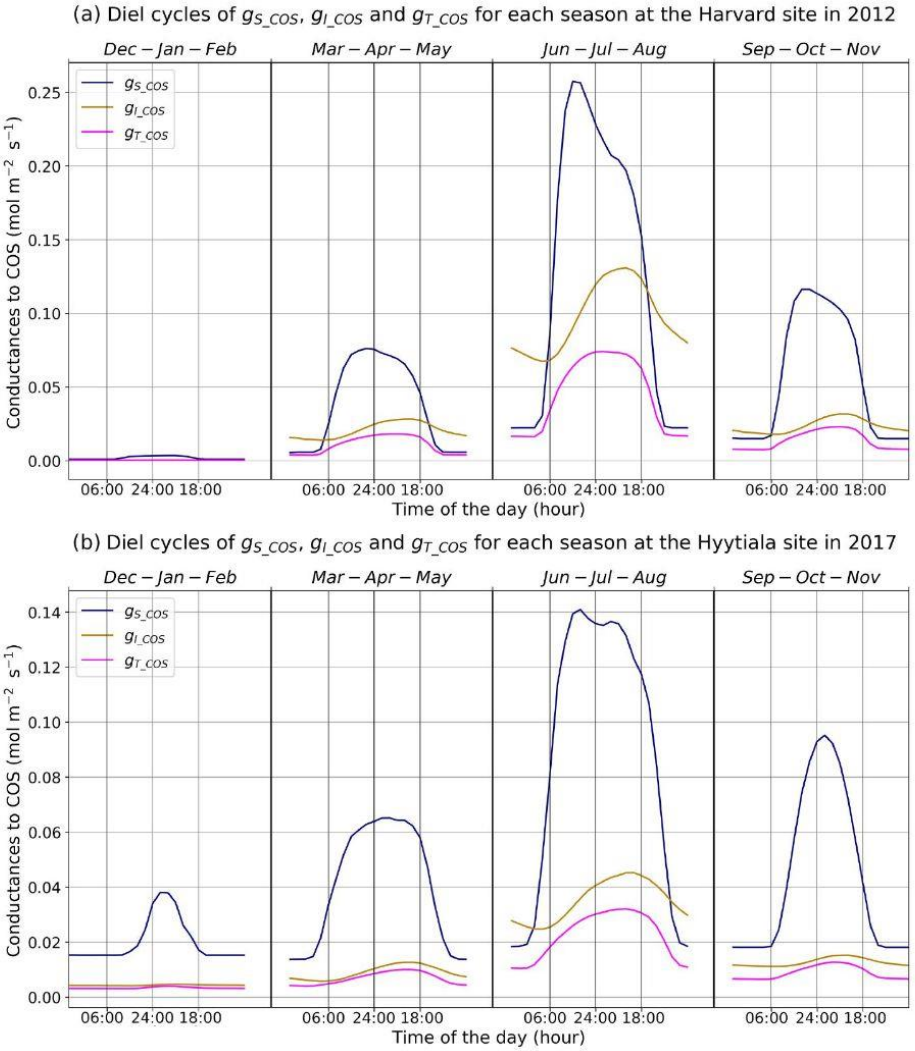

Figure 3. Mean diel cycles of simulated conductances for each season at Harvard Forest in 2012 **(a)** and Hyytiälä in 2017 **(b)**. The area reference for the units is square metres of ground area.

how g_{S_COS} is simulated according to Yin and Struik (2009), there is a linear relationship with the CO_2 assimilation, which depends mainly on PAR.

3.1.3 LRU variability

LRU decreases as a function of PAR, as initially observed by Stimler et al. (2010). Kooijmans et al. (2019) made measurements in two branch chambers installed at the top of the canopy in two Scots pine trees in Hyytiälä. They plotted the response of LRU to light, as quantified by PAR. To compare the ORCHIDEE model behaviour to these field data, we determined an LRU using our modelled COS and GPP fluxes, considering a constant atmospheric concentration of 500 ppt for COS and global yearly values for CO_2.

LRU increases with low PAR values for both branch chambers and for the model and converges towards a constant value for high PAR values (Fig. 4). This demonstrates that assuming a constant value for LRU, and not considering an increase in LRU under low-light conditions, will result in erroneous estimation of COS fluxes. The increasing LRU can be explained by the light dependence of the photosynthesis reaction contrary to the CA activity that is light-independent. Consequently, CO_2 fluxes tend to zero when PAR decreases while COS is still taken up in the dark, leading in theory to infinite values of LRU. The drop of LRU when PAR increases is however much sharper in the model than in the observations. It is to be noted that here we compare LRU values estimated from measurements at the branch level to modelled LRU estimated at canopy level. We conducted a similar modelling study considering only the top-of-canopy level and the associated COS and GPP fluxes, yielding similar results (not shown). This can be linked to the fact that the version of ORCHIDEE we use considers all the incoming light to be diffuse

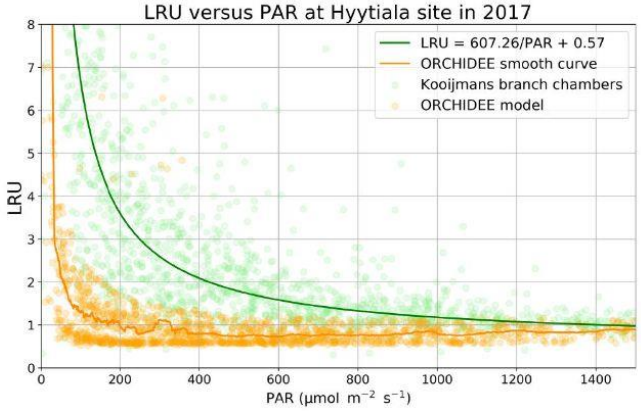

Figure 4. LRU against PAR (Hyytiälä) for ORCHIDEE outputs and measurements (hourly data measured between 18 May and 13 July; Kooijmans et al., 2019). The light green circles represent average LRU values for chambers 1 and 2, and light orange circles represent modelled LRU values. A moving average with a window of 50 points leads to the smooth orange curve for the model. The green line represents the function LRU = 607.26/PAR + 0.57 from Fig. S6 of the Kooijmans et al. (2019) supplement. To focus on LRU behaviour when PAR decreases, we plotted LRU response to PAR for PAR < 1500 µmol m^{-2} s^{-1}.

and does not distinguish between sun and shaded leaves. We thus have similar LRU values at all canopy levels.

Following the model developed in Seibt et al. (2010, their Eq. 8), the LRU explicitly depends on only two variables: the g_{S_COS}-to-g_{I_COS} ratio and the ratio of the CO_2 intracellular concentration, C_i, to $[CO_2]_a$ (equally named C_a) ratio. The modelled daily mean values for the C_i/C_a ratio computed at the two sites vary between 0.68 and 1.00 (Fig. B5). These variations are in agreement with Prentice et al. (2014), who state that the C_i/C_a ratio is pretty stable with only ±30 % variations. These values are in the upper part of the range reported in Seibt et al. (2010, their Table 2); following their Fig. 3, for a given C_i/C_a ratio a larger g_{S_COS}-to-g_{I_COS} ratio implies a lower LRU, consistent with our results.

We also performed a predictor ranking for LRU, as was done previously with conductances. The predictors rank similarly for the two sites. As shown in Fig. B6, the main factors explaining the variability of the simulated LRU at a half-hourly time step are PAR, T_{air}, and LAI.

3.2 Global-scale plant COS fluxes and study of LRU values

3.2.1 Comparison of plant COS uptake sink estimates

The mechanistic approach simulated in the ORCHIDEE model gives a plant COS uptake of −756 Gg S yr^{-1} over the 2000–2009 period. COS fluxes are the strongest in South America, Central Africa, and Southeast Asia (Fig. 5), as expected as these regions are also the most productive ones for GPP.

The more recent studies (Montzka et al., 2007; Suntharalingam et al., 2008; Berry et al., 2013; Launois et al., 2015b) show a higher global plant sink than the one initially found by Kettle et al. (2002) (Table 5). Kettle et al. (2002) used an LRU-like approach, based on net primary productivity (NPP) and on the normalised difference vegetation index (NDVI) temporal evolution, and already acknowledged their estimate was assumed to be a lower-bound one. Estimates from plant chambers and atmospheric measurements (Sandoval et al., 2005; Montzka et al., 2007; Campbell et al., 2008) confirmed that the COS plant sink should be 2-fold to 5-fold larger than estimated in Kettle et al. (2002). Suntharalingam et al. (2008) also found a low estimate of −490 Gg S yr^{-1}, using 3D modelling of COS atmospheric concentrations, constrained by surface site observations. We note that our estimate is similar to the −738 Gg S yr^{-1} found by Berry et al. (2013), which was implemented in the Simple Biosphere (SiB) 3 LSM. The reason for this similarity can be that, on top of using the same mechanistic model for vegetation COS uptake, the leaf photosynthesis and stomatal conductance in both LSMs are derived from the same classical models from Farquhar et al. (1980), Collatz et al. (1992), and Ball et al. (1987).

Launois et al. (2015b) adopted an LRU approach, using constant LRU values for large MODIS vegetation classes, adapted from Seibt et al. (2010). Based on these values and a set of global GPP estimates from three LSMs (ORCHIDEE, LPJ, CLM4), the authors derived the corresponding global vegetation COS uptakes reported in Table 5. The selection of the LSM itself thus introduces an uncertainty on the global vegetation COS uptake of around 40 % in this case.

Applying the LRU values derived from Seibt et al. (2010) (Table 1) to the global GPP simulated in this study leads to the highest plant COS uptake with −1343.3 Gg S yr^{-1}. Seibt et al. (2010) report LRU values for different internal conductance limitations. The LRU values that we used here represent a small limitation of internal conductance to the total COS uptake (the ratio of stomatal to internal conductances is 0.1). A smaller global COS uptake can be expected when the LRU values with a more limiting effect of the internal conductance are used. Applying the LRU values derived from Whelan et al. (2018) (Table 1) leads to an intermediate estimate of −808.3 Gg S yr^{-1}, which is closer to the global uptake obtained with the mechanistic model. This analysis shows that the choice for certain LRU values introduces an uncertainty on the global vegetation COS uptake (around 70 % in this case) and highlights the importance of deriving accurate PFT-dependent LRU values.

3.2.2 Dynamics of simulated LRU values

The PFT distributions of the LRU values, both those computed using Eq. (1) applied to the monthly climatology of mechanistic COS and GPP fluxes over the 2000–2009 pe-

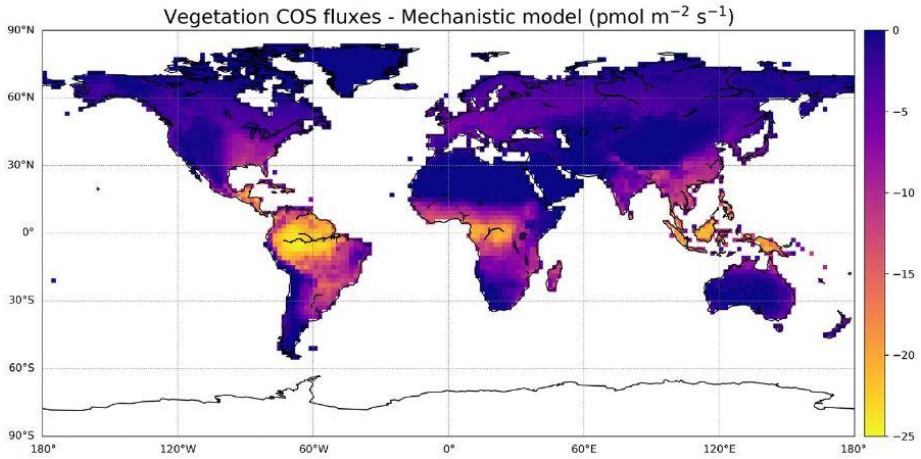

Figure 5. Map of average vegetation COS fluxes over the 2000–2009 period, from the mechanistic model as implemented in ORCHIDEE.

Table 5. Overview of COS plant uptake per year (Gg S yr^{-1}).

	Kettle et al. (2002)	Montzka et al. (2007)	Suntharalingam et al. (2008)	Berry et al. (2013)	Launois et al. (2015b) ORC.	Launois et al. (2015b) LPJ	Launois et al. (2015b) CLM4	This study
Period study	circa 1990–2000	2000–2005	2001–2005	2002–2005	2006–2009			2000–2009
Uptake by plants	−238 (±30)	−730 to −1500	−490 (−460 to −530)	−738	−1335	−1069	−930	−756

riod (LRU_MonthlyFluxes) and the climatological monthly means computed directly from the original half-hourly values (Monthly_LRU), do not support the idea of a constant PFT-dependent LRU value (Fig. 6).

The distributions are usually not Gaussian; nor are they all unimodal, as is the case for PFT 12 C_3 agriculture for C_4 PFTs (PFT 11 C_4 grass and PFT 13 C_4 agriculture) exhibit a large spread. The median values are represented by vertical red bars in Fig. 6 and listed in Table 1. The optimal values (LRU_Opt) obtained by linearly regressing monthly COS fluxes against monthly GPP fluxes multiplied by the ratio of the mean COS to CO_2 concentrations (see Fig. C1) are represented by vertical green bars and also listed in Table 1. They are usually higher than the median values, with a mean difference of 12.1 %. Using either monthly means or yearly means of fluxes gives very similar optimal LRU values, the mean difference being only −0.2 %.

The LRU values from monthly fluxes (LRU_MonthlyFluxes) tend to be lower than the monthly means of the LRU computed at a half-hourly time step (Monthly_LRU). This is visible in Fig. 6 where the blue distributions yield larger LRU values and in the bi-dimensional histogram of LRU_MonthlyFluxes against Monthly_LRU (Fig. C2). The bias is −0.2 and the correlation is 0.67. This shows that LRU is scale-dependent. The values to be considered should be coherent with their usage. For example, the optimal values we computed are lower than values estimated from measurements, but they are adapted to make the link with atmospheric COS studies.

LRU_Opt values are much smaller than LRU_Seibt values for all PFTs, roughly by a factor of 2. They are closer to the LRU_Whelan values, being smaller for all C_3 PFTs except the tropical broadleaved evergreen forests and higher for C_4 PFTs (Table 1). In the LRU_Opt set, the most productive PFTs (tropical forests and C_4 crops) have the highest values around 1.7, while the less productive PFTs (boreal forests and grasses) have the lowest values around 0.9. To the contrary, in the LRU_Seibt set, temperate broadleaved forests have the highest values (3.6) while needleleaf forests have the smallest value around 1.9.

Another way to understand the distribution of LRU values is to look directly at the scatter plots of monthly COS fluxes against GPP fluxes, multiplied by the ratio of COS to CO_2 concentrations (Fig. C1). For most PFTs, it is in fact obvious that the relationship shows non-linear features, disagreeing with the classical linear LRU model. Based on these findings, we fitted a simple exponential model as

$$F_{\mathrm{COS}} = a\left(e^{b\mathrm{GPP}\frac{[\mathrm{COS}]_a}{[\mathrm{CO_2}]_a}} - 1\right),$$

with two parameters a and b. However, given the large spread of the data around the model, the Akaike criterion is always favourable to the LRU linear model, so we will not investigate further with this exponential model. More specific research is needed here in order to bridge this data gap. Still, it is important to note that the larger COS fluxes will in gen-

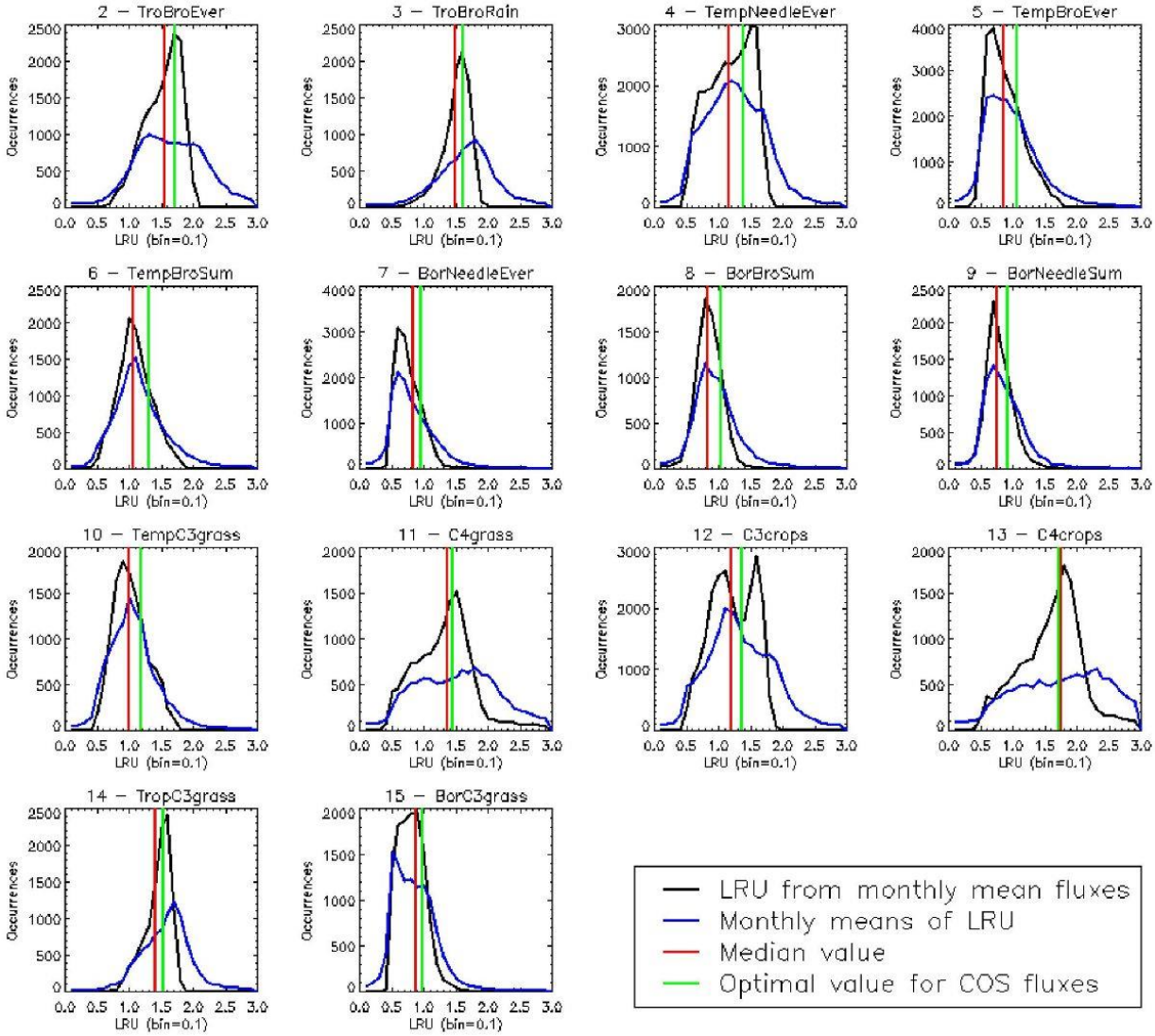

Figure 6. Distributions of the LRU values computed from the mechanistic approach over the 2000–2009 period. Each subplot represents one of the 14 vegetated PFTs used in ORCHIDEE, considering all grid cells where the PFT is present. The x axis represents the LRU value between 0 and 3, with 0.1 bins. The y axis represents the occurrences. For each PFT, the black distribution is computed using a monthly climatology of simulated COS and GPP fluxes (LRU_MonthlyFluxes), the blue distribution is computed using the monthly climatology of LRU values estimated at the original half-hourly time step (Monthly_LRU), the red vertical bar represents the median LRU value for LRU_MonthlyFluxes, and the green vertical bar represents the LRU optimal value that minimises the error between plant COS uptakes estimated at a monthly time step by the mechanistic approach and the LRU approach, for all pixels of the considered PFT (see names and abbreviations in Table 1).

eral be underestimated using a linear LRU approach. It also appears that in certain PFTs (4, 5, 7) small COS fluxes will be underestimated.

We computed mean annual vegetation COS fluxes using our modelled GPP and this new LRU_Opt set of values and compared them to the mechanistic COS fluxes (Fig. 7a).

The maps of differences between the mechanistic and LRU_Opt-based COS fluxes (Fig. 7b), and relative differences (Fig. 7c), provide evidence for the spatial errors introduced by considering a constant LRU value. The differences are always lower than 4 pmol m^{-2} s^{-1} in absolute values and are mainly positive, with the main exception over the Amazon region where the mechanistic approach shows a larger uptake than the linear LRU approach. The difference between the global estimates of the two approaches is less than 2 %; we could still improve the linear regression determining the LRU optimal value by weighting grid-cell fluxes with the corresponding surface of the PFT.

We also compared the mean seasonal cycles of the COS vegetation flux over the 2000–2009 period, for the mecha-

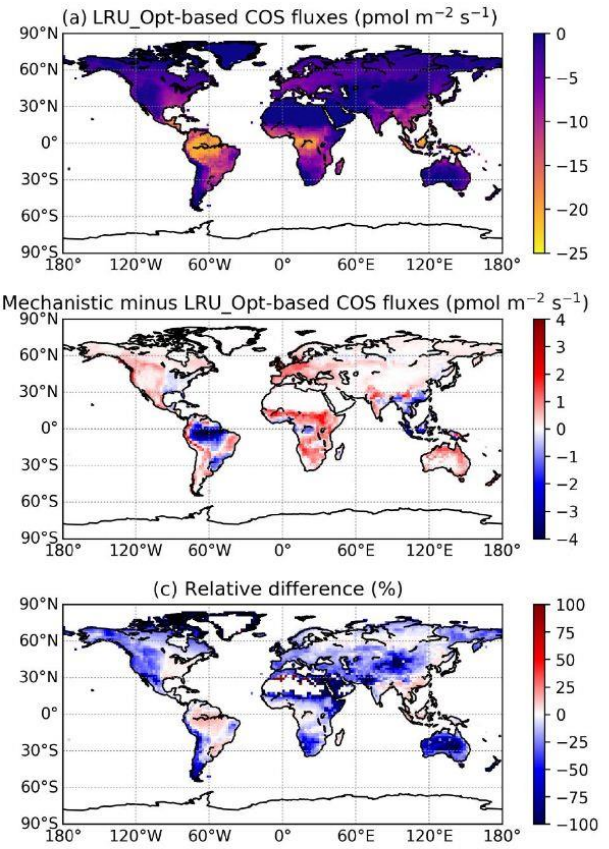

Figure 7. **(a)** Mean annual vegetation COS fluxes for the 2000–2009 period fluxes computed using a linear LRU approach with optimal values for each PFT. **(b)** Differences between mechanistic and LRU-based fluxes. **(c)** Relative difference (%).

nistic approach and the LRU_Opt-based approach, for each PFT (Fig. C3). The seasonal cycles are very similar; for PFT 13 C_4 agriculture, the LRU_Opt-based cycle is slightly in advance compared to the mechanistic cycle.

3.3 Simulating atmospheric COS concentration at surface stations

We transported the global COS and CO_2 fluxes (i.e. the ones obtained from the ORCHIDEE model plus the additional components of each cycle, listed in Tables 2 and 3) with the LMDz6 atmospheric transport model as described in Sect. 2.4.2. We analysed COS concentrations derived from simulated COS fluxes obtained with the mechanistic and LRU approaches with regards to observed COS concentrations from the NOAA at a few selected sites.

Figure 8 shows the detrended temporal evolution of CO_2 and COS concentrations for the mechanistic and LRU approaches at Utqiaġvik (UTK, Alaska) and Mauna Loa (MLO, Hawaii). The MLO site samples air masses coming from all over the Northern Hemisphere (Conway et al., 1994). CO_2 seasonal amplitude at UTK reflects the contributions of surface fluxes from high-latitude ecosystems (Peylin et al., 1999), but also from regions further south due to atmospheric transport (Parazoo et al., 2011; Graven et al., 2013). These two stations have been used to detect large-scale changes in ecosystem functioning (Graven et al., 2013; Commane et al., 2017). In spite of their importance, LMDz driven by the ORCHIDEE vegetation fluxes has difficulties in representing their seasonal cycles. For instance, at MLO, the simulated seasonal amplitude of CO_2 is overestimated and precedes the observations by 1 month.

For COS, the simulated concentrations match relatively well with the observed seasonal variations and seem to be more in phase with the observations than for CO_2. Such a feature could indicate that the phase issues with CO_2 are not primarily driven by GPP issues but by the other CO_2 flux components. The mechanistic model and its LRU optimal equivalent better reproduce the observed 1-month lag between the COS and the CO_2 simulation at MLO (i.e. the minimum COS lags the one of CO_2) than the other LRU approaches with values from Whelan et al. (2018) and Seibt et al. (2010). The simulations differ more in the amplitude than in the phase of their seasonal cycles. The mechanistic approach simulates an amplitude lower than the LRU ones. At MLO for example, the lower amplitude of the mechanistic model is in better agreement with the observations. At UTK, its seasonal amplitude is also lower but is now underestimated. The COS concentration at this station from the mechanistic approach varies between +30 and −50 ppt while it varies between +50 ppt (+37) and −71 ppt (−50) for the simulation based on Seibt et al. (2010) (Whelan et al., 2018). This is a direct consequence of lower COS fluxes with the mechanistic model compared to the fluxes based on the Seibt and Whelan LRU approaches. At both the MLO and UTK sites, the difference between the mechanistic model and its LRU optimal equivalent after being transported is lower than 8 ppt, within the range of the observation uncertainty.

Table 6 presents the NSDs and Pearson correlation coefficients between simulated and observed COS concentrations for the mechanistic and LRU approaches. We see that the simulation with Seibt et al. (2010) intermediate LRU values overestimates the seasonal standard deviation and has the lowest accuracy for most stations. It is difficult to tell whether the mechanistic model is better than the LRU approach based on Whelan values. While the mechanistic approach captures known features of the temporal dynamics of the COS-to-CO_2 flux ratio, it underestimates the simulated concentrations at Alert (ALT, Canada) and Utqiaġvik (UTK, United States). It should be noted that, due to other sources of errors (in particular transport and oceanic emissions), the comparison presented here should be taken as a sensitivity study of the COS seasonal cycle to the vegetation scheme rather than a complete validation of one approach.

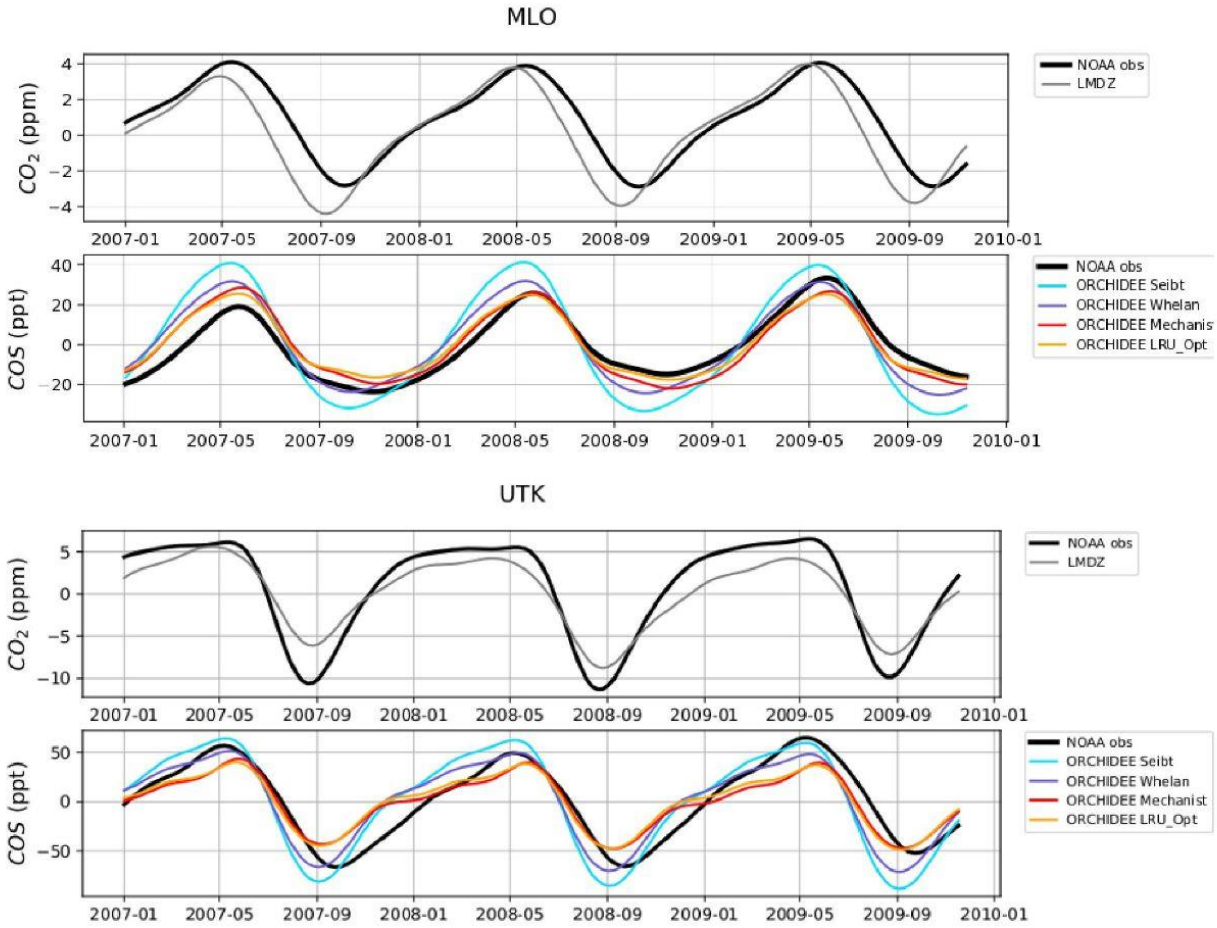

Figure 8. Detrended temporal evolutions of simulated and observed CO_2 and COS concentrations at two selected sites, for the mechanistic (ORCHIDEE Mechanist) and LRU approaches (ORCHIDEE Seibt, ORCHIDEE Whelan, ORCHIDEE LRU_Opt), simulated with LMDz6 transport between 2007 and 2009. The ORCHIDEE LRU_Opt line (orange) corresponds to the concentrations simulated using the optimal LRU values derived from the mechanistic model. Top: Mauna Loa station (MLO, Hawaii); bottom: Utqiaġvik station (UTK, Alaska). The curves have been detrended beforehand and filtered to remove the synoptic variability (see Sect. 2.2.4)

Table 6. Normalised standard deviations (NSDs) of the simulated concentrations by the observed concentrations. Within brackets are the Pearson correlation coefficients (R) between simulated and observed COS concentrations for the mechanistic and LRU approaches, calculated between 2004 and 2009 at 10 NOAA stations. For each station, NSD and R closest to 1 are in bold and the farthest ones are in italic. The time series have been detrended beforehand and filtered to remove the synoptic variability (see Sect. 2.2.4).

	SPO	CGO	SMO	KUM	MLO	NWR	LEF	MHD	UTK	ALT
ORCHIDEE Seibt	*1.15* *(0.96)*	*0.67* *(0.5)*	**0.58** (-0.47)	1.32 *(0.92)*	*1.65* *(0.89)*	*2.12* *(0.50)*	*2.17* *(0.92)*	*1.52* **(0.96)**	1.25 *(0.90)*	1.16 **(0.95)**
ORCHIDEE Whelan	**1.00** (0.97)	0.83 (0.91)	0.40 (0.1)	**1.03** (0.93)	1.23 (0.90)	1.50 (0.52)	1.67 (0.93)	1.26 (0.94)	**1.00** *(0.90)*	**0.92** *(0.94)*
ORCHIDEE Mechanist	1.10 (0.97)	**1.01** **(0.97)**	*0.35* **(0.4)**	0.90 **(0.95)**	**1.05** **(0.92)**	1.26 **(0.63)**	**1.34** **(0.94)**	1.09 *(0.85)*	0.69 **(0.91)**	*0.64* **(0.95)**
ORCHIDEE LRU_Opt	1.02 **(0.98)**	0.98 **(0.97)**	0.34 *(-0.5)*	*0.85* (0.94)	0.94 **(0.92)**	**1.21** *(0.50)*	**1.34** **(0.94)**	**1.04** (0.88)	*0.68* **(0.91)**	*0.64* **(0.95)**

4 Discussion

4.1 How can we use COS fluxes and the mechanistic COS model to improve the simulated GPP?

The mechanistic model links vegetation COS uptake and GPP fluxes through the stomatal conductance model, which includes the minimal conductance as an offset, and the common use of the carboxylation rate of Rubisco, V_{max}, in the internal conductance formulation for COS, and in the Rubisco-limited rate of assimilation for CO_2. The downside is the introduction of the somewhat uncertain α parameter that relates the COS internal conductance to V_{max}. Using COS flux measurements to optimise the parameters of the stomatal and internal conductances would thus in principle benefit the simulated GPP. This optimisation may be done based on appropriate data assimilation techniques; for example, Kuppel et al. (2012) optimised key parameters of the ORCHIDEE model related to several processes including photosynthesis (see their Table 2), by assimilating eddy-covariance flux data over multiple sites. The approach relies on a Bayesian framework where a cost function including uncertainties on observations, model, and parameters is minimised (Tarantola, 1987). The results obtained in this study pave the way for a similar approach using COS fluxes to optimise key parameters controlling GPP; they can be used to define an optimal set-up for the a priori errors and the error correlations in a Bayesian framework. We acknowledge however the scarcity of available measurements for the time being, with no samples for most biomes, a few sites with less than 1 year of data, and only Hyytiälä allowing for interannual variability studies.

4.1.1 First step: improving the mechanistic modelling of vegetation COS fluxes

Without any calibration, the chosen mechanistic model was able to reproduce observed vegetation COS fluxes at the Harvard Forest and Hyytiälä sites with relative RMSDs on the order of 40 %. Regarding conductances, differences are also seen between the diel cycles of simulated and observation-based conductances from Wehr et al. (2017). Diel variations in atmospheric $[COS]_a$, not accounted for in our model, cannot explain these differences, as they would only affect F_{COS} but not the conductances. These discrepancies advocate for the assimilation of COS fluxes to optimise the parameters related to the internal and stomatal conductances. In our modelling framework, the internal conductance is assumed to be the product of V_{max} by the α parameter. This parameter has been calibrated by Berry et al. (2013) using gas exchange measurements of COS and CO_2 uptake (Stimler et al., 2010, 2012). As this α parameter seems much more uncertain compared to the relatively well-known V_{max}, we should first try to optimise α keeping V_{max} fixed.

4.1.2 Exploiting the alternative dominant role between stomatal and internal conductances

Without being perfect, the mechanistic model could reproduce some expected behaviours, such as the limiting role of the internal conductance in winter and then during daytime in the growing season, in relation to the control of CA activity and mesophyll diffusion by temperature, as also depicted in Kooijmans et al. (2019). Determining the limiting conductances to COS uptake depending on the time of day provides useful information, as it can be used to better target which model parameters to optimise, using data assimilation approaches. Thus, observations made in the morning and early afternoon could be used to better constrain the α parameter when the internal conductance limits COS fluxes, at least as modelled on the C_3 species of the two sites, and we could investigate whether the α parameter should be further quantified per PFT rather than simply per photosynthetic pathway. It is to be noted that for C_4 species, the internal conductance is larger than for C_3 species by a factor of 10, so that stomatal conductance is limiting, and it could be difficult and useless to try optimising internal conductance using the α parameter. We have to acknowledge the large uncertainty regarding the modelling of the internal conductance. In parallel to optimising the parameters of the internal conductance, an improvement could thus also be to replace it by the two factors it represents, i.e. the mesophyll conductance and CA activity. A model for the mesophyll conductance is already implemented in ORCHIDEE, with a simple parameter depending on temperature through a multiplication by a modified Arrhenius function following Medlyn and al. (2002) and Yin and Struik (2009). The impact of mesophyll conductance on photosynthesis and water use efficiency is now more studied (e.g. Buckley and Warren, 2014), even if its modelling remains challenging too: the temperature response has notably been reported as highly variable between plant species (von Caemmerer and Evans, 2015), which would imply having PFT-dependent parameters. Regarding measurements, ^{13}C discrimination of the isotopic composition of CO_2 exchanges allows for an estimation of the mesophyll conductance (Stangl et al., 2019). Concerning CA activity, we could test the simple model using a constant value presented in Wehr et al. (2017). Measuring CA activity can be done at a coarse frequency, using different techniques (Henry, 1991).

4.1.3 Exploiting nighttime conductances

Recent studies have shown that nighttime field measurements of stomatal conductances often exhibit larger values than the ones used in models (Caird et al., 2007; Phillips et al., 2010). In the ORCHIDEE model, minimum stomatal conductances to CO_2, g_0, take two different values: 6.25 mmol m^{-2} s^{-1} for C_3 species and 18.75 mmol m^{-2} s^{-1} for C_4 species. However, Lombardozzi et al. (2017), using data from literature,

found that observed nighttime conductances to CO_2 range from 0 to 450 mmol m^{-2} s^{-1} with an overall mean value of 78 mmol m^{-2} s^{-1}. Moreover, they defined a mean value for each PFT (see Table A3) while the ORCHIDEE model uses one value for all C_3 species and another one for all C_4 species. Using higher nighttime stomatal conductances in models has the impact of increasing plant transpiration and reducing available soil moisture, which alters water and carbon budgets, especially in semi-arid regions (Lombardozzi et al., 2017). Lower VPD values at night, which could limit the impact of higher nighttime stomatal conductances, follow an increasing trend however (Sadok and Jagadish, 2020). A better representation of these minimal conductances in the model could then improve the constraint of gas exchange between the atmosphere and the terrestrial biosphere. It is to be noted that Barnard and Bauerle (2013) found, based on sensitivity analyses, that g_0 was the parameter having the largest influence on their modelled transpiration estimates. They also stress that g_0 should maybe be seen as an asymptotic minimal value, rather than an offset. During nighttime, the stomatal conductance limits COS uptake. In the model, the nocturnal stomatal conductance to COS is calculated from the above-mentioned minimum stomatal conductance values. For now, the absolute vegetation COS fluxes at night are slightly overestimated compared to observed fluxes (updated Fig. 1a for Harvard and Fig. B1a for Hyytiälä), thus hinting to overestimated nighttime stomatal conductances. Therefore, nighttime observations of COS fluxes could be used to optimise the minimum stomatal conductance values for each PFT.

We thus see that COS fluxes could be used, through standard data assimilation techniques, to optimise the model parameters related to conductances, thus contributing to the improvement of the GPP. However, many more COS flux measurements are needed over a large variety of biomes, first to assert the validity of the mechanistic COS model at global scale and second to be assimilated in order to improve simulated conductances and GPP estimates.

4.2 The mechanistic versus LRU approach

The mechanistic model is able to reproduce the high-temporal-frequency LRU variations observed at sites. It is thus legitimate to consider this approach to be more accurate than the classical linear LRU approach that uses a time-constant LRU value per PFT to estimate COS fluxes from GPP. Furthermore we have shown that computing LRU values using Eq. (1) applied to monthly mean fluxes yields values lower than computing monthly means of high-frequency LRU values (Fig. 6). This may explain why the LRU values we have thus estimated from monthly mean fluxes show generally lower values than the ones derived from measurements, although these cover a large range from 0.7 to 6.2 (Seibt et al., 2010; Whelan et al., 2018). More recently, Spielman et al. (2019) estimated LRU values from ecosystem and soil measurements: 0.89 for an agricultural soybean field, 1.02 for a temperate C_3 grassland, 2.22 for a temperate beech forest, and 2.27 for a Mediterranean savanna ecosystem; our corresponding PFTs respectively give 1.37 (C_3 crops), 1.18 (Temp C_3 grass), 1.31 (TempBroSum), and 1.06 (TempBroEver), with thus higher estimates for herbaceous plants and lower ones for trees. It is difficult to say whether in situ and laboratory measurements are too sparse and not representative enough of the variability of plants and environmental conditions across the globe to have a reasonable confidence in their derived mean or median LRU values, or whether we can use these LRU values to falsify the modelled COS and/or GPP fluxes. We may also add that LRU values derived from measurements performed in leaf chamber measurements, which are well ventilated and thus associated with large leaf boundary layer conductances, may not be representative of the real-world transfer processes, where the boundary layer conductances vary with wind speed, temporally and within canopy depth (Wohlfahrt et al., 2012).

Without any calibration, the mechanistic approach performs similarly to LRU approaches based on monthly mean fluxes, when COS is transported using all known COS fluxes as inputs, and COS concentrations are evaluated at stations of the NOAA network. We now have a much finer representation of the COS fluxes as, at every time step, the model integrates the plant's response to environmental conditions in the calculation of the internal and stomatal conductances, unlike in the LRU approach which uses constant values for each PFT.

In order to quantify the first-order uncertainty on F_{COS} related to the fact that we have used a constant $[COS]_a$ in our implementation of the Berry model, we computed an alternative $F_{COS'}$, using the LRU approach based on a climatology of hemispheric monthly means of COS atmospheric concentrations (Montzka et al., 2007), the optimal LRU we derived in this study (given in Table 1), average yearly values for CO_2 atmospheric concentrations, and a climatological seasonal cycle of simulated monthly GPP per PFT. Over the 2000–2009 period, the mean difference between the mean seasonal COS fluxes computed with this method ($F_{COS'}$) and the ones simulated with the mechanistic model (F_{COS}) amounts to −7.9 % over the Northern Hemisphere. As expected, the seasonal amplitude of COS fluxes is dampened as $[COS]_a$ decreases with vegetation growth. We thus have to improve our methodology to consider a varying $[COS]_a$ as was done in Berry et al. (2013), either inside the ORCHIDEE model or as post-processing. This requires devising some trade-off between the high-frequency time step of ORCHIDEE and the cost of running the transport model. However, it is to be noted that there is no impact on the derived LRU values as the LRU does not depend on the considered $[COS]_a$, as long as the same one is considered for the computation of the COS fluxes in the mechanistic model (Eq. 3) and for the computation of the LRU (Eq. 1) (i.e. whether fixed or varying monthly).

However, there is currently a larger uncertainty on other COS fluxes in the global COS budget, which have an important impact on simulated COS concentrations (Ma et al., 2020) and their relative seasonal changes. For example, if we use another estimation of the direct oceanic fluxes (Lennartz et al., 2017), which shows a seasonal cycle whose amplitude is comparable to the one from the vegetation in high latitudes, this results in an overestimated seasonal cycle at all sites, with the mechanistic approach having the most realistic seasonal amplitude (see Appendix D1 and Fig. D1). An additional sensitivity test was performed to assess the impact of indirect oceanic emissions via DMS oxidation on simulated seasonal cycles as the importance of these fluxes in the global COS budget is still debated (Whelan et al., 2018). Whereas the impact on northern sites is negligible, the removal of indirect oceanic emissions via the DMS of Kettle et al. (2002) decreases the seasonal amplitude of southern sites (CGO and SPO) in the same proportion in all experiments (see Appendix D2 and Table D2). Transport errors also add uncertainties on the simulated concentrations, especially at elevated continental sites (Remaud et al., 2018). Plus, given the present discrepancies between the GPP estimates of different land surface models, it can be argued that using a mechanistic model instead of an LRU approach when comparing COS concentrations seems to be of a second-order importance (Campbell et al., 2017; Hilton et al., 2017). We nevertheless note in this study that we found an uncertainty on the global vegetation COS uptake of 40 % when considering three different LSMs (Launois et al., 2015b), to be compared to an uncertainty of 70 % when considering three LRU datasets.

Setting aside the uncertainty for the moment, how could we use atmospheric COS concentrations to constrain GPP? A first optimisation was performed with the ORCHIDEE model in Launois et al. (2015b), who optimised a single scaling parameter applied on the vegetation COS fluxes simulated with the LRU approach, thus equivalent to a scaling factor applied on the GPP or the LRU. They assimilated the atmospheric COS concentrations measured at the NOAA air sampling stations, using the LMDz transport model (Hourdin et al., 2006) and a Bayesian framework as in Kuppel et al. (2012). The optimisation reduced in absolute value the estimated global vegetation COS uptake from -1335 to $-708\,\mathrm{Gg\,S\,yr^{-1}}$, more in line with this work's estimate based on a mechanistic modelling of vegetation COS uptake. A mid-term perspective is to go beyond a single scaling parameter and to optimise a set of ORCHIDEE parameters using both atmospheric COS and CO_2 data. Such an approach has been used in several studies with CO_2 data only (e.g. Rayner et al., 2005; Peylin et al., 2016). However, compared to CO_2, the spatial coverage of COS surface observations is still too sparse to accurately constrain the GPP and therefore ORCHIDEE parameters (Ma et al., 2020). There is some hope that new satellite retrievals of COS column content, such as with the IASI (Infrared Atmospheric Sounder Interferometer) instrument, could have enough accuracy to better constrain the surface fluxes (Serio et al., 2020).

5 Conclusions and outlooks

We have implemented the mechanistic model of Berry et al. (2013) inside the ORCHIDEE land surface model for COS uptake by the continental vegetation. Modelled COS fluxes were compared at site scale against measurements at the Harvard temperate deciduous broadleaf forest (USA) and at the Hyytiälä Scots pine forest (Finland), yielding relative RMSDs of around 40 % at both diel and seasonal scales. We found that the mechanistic model yields a lower and thus more limiting internal conductance compared to former works (Seibt et al., 2010; Wehr et al., 2017). The next step is to perform a sensitivity analysis (Morris, 1991; Sobol, 2001) and to optimise the most sensitive parameters related to the modelled fluxes and conductances, to get a better agreement with observations.

Our global estimate of COS uptake by continental vegetation of $-756\,\mathrm{Gg\,S\,yr^{-1}}$ is in the lower range of former studies. An important finding is that the LRU computed from monthly values of the COS and GPP fluxes yields values lower than monthly means of high-frequency LRU values. This has consequences for atmospheric studies where COS concentrations integrate influences from fluxes at large spatial and temporal scales.

Using appropriate LRU values, we transported the monthly mean COS fluxes from the mechanistic and LRU approaches using the LMDz6 model. The evaluation of the modelled COS atmospheric concentrations against observations at stations of the NOAA network yields comparable results for both approaches.

As a general conclusion and for the moment, we can say that the mechanistic model is particularly valuable when studying small timescales or spatial scales using COS fluxes, while for global analyses using COS concentrations, both the mechanistic and LRU approaches give similar results. The fact that the global COS budget has so many components with a large uncertainty (Whelan et al., 2018) limits the use of COS concentrations as a constraint for GPP in land surface models on the global scale, for the present time.

A further development will be to refine the estimation for COS soil fluxes and to implement a mechanistic model for soil COS fluxes inside ORCHIDEE (Ogée et al., 2016; Sun et al., 2015). Having both the vegetation and soil contributions, we will also be able to assimilate ecosystem COS fluxes to optimise COS-related parameters such as α in the internal conductance formulation from the Berry et al. (2013) model for vegetation uptake, and those related to the stomatal conductance (Wehr et al., 2017; Berkelhammer et al., 2020). We will also later look at the complementary constraints on GPP brought by COS and solar-induced fluorescence, another GPP proxy (Bacour et al., 2019; Whelan et al., 2020).

Appendix A: Additional tables related to conductances

Table A1. Ratios of modelled boundary conductance to stomatal conductance and internal conductance at the two studied sites, computed over the year 2012 at Harvard Forest and 2017 at Hyytiälä.

	Harvard Forest		Hyytiälä	
Ratio	Boundary to stomatal	Boundary to internal	Boundary to stomatal	Boundary to internal
Median	28	69	46	228
Minimum	9	20	17	48
Maximum	188	1523	232	9304

Table A2. Partial correlations linking stomatal and internal conductances to photosynthetically active radiation (PAR), air temperature (T_{air}), vapour pressure deficit (VPD), soil moisture (SM), and leaf area index (LAI), computed at a half-hourly time step over the year 2012 at the Harvard Forest site and 2017 at the Hyytiälä site. For each conductance and each site, the largest partial correlation is in bold.

Conductance	Site	PAR	T_{air}	VPD	SM	LAI
g_{S_COS}	Harvard	**0.66**	0.46	−0.61	−0.04	0.33
	Hyytiälä	**0.59**	0.49	−0.47	−0.03	0.25
g_{I_COS}	Harvard	−0.06	**0.68**	0.30	−0.27	0.15
	Hyytiälä	−0.13	**0.74**	0.65	0.32	0.49

Table A3. Minimum stomatal conductance to CO_2 (mmol m^{-2} s^{-1}) for each PFT in Lombardozzi et al. (2017) and ORCHIDEE. No value is given for C_4 crops in Lombardozzi et al. (2017).

	Mean minimum conductance in Lombardozzi et al. (2017)	Minimum conductance in ORCHIDEE
1 – Bare soil	0	0
2 – Tropical broadleaved evergreen forest	90.488	6.25
3 – Tropical broadleaved raingreen forest	109.744	6.25
4 – Temperate needleleaf evergreen forest	16.896	6.25
5 – Temperate broadleaved evergreen forest	34.017	6.25
6 – Temperate broadleaved summergreen forest	72.637	6.25
7 – Boreal needleleaf evergreen forest	8	6.25
8 – Boreal broadleaved summergreen forest	50	6.25
9 – Boreal needleleaf summergreen forest	29	6.25
10 – C_3 grass	157.988	6.25
11 – C_4 grass	93.933	18.75
12 – C_3 agriculture	60.629	6.25
13 – C_4 agriculture	x	18.75

Appendix B: Additional illustrations for results at the site scale

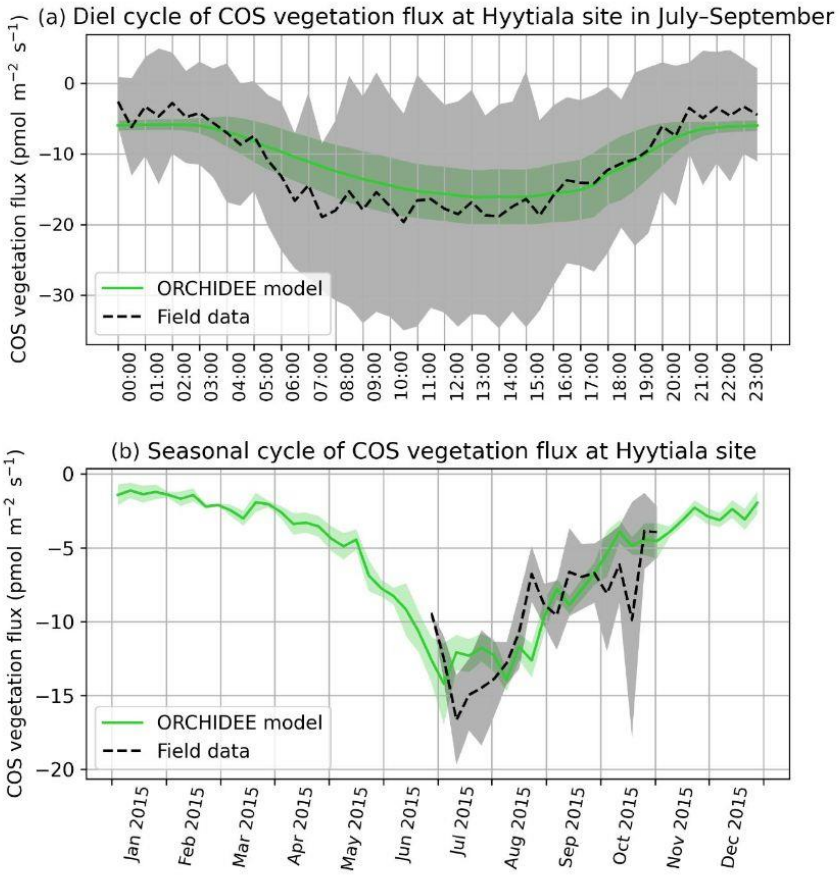

Figure B1. **(a)** Mean diel cycle of observed vegetation COS flux derived from ecosystem COS flux (Kohonen et al., 2020) and soil COS flux (Sun et al., 2018a), and modelled COS vegetation flux in July–September 2015, at Hyytälä, using an atmospheric convention where an uptake of COS by the ecosystem is negative. The shaded areas above and below each curve represent 1 standard deviation of the considered half-hourly values over the July–September period. **(b)** Mean seasonal cycle of simulated and observed weekly average vegetation COS flux in 2015, at Hyytälä. The shaded areas above and below each curve represent 1 standard deviation of the daily means within the considered week.

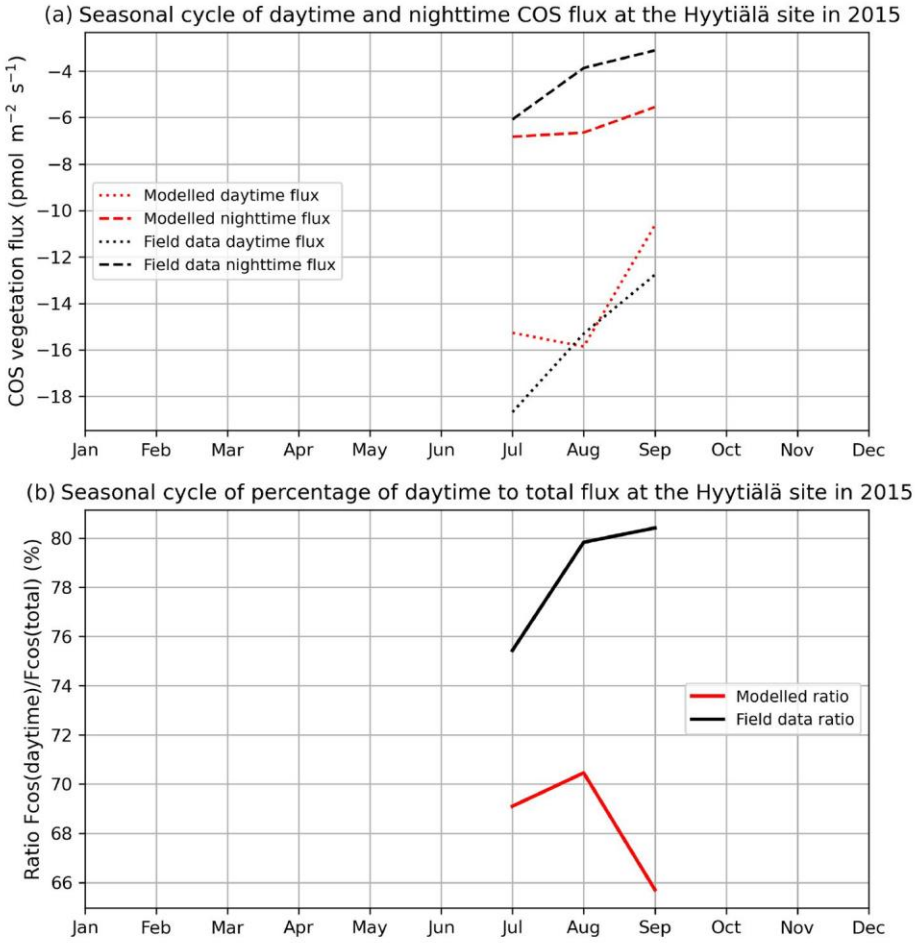

Figure B2. **(a)** Seasonal cycle of daytime (dotted curve) and nighttime (dashed curve) for observed (black) and modelled (red) vegetation COS fluxes. **(b)** Seasonal cycle of percentage of the daytime to the total flux (solid curve), at the Hyytiälä site in 2015.

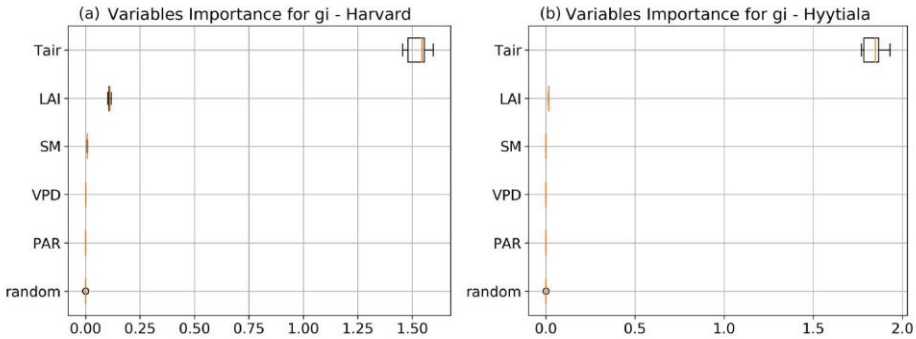

Figure B3. Variables' importance computed using random forests for the internal conductance (gi) at the Harvard Forest site in 2012 (left) and at the Hyytiälä site in 2017 (right). The considered predictors are air temperature (T_{air}), leaf area index (LAI), soil moisture (SM), vapour pressure deficit (VPD), and photosynthetically active radiation (PAR). A random predictor is added to prevent over-fitting.

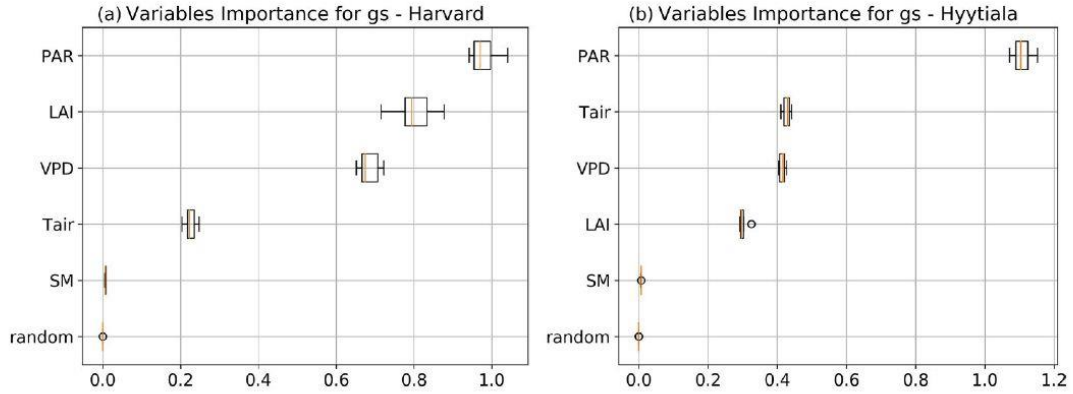

Figure B4. Same as B3 for the stomatal conductance (gs).

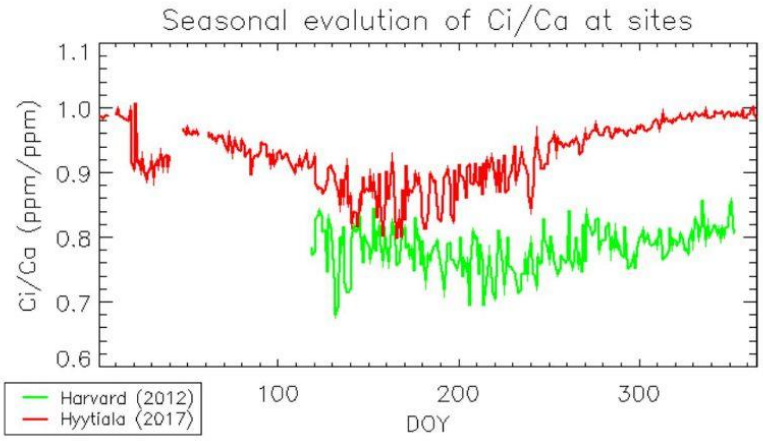

Figure B5. Seasonal evolution of the simulated C_i-to-C_a ratio at the Harvard Forest site in 2012 (green curve) and the Hyytiälä site in 2017 (red curve).

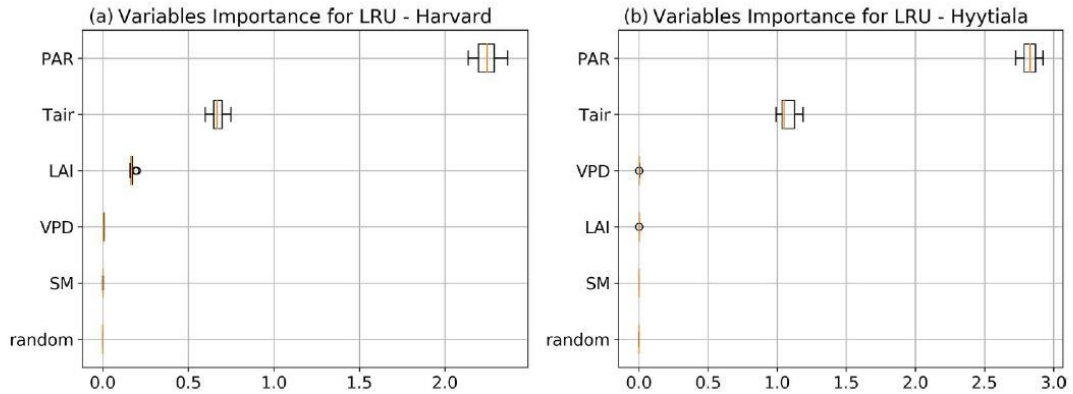

Figure B6. Same as B3 for the leaf relative uptake (LRU).

Appendix C: Additional illustrations for results at the global scale

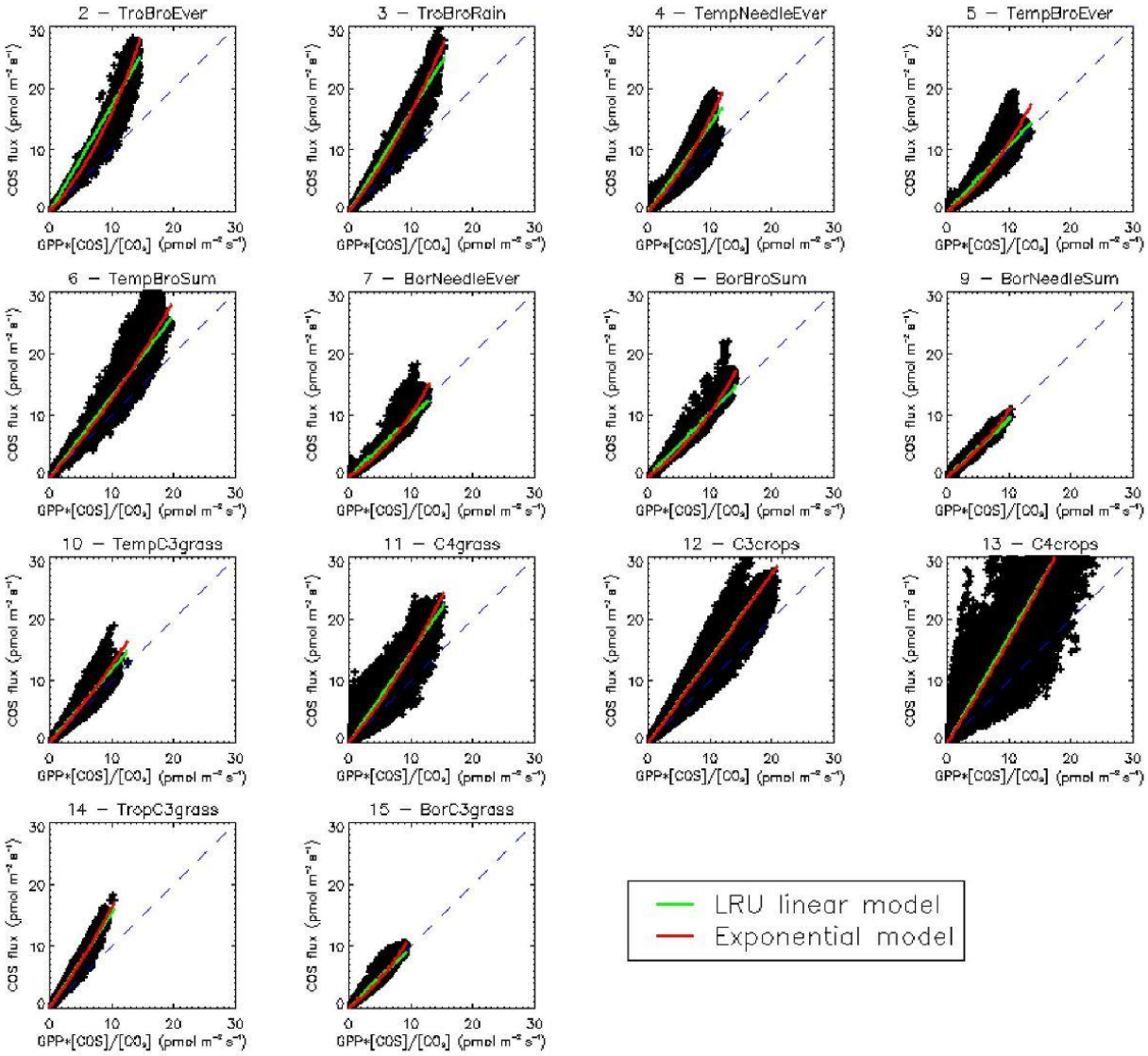

Figure C1. Scatterplots of COS fluxes against GPP multiplied by the ratio of COS to CO2 concentrations, using a climatology of monthly fluxes over the 2000–2009 period and yearly global averages for CO_2 concentrations and a fixed value of 500 ppt for the COS concentration. Each subplot represents one of the 14 vegetated PFTs used in ORCHIDEE. The LRU model in green represents the linear regression, while the exponential model (see text) is represented in red. The blue dashed lines show the 1 : 1 line.

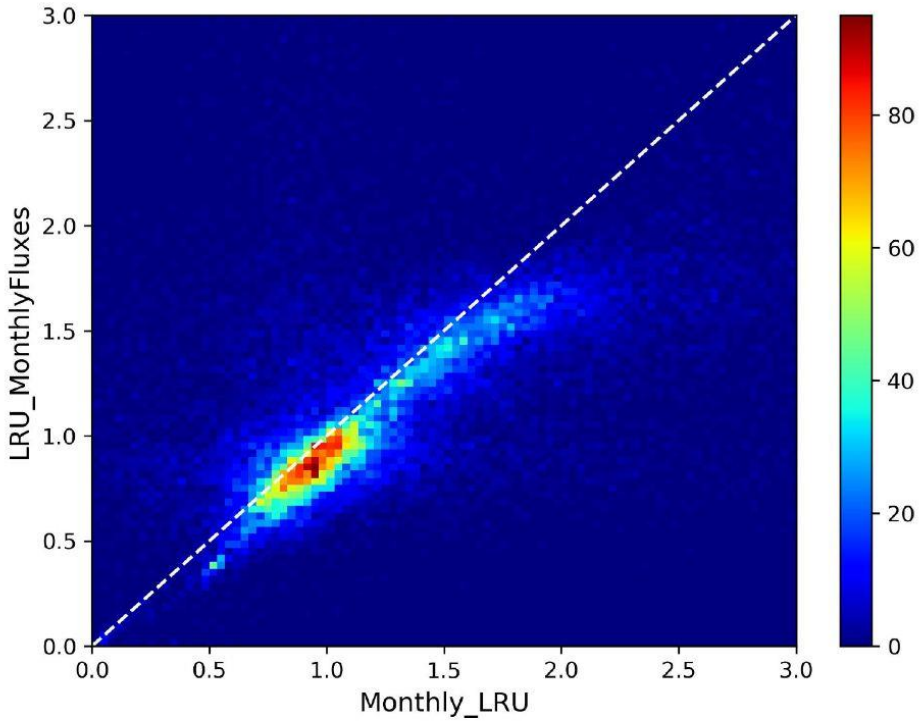

Figure C2. Bi-dimensional histogram of LRU values computed from a climatology of monthly mean fluxes (LRU_MonthlyFluxes) against a climatology of monthly means of LRU computed from original half-hourly values (Monthly_LRU). The colour bar indicates the number of occurrences per bin of 0.1×0.1 size. The white dashed line represents the first bisector.

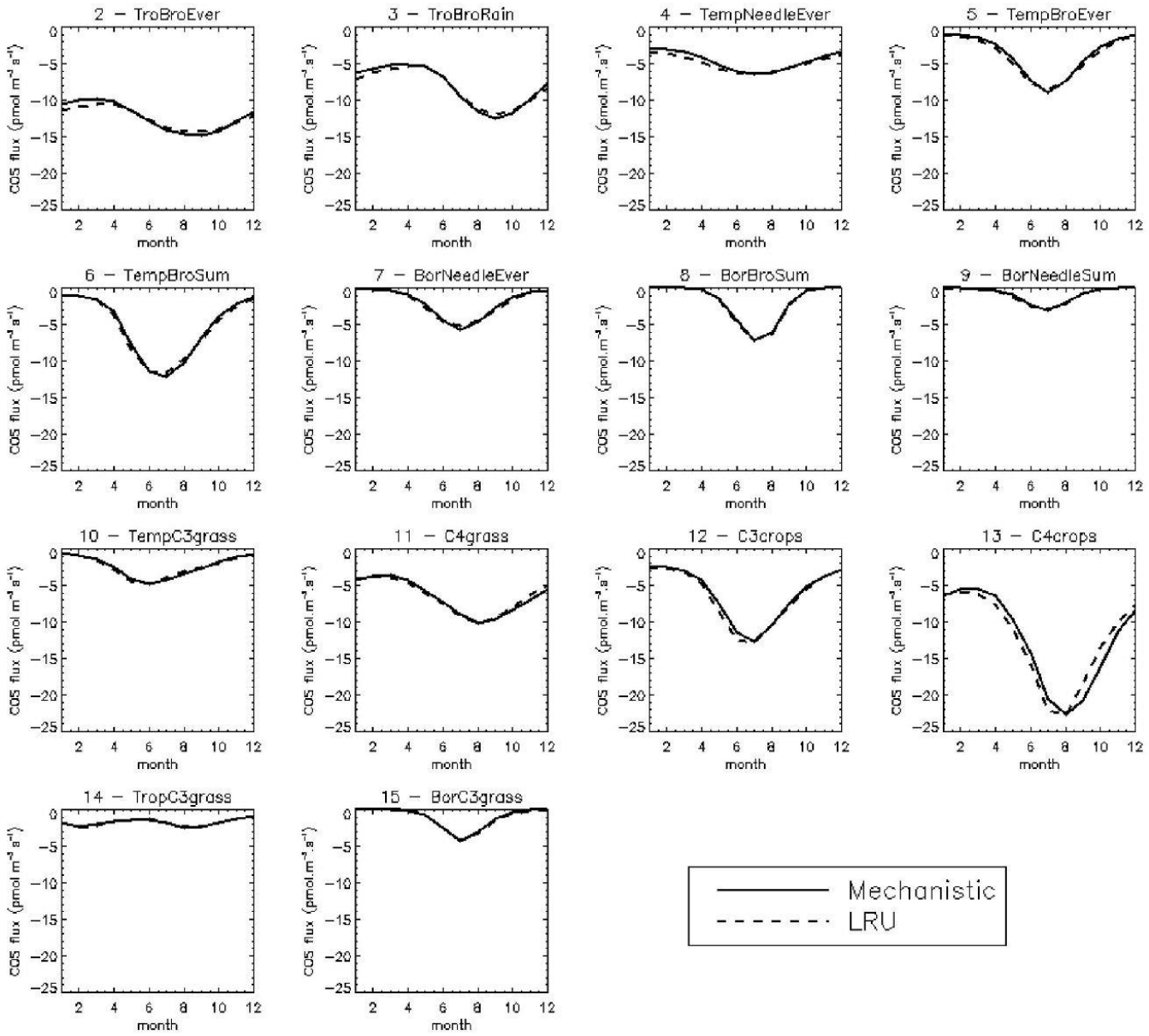

Figure C3. Mean seasonal cycle (monthly means) of COS for each PFT over the Northern Hemisphere for the 2000–2009 period. The solid line represents the mechanistic model, while the dashed line represents the optimal LRU approach.

Appendix D: Sensitivity tests for the modelling of atmospheric COS concentrations

D1 Simulating COS atmospheric concentration at stations: impact of the oceanic emissions

We performed the same experiment as in Sect. 3.4, except that the oceanic fluxes (direct and indirect) here are from Lennartz et al. (2017). In our case, the oceanic emissions (in particular direct oceanic emissions) have more impact than the LRU on the seasonality at surface sites from the NOAA network.

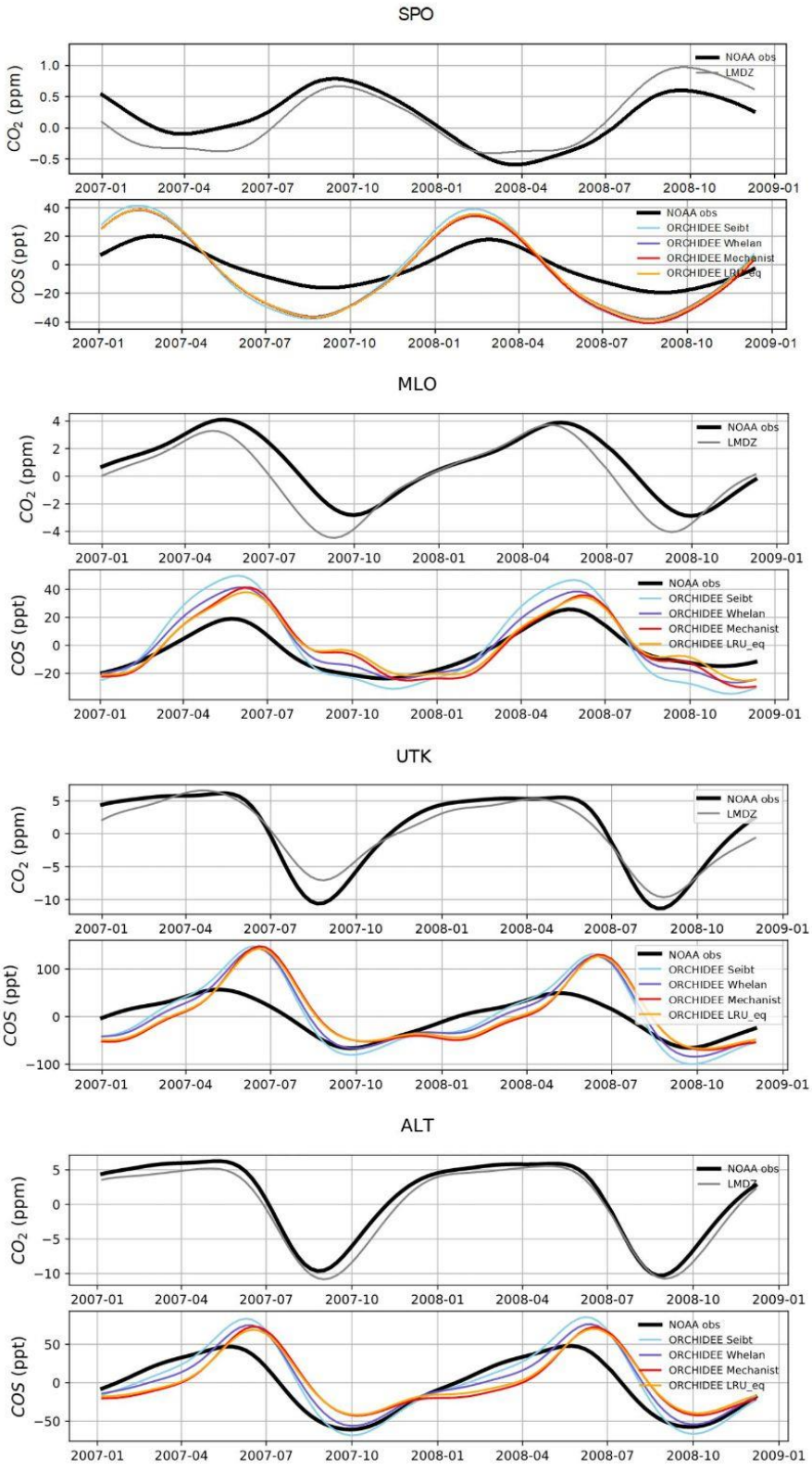

Figure D1. Detrended temporal evolutions of simulated and observed CO_2 and COS concentrations at four selected sites, for the mechanistic (ORCHIDEE Mechanist) and LRU approaches (ORCHIDEE Seibt, ORCHIDEE Whelan, ORCHIDEE LRU_Opt), simulated with LMDz6 transport between 2007 and 2009. The ORCHIDEE LRU_Opt line (orange) corresponds to the concentrations simulated using the optimal LRU values derived from the mechanistic model. The curves have been detrended beforehand and filtered to remove the synoptic variability (see Sect. 2.2.4).

Table D1. Prescribed COS surface fluxes used as model input. Mean magnitudes of different types of fluxes are given for the period 2000–2009.

Type of COS flux	Temporal resolution	Total (Gg S yr^{-1})	Data source
Anthropogenic	Monthly, interannual	337.3	Zumkehr et al. (2018)
Biomass burning	Monthly, interannual	56.3	Stinecipher et al. (2019)
Soil	Monthly, climatological	−409.0	Launois et al. (2015b)
Ocean	Monthly, climatological	344.0	Lennartz et al. (2017)
Vegetation uptake	Monthly, interannual		This work, including mechanistic and LRU approaches (Seibt et al., 2010; Whelan et al., 2018)

D2 DMS sensitivity study

We further tested the impact of the indirect COS fluxes through DMS on the simulated concentrations at NOAA sites. To do that, we compared the atmospheric concentrations given with and without prescribing indirect oceanic fluxes through DMS using the Launois et al. (2015a) oceanic fluxes. In our case, the removal of the DMS oceanic emissions decreases the seasonal amplitude at SPO and CGO but has very few impacts at other sites. We also performed the same experiment using the Lennartz et al. (2017) fluxes and reported no impact of DMS indirect fluxes on simulated concentrations at NOAA sites.

Table D2. Normalised standard deviations (NSDs) of the simulated concentrations by the observed concentrations. Within brackets are the Pearson correlation coefficients (R) between simulated and observed COS concentrations for the mechanistic approach including the DMS or not, calculated between 2004 and 2009 at 10 NOAA stations.

	SPO	CGO	SMO	KUM	MLO	NWR	LEF	MHD	UTK	ALT
ORCHIDEE Mechanist (DMS)	1.10 (0.97)	1.01 (0.97)	0.35 (0.4)	0.90 (0.95)	1.05 (0.92)	1.26 (0.63)	1.34 (0.94)	1.09 (0.85)	0.69 (0.91)	0.64 (0.96)
ORCHIDEE Mechanist (without DMS)	*0.74* (0.91)	*0.53* (0.94)	0.38 (0.20)	0.90 (0.95)	1.04 (0.91)	1.31 (0.64)	1.40 (0.94)	0.93 (0.94)	0.74 (0.90)	0.65 (0.96)

Code availability. The CMIP6 version of the ORCHIDEE model including the COS submodel is available on request to the authors. The LMDz model is available from http://web.lmd.jussieu.fr/LMDZ/LMDZ6/ (last access: 19 April 2021) under the CeCILL v2 Free Software License.

Data availability. For Hyytiälä, we used the 2015 eddy covariance flux data published in Kohonen (2020), the 2015 soil measurements published in Sun et al. (2018b), the 2017 branch chamber and eddy covariance fluxes published in Kooijmans et al. (2018), and local meteorological data available at https://smear.avaa.csc.fi/ (last access: 19 April 2021; Junninen et al., 2009). For Harvard, we used the data published in Commane et al. (2016).

Author contributions. FM and PP devised the research. CA and FM coded the ORCHIDEE developments and made the simulations. MR and PP dealt with the transport model. LMJK and KMK provided the Hyytiälä data. RC and RW provided the Harvard Forest data. JEC, SB, SAM, NR, US, YPS, NV, and MEW were consulted on their respective expertise. FM, CA, and MR analysed the results and wrote the first draft. All authors contributed to the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors thank the reviewers for their constructive and useful comments which helped to further improve this study. The LSCE group thanks the administrative and IT teams for managing the recruitment of Camille Abadie and providing the necessary facilities and tools to run the ORCHIDEE model and analyse the outputs. Operation of the US-Ha1 site is supported by the AmeriFlux Management Project with funding by the U.S. Department of Energy's Office of Science under contract no. DE-AC02-05CH11231 and additionally is a part of the Harvard Forest LTER site supported by the National Science Foundation (DEB-1237491). The authors are very grateful to the ObsPack people who collected and archived the CO_2 surface air samples used in this study.

Financial support. Camille Abadie, Fabienne Maignan, and Philippe Peylin have been mainly supported by the European Commission, Horizon 2020 Framework Programme, 4C (grant no. 821003) and to a smaller extent VERIFY (grant no. 776810). Marine Remaud was funded by the CO_2 Human Emissions (CHE) project which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 776186. Kukka-Maaria Kohonen thanks the Vilho, Yrjö, and Kalle Väisälä Fund and ICOS-FINLAND (319871) for their financial support. Linda M. J. Kooijmans received funding from the ERC project COSOCS under grant no. 742798.

Review statement. This paper was edited by Akihiko Ito and reviewed by Georg Wohlfahrt and one anonymous referee.

References

Allen, M., Babiker, M., Chen, Y., Taylor, M., Tschakert Australia, P., Waisman, H., Warren, R., Zhai, P., Zickfeld, K., Zhai, P., Pörtner, H., Roberts, D., Skea, J., Shukla, P., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J., Chen, Y., Zhou, X., Gomis, M., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T.: IPCCC 1.5C: Summary for Policymakers, Aromar Revi, available at: https://www.ipcc.ch/site/assets/uploads/sites/2/2019/05/SR15_SPM_version_report_LR.pdf (last access: 19 April 2021), 2018.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys., 53, 785–818, https://doi.org/10.1002/2015RG000483, 2015.

Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., Peylin, P., Chevallier, F., Vuichard, N., and Bastrikov, V.: Improving Estimates of Gross Primary Productivity by Assimilating Solar-Induced Fluorescence Satellite Retrievals in a Terrestrial Biosphere Model Using a Process-Based SIF Model, J. Geophys. Res.-Biogeo., 124, 3281–3306, https://doi.org/10.1029/2019JG005040, 2019.

Badger, M. R. and Price, G. D.: The Role of Carbonic Anhydrase in Photosynthesis, Annu. Rev. Plant Physiol. Plant Mol. Biol., 45, 369–392, https://doi.org/10.1146/annurev.pp.45.060194.002101, 1994.

Ball, J. T., Woodrow, I. E., and Berry, J. A.: A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions, in: Progress in Photosynthesis Research, Springer, The Netherlands, 221–224, 1987.

Barkley, M. P., Palmer, P. I., Boone, C. D., Bernath, P. F., and Suntharalingam, P.: Global distributions of carbonyl sulfide in the upper troposphere and stratosphere, Geophys. Res. Lett., 35, L14810, https://doi.org/10.1029/2008GL034270, 2008.

Barnard, D. M. and Bauerle, W. L.: The implications of minimum stomatal conductance on modeling water flux in forest canopies, J. Geophys. Res.-Biogeo., 118, 1322–1333, https://doi.org/10.1002/jgrg.20112, 2013.

Berkelhammer, M., Asaf, D., Still, C., Montzka, S., Noone, D., Gupta, M., Provencal, R., Chen, H., and Yakir, D.: Constraining surface carbon fluxes using in situ measurements of carbonyl sulfide and carbon dioxide, Global Biogeochem. Cy., 28, 161–179, https://doi.org/10.1002/2013GB004644, 2014.

Berkelhammer, M., Alsip, B., Matamala, R., Cook, D., Whelan, M. E., Joo, E., Bernacchi, C., Miller, J., and Meyers, T.: Seasonal Evolution of Canopy Stomatal Conductance for a Prairie and Maize Field in the Midwestern United States from Continuous Carbonyl Sulfide Fluxes, Geophys. Res. Lett., 47, e2019GL085652, https://doi.org/10.1029/2019GL085652, 2020.

Berry, J., Wolf, A., Campbell, J. E., Baker, I., Blake, N., Blake, D., Denning, A. S., Kawa, S. R., Montzka, S. A., Seibt, U., Stimler, K., Yakir, D., and Zhu, Z.: A coupled model of the global cycles of carbonyl sulfide and CO_2: A possible new window on the carbon cycle, J. Geophys. Res.-Biogeo., 118, 842–852, https://doi.org/10.1002/jgrg.20068, 2013.

Bloom, A. A., Exbrayat, J. F., Van Der Velde, I. R., Feng, L., and Williams, M.: The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times, P. Natl. Acad. Sci. USA, 113, 1285–1290, https://doi.org/10.1073/pnas.1515160113, 2016.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P., and Monfray, P.: A global prognostic scheme of leaf onset using satellite data, Global Change Biol., 6, 709–725, https://doi.org/10.1046/j.1365-2486.2000.00362.x, 2000.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J. L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M. A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J. Y., Guenet, B., Guez, L. E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J. B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J. Adv. Model. Earth Syst., 12, e2019MS002010, doi:10.1029/2019MS002010, 2020.

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Buckley, T. N. and Warren, C. R.: The role of mesophyll conductance in the economics of nitrogen and water use in photosynthesis, Photosynth. Res., 119, 77–88, https://doi.org/10.1007/s11120-013-9825-2, 2014.

Caird, M. A., Richards, J. H., and Donovan, L. A.: Nighttime stomatal conductance and transpiration in C_3 and C_4 plants, Plant Physiol., 143, 4–10, https://doi.org/10.1104/pp.106.092940, 2007.

Campbell, J. E., Carmichael, G. R., Chai, T., Mena-Carrasco, M., Tang, Y., Blake, D. R., Blake, N. J., Vay, S. A., Collatz, G. J., Baker, I., Berry, J. A., Montzka, S. A., Sweeney, C., Schnoor, J. L., and Stanier, C. O.: Photosynthetic Control of Atmospheric Carbonyl Sulfide During the Growing Season, Science, 322, 1085–1088, https://doi.org/10.1126/science.1164015, 2008.

Campbell, J. E., Berry, J. A., Seibt, U., Smith, S. J., Montzka, S. A., Launois, T., Belviso, S., Bopp, L., and Laine, M.: Large historical growth in global terrestrial gross primary production, Nature, 544, 84–87, https://doi.org/10.1038/nature22030, 2017.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and other biogeochemical cycles, in: Climate Change 2013, the Physical Science Basis: Working Group I – Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 465–570, https://doi.org/10.1017/CBO9781107415324.015, 2013.

Collatz, G., Ribas-Carbo, M., and Berry, J.: Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C_4 Plants, Funct. Plant Biol., 19, 519–538, https://doi.org/10.1071/pp9920519, 1992.

Commane, R., Meredith, L. K., Baker, I. T., Berry, J. A., Munger, J. W., Montzka, S. A., Templer, P. H., Juice, S. M., Zahniser, M. S., and Wofsy, S. C.: Seasonal fluxes of carbonyl sulfide in a midlatitude forest, P. Natl. Acad. Sci. USA, 112, 14162–14167, https://doi.org/10.1073/pnas.1504131112, 2015.

Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y. W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., Miller, S. M., Parazoo, N. C., Randerson, J. T., Sweeney, C., Tans, P., Thoning, K., Veraverbeke, S., Miller, C. E., and Wofsy, S. C.: Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra, P. Natl. Acad. Sci. USA, 114, 5361–5366, https://doi.org/10.1073/pnas.1618567114, 2017.

Commane, R., Wofsy, S., and Weir, R.: Fluxes of Carbonyl Sulfide at Harvard Forest EMS Tower since 2010 version 4, Environmental Data Initiative [data set], https://doi.org/10.6073/pasta/7ed7b4d1fc7ad303998e76143a3b279a (last access: 19 April 2021), 2016.

Conway, T. J., Tans, P. P., and Waterman, L. S.: Atmospheric CO_2 records from sites in the NOAA/CMDL air sampling network, in: Trends' 93: A Compendium of Data on Global Change, edited by: Boden, T. A., Kaiser, D. P., Sepanski, R. J., and Stoss, F. W., National Laboratory, Oak Ridge, Tennessee, USA, 41–119, 1994.

Cooperative Global Atmospheric Data Integration Project: Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957–2017; obspack_co2_1_GLOBALVIEWplus_v4.0_2018-08-02, NOAA Earth System Research Laboratory, Global Monitoring Division [data set], https://doi.org/10.25925/20180802, 2018.

de Rosnay, P. and Polcher, J.: Modelling root water uptake in a complex land surface scheme coupled to a GCM, Hydrol. Earth Syst. Sci., 2, 239–255, https://doi.org/10.5194/hess-2-239-1998, 1998.

Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5, Clim. Dynam., 40, 2123–2165, https://doi.org/10.1007/s00382-012-1636-1, 2013.

Evans, J. R., von Caemmerer, S., Setchell, B. A., and Hudson, G. S.: The relationship between CO_2 transfer conductance and leaf anatomy in transgenic tobacco with a reduced content of Rubisco, Aust. J. Plant Physiol., 21, 475–495, https://doi.org/10.1071/PP9940475, 1994.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO_2 assim-

ilation in leaves of C_3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Farquhar, G. D., Lloyd, J., Taylor, J. A., Flanagan, L. B., Syvertsen, J. P., Hubick, K. T., Wong, S. C., and Ehleringer, J. R.: Vegetation effects on the isotope composition of oxygen in atmospheric CO_2, Nature, 363, 439–443, https://doi.org/10.1038/363439a0, 1993.

Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry, Atmos. Chem. Phys., 6, 2273–2319, https://doi.org/10.5194/acp-6-2273-2006, 2006.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogée, J., Rödenbeck, C., Running, S. W., Santaren, J. D., Sitch, S., Viovy, N., Ian Woodward, F., and Zaehle, S.: FLUXNET and modelling the global carbon cycle, Global Change Biol., 13, 610–633, https://doi.org/10.1111/j.1365-2486.2006.01223.x, 2007.

Gimeno, T. E., Ogée, J., Royles, J., Gibon, Y., West, J. B., Burlett, R., Jones, S. P., Sauze, J., Wohl, S., Benard, C., Genty, B., and Wingate, L.: Bryophyte gas-exchange dynamics along varying hydration status reveal a significant carbonyl sulphide (COS) sink in the dark and COS source in the light, New Phytol., 215, 965–976, https://doi.org/10.1111/nph.14584, 2017.

Goldan, P. D., Fall, R., Kuster, W. C., and Fehsenfeld, F. C.: Uptake of COS by growing vegetation: A major tropospheric sink, J. Geophys. Res.-Atmos., 93, 14186–14192, https://doi.org/10.1029/JD093iD11p14186, 1988.

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., Welp, L. R., Sweeney, C., Tans, P. P., Kelley, J. J., Daube, B. C., Kort, E. A., Santoni, G. W., and Bent, J. D.: Enhanced seasonal exchange of CO_2 by Northern ecosystems since 1960, Science, 341, 1085–1089, https://doi.org/10.1126/science.1239207, 2013.

Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M. A., Walters, S., Lamarque, J. F., and Holland, E. A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, J. Geophys. Res.-Atmos., 109, D04314, https://doi.org/10.1029/2003JD003957, 2004.

Hauglustaine, D. A., Balkanski, Y., and Schulz, M.: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate, Atmos. Chem. Phys., 14, 11031–11063, https://doi.org/10.5194/acp-14-11031-2014, 2014.

Henry R. P.: Techniques for measuring carbonic anhydrase activity in vitro: the electrometric delta pH and pH stat assay, in: The Carbonic Anhydrases: Cellular Physiology and Molecular Genetics, edited by: Dodgson, S. J., Tashian R. E., Gros, G., and Carter, N. D., New York, Plenum, 11–126, 1991.

Hilton, T. W., Zumkehr, A., Kulkarni, S., Berry, J., Whelan, M. E., and Elliott Campbell, J.: Large variability in ecosystem models explains uncertainty in a critical parameter for quantifying GPP with carbonyl sulphide, Tellus B, 67, 26329, https://doi.org/10.3402/tellusb.v67.26329, 2015.

Hilton, T. W., Whelan, M. E., Zumkehr, A., Kulkarni, S., Berry, J. A., Baker, I. T., Montzka, S. A., Sweeney, C., Miller, B. R., and Elliott Campbell, J.: Peak growing season gross uptake of carbon in North America is largest in the Midwest USA, Nat. Clim. Change, 7, 450–454, https://doi.org/10.1038/nclimate3272, 2017.

Hourdin, F. and Talagrand, O.: Eulerian backtracking of atmospheric tracers, I: Adjoint derivation and parametrization of subgrid-scale transport, Q. J. Roy. Meteor. Soc., 132, 567–583, https://doi.org/10.1256/qj.03.198.A, 2006.

Hourdin, F., Talagrand, O., and Idelkadi, A.: Eulerian backtracking of atmospheric tracers, II: Numerical aspects, Q. J. Roy. Meteor. Soc., 132, 585–603, https://doi.org/10.1256/qj.03.198.B, 2006.

IPCC: Summary for Policymakers, in: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, edited by: Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., and Malley, J., in press., 2019a.

IPCC: Summary for Policymakers, in: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, edited by: Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegri´a, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. M., in press, 2019b.

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y., and Tucker, C.: Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data, Remote Sens.-Basel, 10, 1346, https://doi.org/10.3390/rs10091346, 2018.

Junninen, H., Lauri, A., Keronen, P., Aalto, P., Hiltunen, V., Hari, P., and Kulmala, M.: Smart-SMEAR: on-line data exploration and visualization tool for SMEAR stations, Boreal Environ. Res., 14, 447–457, 2009.

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

Kettle, A. J., Kuhn, U., Von Hobe, M., Kesselmeier, J., and Andreae, M. O.: Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the domi-

nant sources and sinks, J. Geophys. Res.-Atmos., 107, 4658, https://doi.org/10.1029/2002JD002187, 2002.

Kitz, F., Spielmann, F. M., Hammerle, A., Kolle, O., Migliavacca, M., Moreno, G., Ibrom, A., Krasnov, D., Noe, S. M., and Wohlfahrt, G.: Soil COS Exchange: A Comparison of Three European Ecosystems, Global Biogeochem. Cy., 34, e2019GB006202, https://doi.org/10.1029/2019GB006202, 2020.

Kohonen, K.-M.: Data for "Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide", Zenodo [data set], https://doi.org/10.5281/zenodo.3907342 (last access: 19 April 2021), 2020.

Kohonen, K.-M., Kolari, P., Kooijmans, L. M. J., Chen, H., Seibt, U., Sun, W., and Mammarella, I.: Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide, Atmos. Meas. Tech., 13, 3957–3975, https://doi.org/10.5194/amt-13-3957-2020, 2020.

Kooijmans, L. M. J., Uitslag, N. A. M., Zahniser, M. S., Nelson, D. D., Montzka, S. A., and Chen, H.: Continuous and high-precision atmospheric concentration measurements of COS, CO_2, CO and H_2O using a quantum cascade laser spectrometer (QCLS), Atmos. Meas. Tech., 9, 5293–5314, https://doi.org/10.5194/amt-9-5293-2016, 2016.

Kooijmans, L. M. J., Maseyk, K., Seibt, U., Sun, W., Vesala, T., Mammarella, I., Kolari, P., Aalto, J., Franchin, A., Vecchi, R., Valli, G., and Chen, H.: Canopy uptake dominates nighttime carbonyl sulfide fluxes in a boreal forest, Atmos. Chem. Phys., 17, 11453–11465, https://doi.org/10.5194/acp-17-11453-2017, 2017.

Kooijmans, L. M. J., Sun, W., Aalto, J., Erkkilä, K.-M., Maseyk, K., Seibt, U., Vesala, T., Mammarella, I., and Chen, H.: Dataset for "Influences of light and humidity on carbonyl sulfide-based estimates of photosynthesis", Zenodo [data set], https://zenodo.org/record/1211481#.XEC-NVVKhQJ (last access: 19 April 2021), 2018.

Kooijmans, L. M. J., Sun, W., Aalto, J., Erkkilä, K. M., Maseyk, K., Seibt, U., Vesala, T., Mammarella, I., and Chen, H.: Influences of light and humidity on carbonyl sulfide-based estimates of photosynthesis, P. Natl. Acad. Sci. USA, 116, 2470–2475, https://doi.org/10.1073/pnas.1807600116, 2019.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cy., 19, GB1015, https://doi.org/10.1029/2003GB002199, 2005.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9, 3757–3776, https://doi.org/10.5194/bg-9-3757-2012, 2012.

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, Geosci. Model Dev., 7, 2581–2597, https://doi.org/10.5194/gmd-7-2581-2014, 2014.

Landschützer, P., Gruber, N., and Bakker, D. C. E.: A 30 years observation-based global monthly gridded sea surface pCO_2 product from 1982 through 2011 (NCEI Accession 0160558) [Dataset], Version 2.2, NOAA National Centers for Environmental Information, available at: https://doi.org/10.3334/cdiac/otg.spco2_1982_2011_eth_somffn, 2015.

Lardy, R., Bellocchi, G., and Soussana, J.: A new method to determine soil organic carbon equilibrium, Environ. Modell. Softw., 26, 1759–1763, https://doi.org/10.1016/j.envsoft.2011.05.016, 2011.

Launois, T., Belviso, S., Bopp, L., Fichot, C. G., and Peylin, P.: A new model for the global biogeochemical cycle of carbonyl sulfide – Part 1: Assessment of direct marine emissions with an oceanic general circulation and biogeochemistry model, Atmos. Chem. Phys., 15, 2295–2312, https://doi.org/10.5194/acp-15-2295-2015, 2015a.

Launois, T., Peylin, P., Belviso, S., and Poulter, B.: 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, Atmos. Chem. Phys., 15, 9285–9312, https://doi.org/10.5194/acp-15-9285-2015, 2015b.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Lee, E., Zeng, F.-W., Koster, R. D., Weir, B., Ott, L. E., and Poulter, B.: The impact of spatiotemporal variability in atmospheric CO_2 concentration on global terrestrial carbon fluxes, Biogeosciences, 15, 5635–5652, https://doi.org/10.5194/bg-15-5635-2018, 2018.

Lennartz, S. T., Marandino, C. A., von Hobe, M., Cortes, P., Quack, B., Simo, R., Booge, D., Pozzer, A., Steinhoff, T., Arevalo-Martinez, D. L., Kloss, C., Bracher, A., Röttgers, R., Atlas, E., and Krüger, K.: Direct oceanic emissions unlikely to account for the missing source of atmospheric carbonyl sulfide, Atmos. Chem. Phys., 17, 385–402, https://doi.org/10.5194/acp-17-385-2017, 2017.

Lombardozzi, D. L., Zeppel, M. J. B., Fisher, R. A., and Tawfik, A.: Representing nighttime and minimum conductance in CLM4.5: global hydrology and carbon sensitivity analysis using observational constraints, Geosci. Model Dev., 10, 321–331, https://doi.org/10.5194/gmd-10-321-2017, 2017.

Ma, J., Kooijmans, L. M. J., Cho, A., Montzka, S. A., Glatthor, N., Worden, J. R., Kuai, L., Atlas, E. L., and Krol, M. C.: Inverse modelling of carbonyl sulfide: implementation, evaluation and implications for the global budget, Atmos. Chem. Phys., 21, 3507–3529, https://doi.org/10.5194/acp-21-3507-2021, 2021.

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., and Ciais, P.: Using satellite data to improve the leaf phenology of

a global terrestrial biosphere model, Biogeosciences, 12, 7185–7208, https://doi.org/10.5194/bg-12-7185-2015, 2015.

Maseyk, K., Berry, J. A., Billesbach, D., Campbell, J. E., Torn, M. S., Zahniser, M., and Seibt, U.: Sources and sinks of carbonyl sulfide in an agricultural field in the Southern Great Plains, P. Natl. Acad. Sci. USA, 111, 9064–9069, https://doi.org/10.1073/pnas.1319132111, 2014.

Medlyn, B. E., Dreyer, E., Ellsworth, D., Forstreuter, M., Harley, P. C., Kirschbaum, M. U. F., Le Roux, X., Montpied, P., Strassemeyer, J., Walcroft, A., Wang, K., and Loustau, D.: Temperature response of parameters of a biochemically based model of photosynthesis, II. A review of experimental data, Plant Cell Environ., 25, 1167–1179, https://doi.org/10.1046/j.1365-3040.2002.00891.x, 2002.

Montzka, S. A., Calvert, P., Hall, B. D., Elkins, J. W., Conway, T. J., Tans, P. P., and Sweeney, C. S.: On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO_2, J. Geophys. Res.-Atmos., 112, D09302, https://doi.org/10.1029/2006JD007665, 2007.

Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33, 161–174, https://doi.org/10.2307/1269043, 1991.

Norton, A. J., Rayner, P. J., Koffi, E. N., Scholze, M., Silver, J. D., and Wang, Y.-P.: Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model, Biogeosciences, 16, 3069–3093, https://doi.org/10.5194/bg-16-3069-2019, 2019.

Notni, J., Schenk, S., Protoschill-Krebs, G., Kesselmeier, J., and Anders, E.: The Missing Link in COS Metabolism: A Model Study on the Reactivation of Carbonic Anhydrase from its Hydrosulfide Analogue, ChemBioChem, 8, 530–536, https://doi.org/10.1002/cbic.200600436, 2007.

Ogawa, T., Noguchi, K., Saito, M., Nagahata, Y., Kato, H., Ohtaki, A., Nakayama, H., Dohmae, N., Matsushita, Y., Odaka, M., Yohda, M., Nyunoya, H., and Katayama, Y.: Carbonyl sulfide hydrolase from Thiobacillus thioparus strain THI115 is one of the β-carbonic anhydrase family enzymes, J. Am. Chem. Soc., 135, 3818–3825, https://doi.org/10.1021/ja307735e, 2013.

Ogée, J., Sauze, J., Kesselmeier, J., Genty, B., Van Diest, H., Launois, T., and Wingate, L.: A new mechanistic framework to predict OCS fluxes from soils, Biogeosciences, 13, 2221–2240, https://doi.org/10.5194/bg-13-2221-2016, 2016.

Parazoo, N. C., Denning, A. S., Berry, J. A., Wolf, A., Randall, D. A., Kawa, S. R., Pauluis, O., and Doney, S. C.: Moist synoptic transport of CO_2 along the mid-latitude storm track, Geophys. Res. Lett., 38, 2011GL047238, https://doi.org/10.1029/2011GL047238, 2011.

Peters, W., van der Velde, I. R., van Schaik, E., Miller, J. B., Ciais, P., Duarte, H. F., van der Laan-Luijkx, I. T., van der Molen, M. K., Scholze, M., Schaefer, K., Vidale, P. L., Verhoef, A., Wårlind, D., Zhu, D., Tans, P. P., Vaughn, B., and White, J. W. C.: Increased water-use efficiency and reduced CO_2 uptake by plants during droughts at a continental scale, Nat. Geosci., 11, 744–748, https://doi.org/10.1038/s41561-018-0212-7, 2018.

Peylin, P., Ciais, P., Denning, A. S., Tans, P. P., Berry, J. A., and White, J. W. C.: A 3-dimensional study of $\delta^{18}O$ in atmospheric CO_2: Contribution of different land ecosystems, Tellus B, 51, 642–667, https://doi.org/10.3402/tellusb.v51i3.16452, 1999.

Phillips, N. G., Lewis, J. D., Logan, B. A., and Tissue, D. T.: Inter- and intra-specific variation in nocturnal water transport in Eucalyptus, Tree Physiol., 30, 586–596, https://doi.org/10.1093/treephys/tpq009, 2010.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative, Geosci. Model Dev., 8, 2315–2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., and Wright, I. J.: Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology, Ecol. Lett., 17, 82–91, https://doi.org/10.1111/ele.12211, 2014.

Protoschill-Krebs, G., Wilhelm, C., and Kesselmeier, J.: Consumption of carbonyl sulphide (COS) by higher plant carbonic anhydrase (CA), Atmos. Environ., 30, 3151–3156, https://doi.org/10.1016/1352-2310(96)00026-X, 1996.

Rastogi, B., Berkelhammer, M., Wharton, S., Whelan, M. E., Itter, M. S., Leen, J. B., Gupta, M. X., Noone, D., and Still, C. J.: Large Uptake of Atmospheric OCS Observed at a Moist Old Growth Forest: Controls and Implications for Carbon Cycle Applications, J. Geophys. Res.-Biogeo., 123, 3424–3438, https://doi.org/10.1029/2018JG004430, 2018.

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

Remaud, M., Chevallier, F., Cozic, A., Lin, X., and Bousquet, P.: On the impact of recent developments of the LMDz atmospheric general circulation model on the simulation of CO_2 transport, Geosci. Model Dev., 11, 4489–4513, https://doi.org/10.5194/gmd-11-4489-2018, 2018.

Sadok, W. and Jagadish, S. V. K.: The Hidden Costs of Nighttime Warming on Yields, Trends Plant Sci., 25, 644–651, https://doi.org/10.1016/j.tplants.2020.02.003, 2020.

Sandoval-Soto, L., Stanimirov, M., von Hobe, M., Schmitt, V., Valdes, J., Wild, A., and Kesselmeier, J.: Global uptake of carbonyl sulfide (COS) by terrestrial vegetation: Estimates corrected by deposition velocities normalized to the uptake of carbon dioxide (CO_2), Biogeosciences, 2, 125–132, https://doi.org/10.5194/bg-2-125-2005, 2005.

Seibt, U., Kesselmeier, J., Sandoval-Soto, L., Kuhn, U., and Berry, J. A.: A kinetic analysis of leaf uptake of COS and its relation to transpiration, photosynthesis and carbon isotope fractionation, Biogeosciences, 7, 333–341, https://doi.org/10.5194/bg-7-333-2010, 2010.

Serio, C., Masiello, G., Mastro, P., Cersosimo, A., Pasquariello, P., and Venafra, S.: Simultaneous retrieval of OCS, and CO_2 from the IASI shortwave spectral band: assessment of the accuracy of the retrieval products and validation with in situ observations, in: Proceedings Society of Photo-Optical Instrumentation Engineers (SPIE), Remote Sensing of Clouds and the Atmosphere XXV, Online Only, 20 September 2020, 11531, 1153107, https://doi.org/10.1117/12.2573351, 2020.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Sobol, I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul., 55, 271–280, https://doi.org/10.1016/S0378-4754(00)00270-6, 2001.

Stangl, Z. R., Tarvainen, L., Wallin, G., Ubierna, N., Räntfors, M., and Marshall, J. D.: Diurnal variation in mesophyll conductance and its influence on modelled water-use efficiency in a mature boreal Pinus sylvestris stand, Photosynth. Res., 141, 53–63, https://doi.org/10.1007/s11120-019-00645-6, 2019.

Stimler, K., Montzka, S. A., Berry, J. A., Rudich, Y., and Yakir, D.: Relationships between carbonyl sulfide (COS) and CO_2 during leaf gas exchange, New Phytol., 186, 869–878, https://doi.org/10.1111/j.1469-8137.2010.03218.x, 2010.

Stimler, K., Berry, J. A., Montzka, S. A., and Yakir, D.: Association between carbonyl sulfide uptake and 18Δ during gas exchange in C_3 and C_4 leaves, Plant Physiol., 157, 509–517, https://doi.org/10.1104/pp.111.176578, 2011.

Stinecipher, J. R., Cameron-Smith, P. J., Blake, N. J., Kuai, L., Lejeune, B., Mahieu, E., Simpson, I. J., and Campbell, J. E.: Biomass Burning Unlikely to Account for Missing Source of Carbonyl Sulfide, Geophys. Res. Lett., 46, 14912–14920, https://doi.org/10.1029/2019GL085567, 2019.

Sun, W., Maseyk, K., Lett, C., and Seibt, U.: A soil diffusion–reaction model for surface COS flux: COSSM v1, Geosci. Model Dev., 8, 3055–3070, https://doi.org/10.5194/gmd-8-3055-2015, 2015.

Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., Levula, J., Keskinen, H., and Seibt, U.: Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland, Atmos. Chem. Phys., 18, 1363–1378, https://doi.org/10.5194/acp-18-1363-2018, 2018a.

Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., Levula, J., Keskinen, H., and Seibt, U.: Dataset for "Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland", Zenodo [data set], https://zenodo.org/record/322936#.YH7KLZ9xeUk (last access: 19 April 2021), 2018b.

Sun, W., Maseyk, K., Lett, C., and Seibt, U.: Stomatal control of leaf fluxes of carbonyl sulfide and CO_2 in a Typha freshwater marsh, Biogeosciences, 15, 3277–3291, https://doi.org/10.5194/bg-15-3277-2018, 2018c.

Suntharalingam, P., Kettle, A. J., Montzka, S. M., and Jacob, D. J.: Global 3-D model analysis of the seasonal cycle of atmospheric carbonyl sulfide: Implications for terrestrial vegetation uptake, Geophys. Res. Lett., 35, L19801, https://doi.org/10.1029/2008GL034332, 2008.

Tarantola, A.: Inverse problem theory: methods for data fitting and model parameter estimation, Elsevier, Amsterdam, The Netherlands, 1987.

Thoning, K. W., Tans, P. P., and Komhyr, W. D.: Atmospheric carbon dioxide at Mauna Loa Observatory, 2. Analysis of the NOAA GMCC data 1974–1985, J. Geophys. Res.-Atmos., 94, 8549–8565, https://doi.org/10.1029/JD094iD06p08549, 1989.

von Caemmerer, S. and Evans, J. R.: Temperature responses of mesophyll conductance differ greatly between species, Plant. Cell Environ., 38, 629–637, https://doi.org/10.1111/pce.12449, 2015.

Wehr, R., Commane, R., Munger, J. W., McManus, J. B., Nelson, D. D., Zahniser, M. S., Saleska, S. R., and Wofsy, S. C.: Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake, Biogeosciences, 14, 389–401, https://doi.org/10.5194/bg-14-389-2017, 2017.

Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R., Schaefer, K., Jacobson, A. R., Lu, C., Tian, H., Ricciuto, D. M., Cook, R. B., Mao, J., and Shi, X.: The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data, Geosci. Model Dev., 7, 2875–2893, https://doi.org/10.5194/gmd-7-2875-2014, 2014.

Welp, L. R., Keeling, R. F., Meijer, H. A. J., Bollenbacher, A. F., Piper, S. C., Yoshimura, K., Francey, R. J., Allison, C. E., and Wahlen, M.: Interannual variability in the oxygen isotopes of atmospheric CO_2 driven by El Niño, Nature, 477, 579–582, https://doi.org/10.1038/nature10421, 2011.

Whelan, M. E., Min, D. H. and Rhew, R. C.: Salt marsh vegetation as a carbonyl sulfide (COS) source to the atmosphere, Atmos. Environ., 73, 131–137, doi:10.1016/j.atmosenv.2013.02.048, 2013.

Whelan, M. E., Hilton, T. W., Berry, J. A., Berkelhammer, M., Desai, A. R., and Campbell, J. E.: Carbonyl sulfide exchange in soils for better estimates of ecosystem carbon uptake, Atmos. Chem. Phys., 16, 3711–3726, https://doi.org/10.5194/acp-16-3711-2016, 2016.

Whelan, M. E., Lennartz, S. T., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L. M. J., Hilton, T. W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K.-F., Yakir, D., Zumkehr, A., Katayama, Y., Ogée, J., Spielmann, F. M., Kitz, F., Rastogi, B., Kesselmeier, J., Marshall, J., Erkkilä, K.-M., Wingate, L., Meredith, L. K., He, W., Bunk, R., Launois, T., Vesala, T., Schmidt, J. A., Fichot, C. G., Seibt, U., Saleska, S., Saltzman, E. S., Montzka, S. A., Berry, J. A., and Campbell, J. E.: Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles, Biogeosciences, 15, 3625–3657, https://doi.org/10.5194/bg-15-3625-2018, 2018.

Whelan, M. E., Anderegg, L. D. L., Badgley, G., Elliott Campbell, J., Commane, R., Frankenberg, C., Hilton, T. W., Kuai, L., Parazoo, N., Shiga, Y., Wang, Y., and Worden, J.: Scientific communities striving for a common cause, B. Am. Meteorol. Soc., 101, 1537–1543, https://doi.org/10.1175/BAMS-D-19-0306.1, 2020.

Wohlfahrt, G., Brilli, F., Hörtnagl, L., Xu, X., Bingemer, H., Hansel, A., and Loreto, F.: Carbonyl sulfide (COS) as a tracer for canopy photosynthesis, transpiration and stomatal conductance: potential and limitations, Plant. Cell Environ., 35, 657–667, https://doi.org/10.1111/j.1365-3040.2011.02451.x, 2012.

Yang, F., Qubaja, R., Tatarinov, F., Rotenberg, E., and Yakir, D.: Assessing canopy performance using carbonyl sul-

fide measurements, Global Change Biol., 24, 3486–3498, https://doi.org/10.1111/gcb.14145, 2018.

Yin, X. and Struik, P. C.: C_3 and C_4 photosynthesis models: An overview from the perspective of crop modelling, NJAS-Wagen. J. Life Sc., 57, 27–38, https://doi.org/10.1016/j.njas.2009.07.001, 2009.

Zobler, L.: A World Soil File for Global Climate Modelling, in: NASA Technical Memorandum 87802, NASA Goddard Institute for Space Studies, New York, USA, 1986.

Zumkehr, A., Hilton, T. W., Whelan, M., Smith, S., Kuai, L., Worden, J., and Campbell, J. E.: Global gridded anthropogenic emissions inventory of carbonyl sulfide, Atmos. Environ., 183, 11–19, https://doi.org/10.1016/j.atmosenv.2018.03.063, 2018.

9.2 Carbonyl sulfide (COS) emissions in two agroecosystems in central France

PLOS ONE

RESEARCH ARTICLE

Carbonyl sulfide (COS) emissions in two agroecosystems in central France

Sauveur Belviso[1]*, Camille Abadie[1], David Montagne[2], Dalila Hadjar[2], Didier Tropée[3], Laurence Vialettes[1], Victor Kazan[1], Marc Delmotte[1], Fabienne Maignan[1], Marine Remaud[1], Michel Ramonet[1], Morgan Lopez[1], Camille Yver-Kwok[1], Philippe Ciais[1]

1 Laboratoire des Sciences du Climat et de l'Environnement, Université Paris-Saclay, CEA-CNRS-UVSQ, UMR8212, Gif-sur-Yvette, France, **2** UMR ECOSYS, Université Paris-Saclay, INRAE, AgroParisTech, Thiverval-Grignon, France, **3** UMR GQE, Université Paris-Saclay, INRAE-CNRS, AgroParisTech, Gif-sur-Yvette, France

* sauveur.belviso@lsce.ipsl.fr

Citation: Belviso S, Abadie C, Montagne D, Hadjar D, Tropée D, Vialettes L, et al. (2022) Carbonyl sulfide (COS) emissions in two agroecosystems in central France. PLoS ONE 17(12): e0278584. https://doi.org/10.1371/journal.pone.0278584

Editor: Tanvir Shahzad, Government College University Faisalabad, PAKISTAN

Received: February 1, 2022

Accepted: November 19, 2022

Published: December 6, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0278584

Data Availability Statement: All relevant data are within the manuscript and the publicly available external repository available here: https://mycore.core-cloud.net/index.php/s/wUsUpMYrW9FUniz

Abstract

Carbonyl sulfide (COS) fluxes simulated by vegetation and soil component models, both implemented in the ORCHIDEE land surface model, were evaluated against field observations at two agroecosystems in central France. The dynamics of a biogenic process not yet accounted for by this model, i.e., COS emissions from croplands, was examined in the context of three independent and complementary approaches. First, during the growing seasons of 2019 and 2020, monthly variations in the nighttime ratio of vertical mole fraction gradients of COS and carbon dioxide measured between 5 and 180 m height ($Grad_{COS}$/$Grad_{CO2}$), a proxy of the ratio of their respective nocturnal net fluxes, were monitored at a rural tall tower site near Orléans (i.e., a "profile vs. model" approach). Second, field observations of COS nocturnal fluxes, obtained by the Radon Tracer Method (RTM) at a sub-urban site near Paris, were used for that same purpose (i.e., a "RTM vs. model" approach of unaccounted biogenic emissions). This site has observations going back to 2014. Third, during the growing seasons of 2019, 2020 and 2021, horizontal mole fraction gradients of COS were calculated from downwind-upwind surveys of wheat and rapeseed crops as a proxy of their respective exchange rates at the plot scale (i.e., a "crop based" comparative approach). The "profile vs. model" approach suggests that the nocturnal net COS uptake gradually weakens during the peak growing season and recovers from August on. The "RTM vs. model" approach suggests that there exists a biogenic source of COS, the intensity of which culminates in late June early July. Our "crop based" comparative approach demonstrates that rapeseed crops shift from COS uptake to emission in early summer during the late stages of growth (ripening and senescence) while wheat crops uptake capacities lower markedly. Hence, rapeseed appears to be a much larger source of COS than wheat at the plot scale. Nevertheless, compared to current estimates of the largest COS sources (i.e., marine and anthropogenic emissions), agricultural emissions during the late stages of growth are of secondary importance.

(https://doi.org/10.14768/6800b065-dcec-4006-ada5-b5f62a4bb832).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exis.

Introduction

The uptake of atmospheric carbonyl sulfide (COS) through stomata in the plant leaves and subsequent irreversible hydrolysis by enzymatic reaction with carbonic anhydrase ($F_{COSveg.}$) is the largest sink in the global budget of this compound, yet estimates differ by as much as a factor of 5.6 as discussed in a recent review [1]. However, if estimates of $F_{COSveg.}$ based on net primary production (NPP) scaling are excluded, since it is now known that $F_{COSveg.}$ has to be scaled rather to gross primary production (GPP) than NPP, the rest of the estimates based on GPP scaling or mechanistic models vary by only a factor of two [1]. A larger uncertainty is attached to the annual amount of COS exchanged by soil, by as much as a factor of 14 as shown in Table 3 of [2]. However, most recent simulations using mechanistic vegetation and soil COS models of [3–6] implemented in the ORCHIDEE land surface model show that the relative contribution of soils in the global budget is secondary, taking up 19 times less COS from the atmosphere than vegetation on an annual basis (i.e., -30 GgS yr^{-1} vs. -576 GgS yr^{-1}, respectively) [2]. This is because models consider (1) that the uptake capacity of oxic soils can be partly countered by a production mechanism, the seasonality of which is mainly driven by temperature, and (2) that anoxic soils behave as sources of COS [5, 6]. Although laboratory and field observations have shown in rare cases that vascular plants could play also a role in COS production [7–9], the processes by which plants emit this gas have yet to be considered in global modeling studies. COS production has also been reported for mosses [10] and lichens [11]. Any method suitable for assessing plant emissions at different spatial scales would help to resolve uncertainties in the global COS budget and to correct the imbalance between bottom-up estimates of total sources and sinks, the latter generally exceeding the former by hundreds of GgS yr^{-1} before correction [12, 13].

Our goal is to examine the capacity of agricultural ecosystems and of some specific crops in producing COS, which in mid-latitude regions exhibit high productivity and cover very large areas. Moreover, shifts from sink to source have already been reported in such ecosystems [8, 14]. For that purpose, we used field observations of (1) vertical and horizontal concentration gradients of COS and (2) COS fluxes obtained by the Radon Tracer Method (RTM). From the comparison of observed and modelled COS fluxes, we propose an empirical function suitable for inventorying crop emissions at different spatial scales.

Materials and methods

No specific permissions were required for these locations/activities. I confirm that the field studies did not involve endangered or protected species.

Experimental sites / applied flux quantification methods

Atmospheric boundary layer mixing ratios were monitored at two sites, Gif-sur-Yvette (GIF) and Trainou (TRN tall tower), both in central France (Fig 1).

TRN site: Nocturnal vertical gradients

The Trainou 180-m tall tower atmospheric observatory is located about 80 km south of GIF (Fig 1). A description of the TRN site is available online at https://icos-atc.lsce.ipsl.fr/panelboard/TRN. At the TRN tall tower, atmospheric COS was measured with the Aerodyne Research quantum cascade laser (mini-QCL) formerly deployed at the SAC station [14]. The sampling lines, made of Synflex tubing, collected air at 4 heights (5 m, 50 m, 100 m and 180 m) sequentially, the total sequence lasting for 80 minutes. Because the TRN station is operated in the framework of ICOS European Infrastructure [15], the mini-QCL was synchronized with

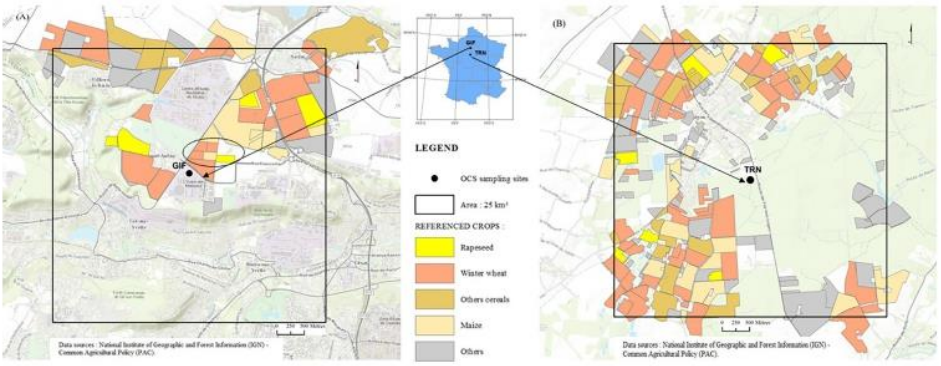

Fig 1. Maps showing the location of the GIF/TRN sampling sites and Common Agricultural Policy (CAP) declared crops (year 2019). The black squares delineate an area of 25 km^2 in the immediate proximity of both stations. The ellipse on the left map delineates the agricultural fields where the uptake/emission of COS has been surveyed for 3 consecutive years. The winter rapeseed (WR) plot area was 4.5 ha in 2019. In 2020 and 2021, WR was grown at the maize (MI) plot. The winter wheat (WW) plots sampled in 2020 and 2021 are those located NE and W of the MI field, respectively. Both have the same size, 3.5 ha each. This figure is similar but not identical to the original image and is therefore for illustrative purposes only.

https://doi.org/10.1371/journal.pone.0278584.g001

an ICOS analyzer measuring CO_2, CH_4, CO and H_2O (Picarro Model G2401). A proxy of the ratio of nocturnal ecosystem fluxes (F_{COS}/F_{CO2}) was calculated from nighttime vertical gradients of COS and CO_2 ($Grad_{COS}/Grad_{CO2}$ from data collected at 5 m and 180 m height), assuming similarity of COS fluxes and gradients to CO_2 fluxes and gradients. Data used to calculate the linear regression slope between $Grad_{COS}$ and $Grad_{CO2}$, were selected according to time of day as follows: [21; 4] (UTC) in spring and summer since we only aimed at examining the capacity of local agricultural ecosystems in exchanging COS with the atmosphere. The linear regression slope between $Grad_{COS}$ and $Grad_{CO2}$ is forced through zero because vertical gradients tend to zero during windy nights. The relative vertical gradients of COS measured at TRN between 5 and 180 m can be used to assess local exchange rates provided that the highest inlet height remains within the nocturnal boundary layer (NBL) throughout the night. This necessary condition was achieved during the months of May through August of 2019 and 2020 as shown later from an analysis of the diurnal variations in monthly mean CO_2 profiles. Emission and uptake rates of COS from nearby crops were not investigated from horizontal mole fraction gradients there.

GIF site: RTM method and horizontal concentration gradients

The GIF monitoring station is located at LSCE, l'Orme des Merisiers, about 2 km south of the Saclay station (SAC) part of the Integrated Carbon Observation System (ICOS) network (https://icos-atc.lsce.ipsl.fr/panelboard/SAC), and next to the village of Saint Aubin. For more information about land cover classification, see Fig 1 of [14] and the EU Common Agricultural Policy (CAP) maps shown in Fig 1.

At GIF, atmospheric COS was measured at 7 m, on line every hour using an automated sampling system (Entech P7100) coupled to a gas chromatograph (GC, Varian 3800) [14, 16]. The GIF time series currently covers seven years (from August 2014 to September 2021). COS fluxes were computed with RTM as by [14], an approach assuming comparably homogeneous spatial distributions of the soil ^{222}Rn source and the ecosystem COS sink/source. COS fluxes are computed as the product of $COS/^{222}Rn$ slopes during nighttime inversion and ^{222}Rn local exhalation rates. The latter (i.e., 52 Bq m^{-2} h^{-1} with a seasonal cycle amplitude of +25% in summer and -25% in winter driven by soil moisture) was taken from [17]. The RTM allowed us to

make 400 determinations of nocturnal fluxes of COS throughout the 2014–2021 period as detailed later.

The small ellipse on the left panel of Fig 1 delineates the fields from which COS fluxes were assessed indirectly from downwind minus upwind horizontal concentration gradients along the wind direction (WD). First, wind speed (WS) and WD data from a weather station located on the LSCE roof top (Vaisala model WXT530) was used to approximate local meteorology. In the field, we used the facilities of a local meteorological station belonging to the Institut National de la Recherche Agronomique (INRA), equipped with a wind vane providing a continuous visualization at the local scale of the WD. WS data from the INRA weather station was downloaded later. Horizontal COS concentration gradients were determined discontinuously in spring-summer of 2019, 2020 and 2021. They were documented between 9:00 and 11:00 AM (local time), during non-rainy days, from flask-air samples collected in pairs each upwind and downwind ($2 < WS < 14$ km h^{-1}, mean = 6 km h^{-1}, SD = 3.6 km h^{-1}) of selected crops fields surrounded by the ellipse in Fig 1A. Although the total duration of the crop surveys was 2 hr, each horizontal gradient was documented within 40 minutes. We ensured that the concentration gradient was not affected by the residual nocturnal boundary layer by comparing samples collected on an hourly basis at 7 m (GIF time series) with upwind flask samples collected at canopy height. The sampling device has been described by [18]. Once filled with sample air, flasks were transferred within hours to the laboratory nearby and analyzed for COS with the Entech/Varian instruments described below. Precision for GC measurements of flask-air samples is reported as box plots in Fig 2.

During May to July of 2021, 29 pairs of flasks were collected and each flask was analyzed twice consecutively for COS. The average difference between duplicated analyses was 9.3 ± 7.2 ppt (median = 7.5 ppt, interquartile range (IQR) = 4.5–12.5 ppt, n = 58). The average difference between flasks of the same pair, assessed as the median of flask 1 minus that of flask 2, was 5.5 ± 4.5 ppt (median = 4.9 ppt, IQR = 2.2–8.0 ppt, n = 29). In 90% of the analyses, the difference between flasks of the same pair was less than 14 ppt.

Our study of uptake/emission from the rapeseed and wheat fields covers a measurement period from March 30th to July 29th, 2021, encompassing most important growth stages, i.e., stem elongation/extension, flowering, ripening and senescence. Winter rapeseed (*Brassica napus*, Blackbuzz variety) was sown on September 5th, 2020 and harvested August 6th, 2021. A mixture of four varieties of winter wheat (*Triticum aestivum*; Renan, Gwenn, LG Absalon and Chevignon varieties) was sown on October 31st, 2020 and harvested August 14th, 2021. Agricultural treatments, including herbicides, insecticides and nitrogen fertilizers, were applied to each crop in autumn and winter. The wheat field was fertilized with sulfur in early March 2021. Wheat and rapeseed yields were 6.8 and 3.7 t ha^{-1} in 2021, respectively. COS exchanges could not be investigated in 2020 as in 2021 because of the French COVID-19 lockdown. All the cultivated plots sampled in our study area, which are less than 500 meters apart (Fig 1A), show very similar soils and the rotation of cultivars is a general feature of agricultural practices in the GIF area.

Simulations of ecosystem fluxes at the GIF and TRN sites

The ORCHIDEE Land Surface Model is developed at the Institut Pierre Simon Laplace (IPSL). It computes the carbon, water and energy balances over land surfaces [19]. Biome types are grouped into 15 Plant Functional Types (PFTs), including bare soil. At the GIF and TRN sites, the vegetation distribution was prescribed based on the respective land cover map and the soil texture is defined as silt loam in texture classification of the United States Department of Agriculture (USDA) [14].

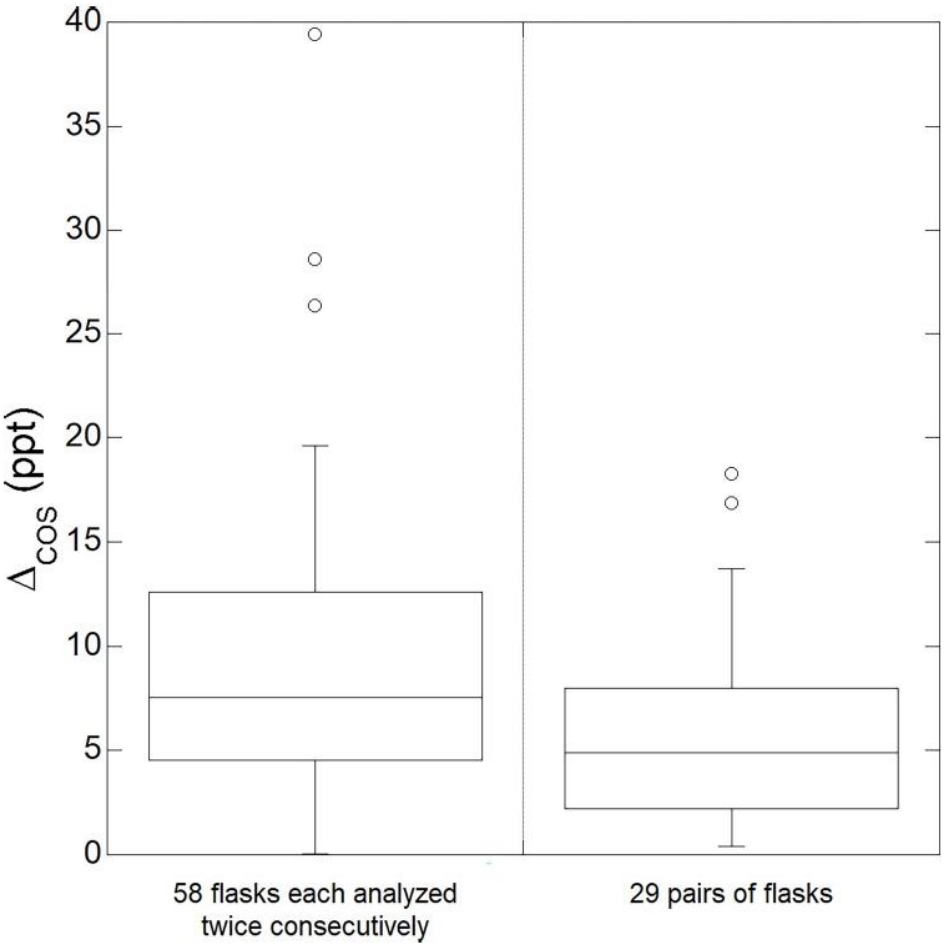

Fig 2. Analysis precision for COS measurements from flask-air samples collected upwind and downwind of selected crops. Left box plot: difference in COS (ppt dry air) between two consecutive analyses of each flask content (n = 58). Right box plot: difference in COS (ppt dry air) between flasks of the same pair (n = 29). Shown are10th, 25th, median, 75th and 90th percentile. Circles correspond to outliers. The study period extends from late March to late July 2021.

https://doi.org/10.1371/journal.pone.0278584.g002

ORCHIDEE was forced by 0.25˚x0.25˚ hourly reanalysis fields from the fifth generation of meteorological analysis of the European Centre for Medium-Range Weather Forecast (ECMWF ERA5, [20]). Near-surface COS concentrations were prescribed using monthly averages of tower atmospheric concentration measurements at these sites for the available years, or using simulated monthly average concentrations by the Laboratoire de Météorologie Dynamique atmospheric transport model (LMDz, [13]). Near-surface CO_2 concentrations are estimated using global annual-mean values provided by the TRENDY project [21].

A "spin-up" phase was first performed for each site, by cycling over the available forcing years for a total of 340 years. This enables all carbon pools to stabilize and the net ecosystem production to oscillate around zero [22]. Then, a transient phase of 40 years was run to introduce anthropogenic disturbances. Finally, simulations of soil and vegetation COS and CO_2 fluxes were run until 2020.

Nocturnal ecosystem COS fluxes take into account plant COS uptake and soil COS uptake and production, computed with mechanistic-based models recently implemented in ORCHIDEE (see [2, 4] for a detailed description of the vegetation and soil models and Table A3 of [4]

where the nocturnal stomatal conductances adopted in ORCHIDEE tend to lower the vegetation uptake of COS at night compared to other approaches).

Results and discussion

A profile vs. model approach of COS nocturnal exchange rates at TRN

The diurnal and seasonal variability of the atmospheric boundary layer (ABL) depth (z_i) at the TRN site has been extensively investigated by [23] from lidar measurements, vertical profiles of CO_2 at 5, 50, 100 and 180 m along the 207 m tall tower, and ^{222}Rn measurements at 180 m. The authors of [23] showed that, in 33.9% of cases, z_i was below the top of the tower during the summer (JJA) nights of 2011 ([23], cf. Fig 5C). As shown in S1 Fig, during the months of May through August of 2019 and 2020, the highest inlet height remained within the NBL during the entire night because the nocturnal CO_2 monthly means at 180 m were 1.7 to 7.7 ppm (in 2019) and 2.7 to 6.2 ppm (in 2020) higher during the night (black curves, strong stratification) than during the day (blue curves, strong vertical mixing). Thus (1) the 180 m inlet did not sample air from the residual layer (RL) by night, in support of Pal et al. (2015)'s scenario B ([23], cf. Fig A1), and (2) the concentration gradients between the two levels indicate a single land use type, i.e. that they are driven by local processes.

Data collected in June-July 2020 at the TRN site (S2 Fig), although after the 2019 survey (Fig 3), should be taken as a preliminary illustration of the approach aiming at assessing monthly changes in net nocturnal exchange rates from the survey of relative vertical gradients of COS (GradCOS / $GradCO_2$). Indeed, we were unable to fully apply that approach during summer in 2020 as we did in 2019 because the IR-laser of our mini-QCL failed in August 2020. Moreover, data selection is applied to the 2020 record (S2 Fig) whereas temporal variations in the slope of linear regressions (ppt/ppm) are reported strictly on a monthly basis (Fig 3A) as when these are computed by the ORCHIDEE model (Fig 3B).

Whereas ecosystem respiration accounts for the nighttime build-up of CO_2 near the ground in agroecosystems during the growing season, net fluxes of COS behave in the opposite way as shown in panels A and B of S2 Fig. During the 3rd week of June 2020, nocturnal COS losses were also observed at 50 m and, in a lesser extent, at 100 m (data not shown). Within the nocturnal boundary layer (see z_i discussion above), 180 m is the sampling height up to which the nocturnal COS drawdown almost never propagates. That is why reference air is taken at 180 m. The average nocturnal COS gradient between 5 and 180 m is generally negative whereas that of CO_2 is always positive (S2C Fig). When plotted against one another, the slopes of the linear regressions forced through zero exhibit an abrupt change during the second half of June 2020, from -1.17 ppt/ppm to -0.40 ppt/ppm. This finding suggests that the efficiency of the ecosystem to absorb COS from the atmosphere by night declines as the growing season proceeds. However, 2019 data shows that the latter recovers in August (Fig 3A, slope = -0.55 ppt/ppm, to be compared with the July one equal to -0.21 ppt/ppm). Nocturnal ecosystem fluxes of COS and CO_2 simulated by the ORCHIDEE model are strongly correlated throughout the growing season, except during a heatwave which lasted a few days in July 2019 and resulted in enhanced soil production of COS (Fig 3B). This strong correlation describes a proportionality between the CO_2 and COS fluxes that remains constant over all summer months, as opposite to what is found for the concentration gradients (Fig 3A) and illustrated by a different linear regression each month. If a decrease in ecosystem efficiency to absorb COS was simulated until July, followed by a recovery in August, we would not expect a single linear regression to be able to fit the simulated COS versus CO_2 fluxes. Hence, this simulated COS exchange, i.e., summed changes in nocturnal stomatal conductance and soil fluxes, failed to reproduce the

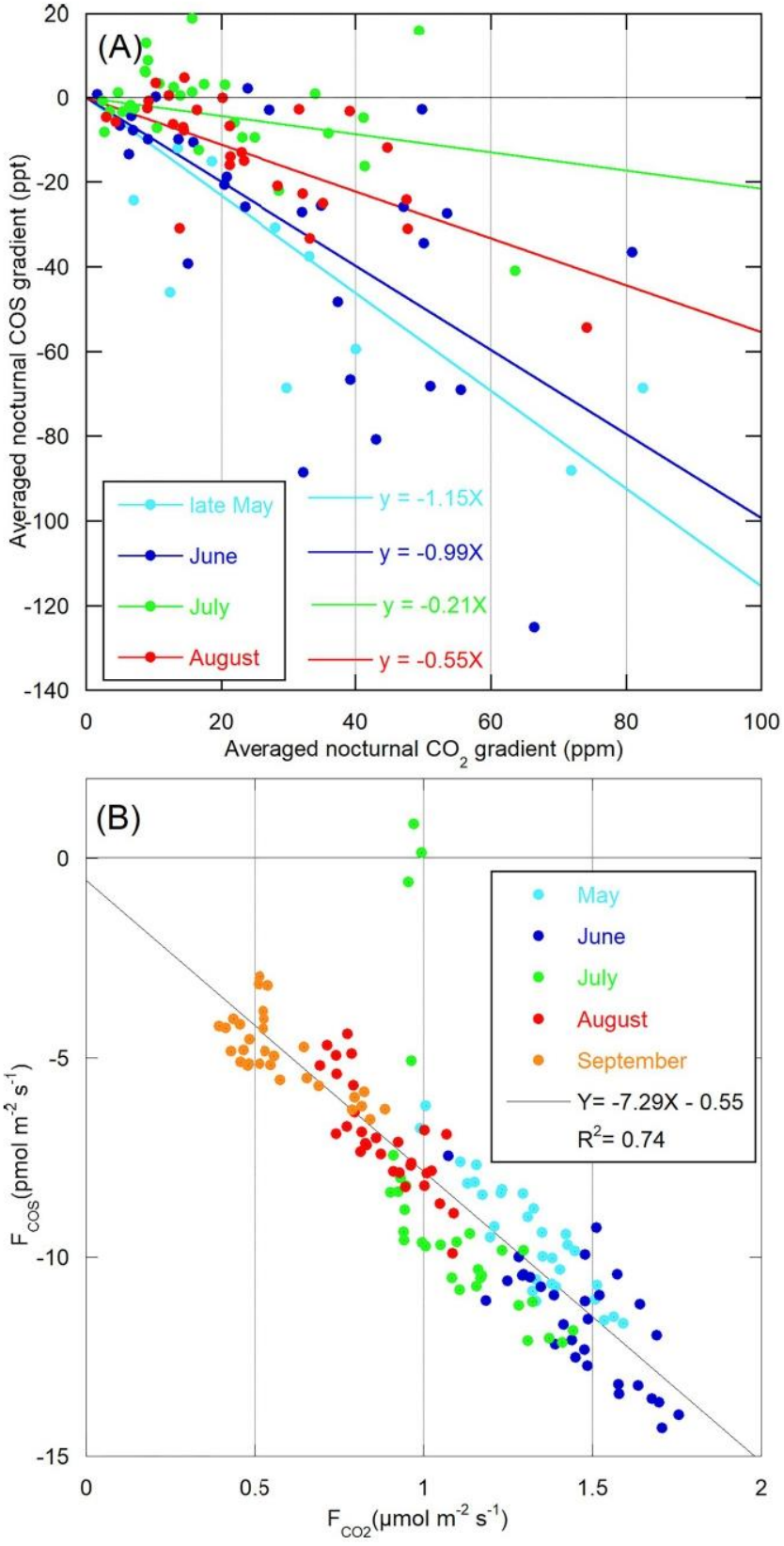
(A)
Averaged nocturnal COS gradient (ppt)
Averaged nocturnal CO_2 gradient (ppm)
late May
June
July
August
y = -1.15X
y = -0.99X
y = -0.21X
y = -0.55X
(B)
F_{COS}(pmol m^{-2} s^{-1})
F_{CO2}(μmol m^{-2} s^{-1})
May
June
July
August
September
Y= -7.29X - 0.55
R^2= 0.74

Fig 3. Relative changes in observed COS vertical gradients and simulated fluxes at the TRN site (2019). The nocturnal observed vertical gradients between 5 and 180m (A; $\Delta COS_{5m\text{-}180m}$) and simulated fluxes of COS (B) are plotted respectively against those of CO_2. Months are color coded.

https://doi.org/10.1371/journal.pone.0278584.g003

decrease in the net sink of COS at TRN during the spring-summer 2019 and its recovery at the end of summer.

Multi-year variations of COS mixing ratio and exchange rates

The GIF time series presently includes 45,000 hourly measurements of atmospheric COS mixing ratio (S3 Fig). It exhibits seasonal variations and a long-term decreasing trend as elsewhere in the northern hemisphere (e.g., https://gml.noaa.gov/dv/iadv/graph.php?code=MLO&program=hats&type=ts, [24]). Indeed, during spring both in 2020 and 2021, COS rarely exceeded 500 ppt at the GIF station, a feature corroborated by data collected at 5 m agl at the TRN site where the mini-QCL remained operative during the French COVID-19 lockdown (S4 Fig).

A plot of the updated time series from [14] of nocturnal fluxes computed with RTM at GIF site is provided in S5 Fig. After being updated, the general features put forward in [14] remain essentially the same. In 365 out of 400 cases, the site is a net sink with a median value of -5.8 pmol m^{-2} s^{-1}. The median emission of the 35 nocturnal events recorded during about 6.5 years is 9.8 pmol m^{-2} s^{-1}. More precisely, emission episodes were systematically recorded during the months of May through July over six consecutive years. Such emission events represent about three quarters of all emission episodes. Model simulations have been used to assess their origin.

A RTM vs. model approach of COS nocturnal exchange rates

At GIF site, the temporal variations in nocturnal ecosystem exchange rates of COS and CO_2 simulated by ORCHIDEE have opposite signs (Fig 4A).

The difference between RTM fluxes and simulations is quantified in Fig 4B. Here, as we aim at finding an easy-to-use empirical function to describe this difference, May to July data shown in Fig 4C has been fit with a polynomial function of the following form applied from May 1st,

$$Y = \frac{A * \alpha^2}{\alpha^2 + (x - x_0)^2}$$

with coefficients x_0, A and α equal to 63.1 ± 0.6, 22.7 ± 0.6 and 15.3 ± 0.4, respectively. This implies that simulated ecosystem exchanges of COS at the GIF site are unable to account for observations. Furthermore, it reveals the existence of a missing source, the intensity of which culminates in late June early July. Its origin is elucidated from investigations of agricultural crops as shown below.

A comparative study of COS exchange by wheat and rapeseed

The direction and magnitude of daytime COS exchange by wheat and rapeseed fields have been assessed in 2020 and 2021 from downwind-upwind differences in COS concentration measured at the top of the canopy ($\Delta COS_{downwind\text{-}upwind}$, S6 Fig). A first attempt carried out in 2019 served as a test of the methodological approach and indicated that rapeseed was a potential source of COS (S6 Fig). In 2021, the rapeseed and wheat fields shifted from net uptake to net emission at DOY 155 and DOY 180, respectively (S6 Fig). The significance of the

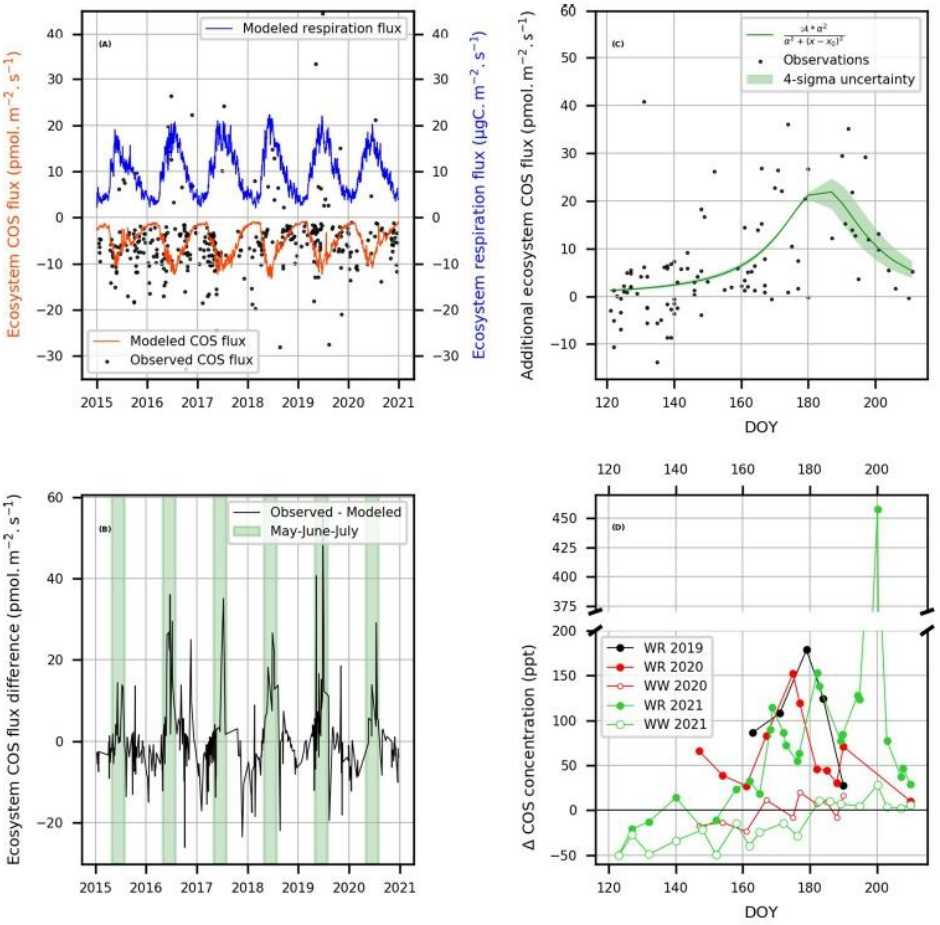

Fig 4. Comparison between observed and simulated nighttime fluxes at the GIF site. (A) Comparison of multi-year variations in simulated nocturnal ecosystem exchange rates of COS (red line) and CO_2 (blue line) with COS fluxes estimated by RTM (black dots, 2014–2020). (B) Difference between "observed" and simulated COS fluxes over 6 years. The period May to July is depicted with a green band. (C) Selected (May to July) data gathered from panel (B) then fitted with a polynomial function of the type depicted in this panel which estimates the dynamics of an additional, yet unidentified, source of COS. Fit: x0 = 63.1 ± 0.6, A = 22.7 ± 0.6, alpha = 15.3 ± 0.4, R^2 = 0.33. The light green shading corresponds to 4-sigma uncertainty. (D) Comparison of the dynamics of COS exchange between crops (WR: winter rapeseed; WW: winter wheat) and the atmosphere, assessed indirectly from horizontal concentration gradients (ΔCOS) downwind and upwind of selected plots (Fig 1). The full record including data collected between late March and late April are displayed in S6 Fig. In 2021, the survey of crops was interrupted before harvest, whereas in 2020, the last sample was collected after harvest. Rapeseed growth stages–year 2021: day of year (DOY) >110–140, flowering; >140–160, development of fruit; >160–190, ripening; >190–210, senescence.

https://doi.org/10.1371/journal.pone.0278584.g004

differences between COS enhancements over the wheat and rapeseed fields was assessed by running a Wilcoxon rank sign test for paired data. After DOY 180, when both crops became net sources of COS, the enhancements downwind the rapeseed and wheat plots were equal to 132 ± 125 ppt (1 SD) and 9.1 ± 8.2 ppt, respectively (P = 0.016). In 2020, when the wheat plot shifted from net sink to net source by DOY 165, the rapeseed plot was already producing COS for at least 15 days. After DOY 165 of 2020, when both crops were sources of COS of contrasted importance, the enhancements over the rapeseed and wheat plots were equal to 78 ± 44 ppt and 5.4 ± 11.2 ppt, respectively (P = 0.016). Moreover, the lowest COS enhancement of the 2020 survey of the rapeseed field was measured after harvest (i.e., 10 ppt vs. 25–150 ppt during the ripening and senescence). These observations are of crucial importance

because they very likely tell that the COS source is in the plant not in the soil. Hence, it appears that the net production of COS by the rapeseed field lasted 15–25 day longer and its intensity was about 14 times higher than that of the wheat plot. The shift of wheat from net uptake to net production occurred about 2 weeks later in 2021 than in 2020, probably because crops experienced a summer drought in 2020 and wetter conditions in 2021 as shown by the soil wetness index for the Ile-de-France region available in French at https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=129&id_rubrique=29 (S7 Fig). For rapeseed crops, it appears that COS production in 2020 was delayed by at least one week. During the early growth stages (DOY $\leq$ 140), when both plants had leaves, horizontal gradients were negative indicating that COS was taken up from the atmosphere (S6 Fig). However, the difference between medians (wheat = -28.6 ppt; rapeseed = -12.9 ppt) suggests that rapeseed plots are smaller net sinks of COS than wheat plots (P = 0.027). From the comparison of Fig 4C and 4D, it appears that COS production from rapeseed crops in the ripening and senescence phases of growth is the biogenic process that is yet unaccounted for by the vegetation and soil models implemented in the ORCHIDEE land surface model, because the temporal variations of the additional ecosystem COS flux (Fig 4C) and emissions from rapeseed plots (Fig 4D) are synchronous at the GIF site.

Maseyk et al. [8] were the first to report that wheat plots in Oklahoma in May-June 2012 act as a sink of COS during the early stages of growth (green plants), then shifted to COS release during the ripening and senescence phases. However, we are unaware of any comparison between simulated and observed fluxes in north-central Oklahoma, where Maseyk et al. [8] documented ecosystem and soil fluxes of COS. Hence, it is not known whether discrepancies exist between simulated and observed fluxes that could arise from poorly calibrated plant physiological or soil parameters rather than from COS emissions from senescent plants yet unaccounted for by models. The difference between net ecosystem COS fluxes measured by eddy covariance (EC) and soil fluxes calculated at below-canopy soil temperature and water content has been tentatively estimated, from a visual inspection of their Fig 1C, to be in the range 10–15 pmol m^{-2} s^{-1} during the whole senescence phase (DOY 135–145). Moreover, the authors did not report on the existence of diurnal variations in COS emission from wheat plants during the senescence phase (i.e., before harvest). Our field observations support qualitatively their results. However, here we provide evidence that the contribution at plot scale of rapeseed releasing COS during the ripening and senescence phases is much more important than that of wheat because rapeseed COS emission is stronger and lasts longer.

Evidence of COS production by rapeseed has been provided in the past from greenhouse experiments where plants in the early stage of growth (stem elongation and flowering) were exposed to unrealistic low levels of atmospheric COS (i.e., < 150 ppt, [25]) or to fungal infections [7]. Here COS production has been documented in the field from late March to late July, along each stage of growth during which farmers did not have to apply specific biochemical treatments against fungal infections. That is, our surveys of agricultural crops were carried out in conditions not perturbed by fungal infections.

A previous comparative study of biogenic volatile organic compounds (VOC) fluxes by wheat and rapeseed plants, with surveys near Paris, although at slightly different growth stages than ours (ripening and fruit development), showed that the total net VOC flux of rapeseed was about 6 times higher than that of wheat [26]. Moreover, VOC emissions of winter wheat increased twofold during the senescence stage compared to the maturation stage [27]. This is qualitatively consistent with Maseyk et al. (2014)'s observations over wheat fields [8] and ours over rapeseed crops. Although those datasets suggest rapeseed to be a stronger emitter of volatile compounds than wheat, comparison with Maseyk et al. [8] suggest that rapeseed and wheat emit COS during the senescence phase at about the same rates. However, both

approaches have estimated COS emission rates by senescent plants indirectly, i.e., from the difference between EC and soil flux measurements [8] and from the difference between fluxes estimated by the RTM method and simulated using the ORCHIDEE model (this work). There is a lack of comparative studies of COS fluxes by wheat, rapeseed and potentially by other plants using dynamic chambers as for VOC fluxes [26]. Maize would deserve special attention because it has been suggested that emission of COS from a largely senesced maize field in Bondville (USA) can disguise any leaf uptake of this gas at the end of the season [28]. The COS emissions at Bondville can account for 25% or so of the total flux (Mary Whelan, pers. com. dated October 2022).

Implications for the global budget of COS

The COS production path in plants is not completely known. COS is thought to be produced from isothiocyanates and thiocyanate ions (SCN-) which are secondary products of the action of the myrosinase enzyme onto a large variety of forms of glucosinolates, which themselves are naturally biosynthesized in plants of the "cabbage order", Brassicales [29]. These authors also provided a list of about sixty Brassicaceae sharing the potential to form isothiocyanates. Hence, rapeseed may be only one among many Brassicaceae sharing the potential to form COS during the late stages of growth. COS is also produced from wheat [8] although it lacks the capacity to produce glucosinolates and myrosinase [30].

The empirical production function from the comparison of observed and modelled COS non-photosynthetic fluxes (Fig 4C) is suitable for inventorying COS emissions from rapeseed crops at different spatial scales. However, because its generality has not been tested yet, it should be used with caution. Indeed, we don't know if the temporal features described by this equation apply to other rapeseed fields. The magnitude of the peak emission, the date it occurs, and how quickly it wanes may vary from site to site. Nevertheless, rapeseed is mainly grown in Canada, China, the European Union, India and Australia. The total harvested areas represented about 32 million ha in 2015 [31]. Assuming that COS emissions remain invariant all day long, our empirical function, integrated over 3 months, yields a yearly total emission of 0.84 ± 0.13 GgS yr^{-1} which, as such, does not make a significant contribution to the global budget of COS because the net uptake of COS by soils and vegetation is estimated by ORCHIDEE to be 606 GgS yr^{-1} [2]. Wheat total harvested areas represent about 215 million ha (https://wheat.org/wheat-in-the-wo4rld/). Assuming that rapeseed and wheat share the same capacity of COS production during their late stages of growth, a yearly total emission of 4.7 ± 0.7 GgS yr^{-1} is obtained. If all C3 crops would share that capacity, a weak assumption, a yearly total emission of 33.3 ± 5.1 GgS yr^{-1} is obtained.

Conclusions

Nocturnal enhancements of atmospheric COS observed each year during springtime and early summer in the GIF area originate from agricultural crops of wheat and rapeseed which shift from uptake to release during ripening and senescence. At the plot scale, COS concentration enhancements from rapeseed largely surpass those of wheat. At the ecosystem scale, COS emissions from agricultural crops of wheat and rapeseed either partly compensate the net nocturnal uptake of COS by vegetation and soil, as observed indirectly from measurements of COS and CO_2 vertical gradients at the TRN site, or largely surpass it as the RTM and model joint approach shows. A COS empirical production function is proposed which generality needs to be tested in other rapeseed fields especially in terms of duration and magnitude of the peak emission. The role of other plants of the "cabbage order" needs also to be addressed experimentally in the fields with dynamic chambers or by eddy covariance. Although our field

observations support qualitatively those of [8], some inconsistencies remain as to the relative importance of COS emissions from senescent wheat at the plot scale on both sides of the Atlantic. For now, it is suggested that emissions from rapeseed alone and C3 crops in general cannot account for the missing source of COS in the global budget of this gas.

Supporting information

S1 Fig. CO_2 vertical gradient, per month and per daily period at the TRN site for the years 2019 (upper plot) and 2020 (lower plot). These are standard ICOS products generated by the ICOS database.
(PDF)

S2 Fig. Variations in mixing ratios and in relative COS nocturnal vertical gradient. (A) CO_2 and (B) COS (OCS) mixing ratios measured at 5 and 180 m roughly on an hourly basis at the TRN site in June-July 2020. (C) Correlation of COS (OCS) and CO_2 averaged nocturnal vertical gradients measured on a daily basis. Those linear regressions forced through zero are not calculated strictly on a monthly basis but after data selection because the transition from one regime (slope = -1.17 ppt/ppm) to another (slope = -0.4 ppt/ppm) took place the night of June 21st to 22nd as shown in panels A and B.
(PDF)

S3 Fig. Multi-year variations in COS mixing ratio at GIF with hourly resolution. The data gaps in summer/early autumn of 2017 and spring 2020 being a failure of the Entech preconcentrator and the consequence of the French lockdown, respectively. The full COS records are now available from https://doi.org/10.14768/6800b065-dcec-4006-ada5-b5f62a4bb832.
(PDF)

S4 Fig. Hourly variations in COS mixing ratio at the GIF (7m agl) and TRN (5m agl) sites. The mini-QCL remained operative at TRN during the French lockdown while GC measurements at GIF were stopped for about two months.
(PDF)

S5 Fig. Multi-year variations in COS exchange rates at the GIF site. These are nocturnal COS (OCS) fluxes obtained by the Radon Tracer Method.
(PDF)

S6 Fig. Variations during the growing season in uptake/emission regimes by wheat and rapeseed assessed from horizontal gradients of COS. The difference in COS concentrations measured downwind and upwind of selected plots is plotted against day of year (DOY). Measurements were carried out in the morning between 9:00 and 11:00 (local time), during no rainy days and roughly in similar meteorological conditions according to wind speed ($2 < WS < 14$ km h^{-1}, mean = 6 km h^{-1}, SD = 3.6 km h^{-1}). In 2021, the survey of crops was interrupted before harvest, whereas in 2020, the last sample was collected after harvest. The lag in 2021 of the shift from net uptake to net production (upt-to-prod) for either WW or WR is depicted by an horizontal double arrow. Rapeseed growth stages—year 2021: DOY<110, inflorescence emergence and elongation; >110–140, flowering; >140–160, development of fruit; >160–190, ripening; >190–210, senescence. We have zoomed in the May-to-July period in Fig 4D.
(PDF)

S7 Fig. Soil wetness index (SWI) for the Ile-de-France region. The SWI is a soil moisture index documented in the scientific literature. It represents, over a depth of about two meters, the state of the water reserve of the soil in relation to the useful reserve (water available for

plant nutrition). Plots downloaded from https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=129&id_rubrique=29. First, we selected Bulletin climatique mensuel régional (à partir de janvier 2020), then Ile-de-France region from the drop-down menu, then we downloaded reports for the months of July 2020 and July 2021, then compared graphs entitled "Indice d'humidité des sols" in page 4 of 5). Upper panel: March 1st to July 31st, 2020. Lower panel: March 1st to July 31st, 2021. Refer only to the purple curves.
(PDF)

Acknowledgments

SB expresses its special thanks to Mark Zahniser at Aerodyne Research for its unconditional and enthusiastic support during operation of the mini-QCL. We thank Nicolas Vuichard and Tanguy Martinez for preparing the ERA5 forcing files for the ORCHIDEE model and the ICOS data products displayed in the supplements, respectively. We wish to thank the two reviewers, including Mary Whelan, for their helpful suggestions to improve the paper.

Author Contributions

Conceptualization: Sauveur Belviso, Camille Abadie, David Montagne, Fabienne Maignan.

Data curation: Sauveur Belviso, Camille Abadie.

Funding acquisition: Michel Ramonet, Philippe Ciais.

Investigation: Sauveur Belviso, Camille Abadie, David Montagne, Dalila Hadjar, Didier Tropée, Laurence Vialettes, Victor Kazan, Marc Delmotte, Fabienne Maignan, Marine Remaud, Michel Ramonet, Morgan Lopez.

Methodology: Sauveur Belviso, Camille Abadie, David Montagne, Marc Delmotte, Fabienne Maignan, Michel Ramonet, Morgan Lopez.

Software: Fabienne Maignan, Marine Remaud, Camille Yver-Kwok.

Writing – original draft: Sauveur Belviso, Camille Abadie, David Montagne.

Writing – review & editing: Marc Delmotte, Fabienne Maignan, Marine Remaud, Michel Ramonet, Morgan Lopez, Camille Yver-Kwok, Philippe Ciais.

References

1. Kooijmans LMJ, Cho A, Ma J, Kaushik A, Haynes KD, Baker I, et al. Evaluation of carbonyl sulfide biosphere exchange in the Simple Biosphere Model (SiB4). Biogeosciences. 2021; 18. https://doi.org/10.5194/bg-18-6547-2021
2. Abadie C, Maignan F, Remaud M, Ogée J, Campbell JE, Whelan ME, et al. Global modelling of soil carbonyl sulfide exchanges. Biogeosciences. 2021; 19. https://doi.org/10.5194/bg-19-2427-2022
3. Berry J, Wolf A, Campbell JE, Baker I, Blake N, Blake D, et al. A coupled model of the global cycles of carbonyl sulfide and CO_2: A possible new window on the carbon cycle, Journal of Geophysical Research: Biogeosciences. 2013; 118. https://doi.org/10.1002/jgrg.20068
4. Maignan F, Abadie C, Remaud M, Kooijmans LMJ, Kohonen KM, Commane R, et al. Carbonyl sulfide: comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach. Biogeosciences. 2021; 18. https://doi.org/10.5194/bg-18-2917-2021
5. Ogée J, Sauze J, Kesselmeier J, Genty B, van Diest H, Launois T, et al. A new mechanistic framework to predict OCS fluxes from soils, Biogeosciences. 2016; 13. https://doi.org/10.5194/bg-13-2221-2016
6. Whelan ME, Hilton TW, Berry JA, Berkelhammer M, Desai AR., Campbell EJ. Carbonyl sulfide exchange in soils for better estimates of ecosystem carbon uptake. 2016; 16. https://doi.org/10.5194/acp-16-3711-2016

7. Bloem E, Haneklaus S, Kesselmeier J, Schnug E. Sulfur fertilization and fungal infections affect the exchange of H_2S and COS from agricultural crops. Journal of Agricultural and Food Chemistry. 2012; 60. https://doi.org/10.1021/jf301912h

8. Maseyk K, Berry JA, Billesbach D, Campbell JE, Torn MS, Zahniser M, et al. Sources and sinks of carbonyl sulfide in an agricultural field in the Southern Great Plains. Proceedings of the National Academy of Sciences. 2014; 111. https://doi.org/10.1073/pnas.1319132111

9. Geng C, Mu Y. Carbonyl sulfide and dimethyl sulfide exchange between trees and the atmosphere. Atmospheric Environment. 2006; 40. https://doi.org/10.1016/j.atmosenv.2005.10.023

10. Gimeno TE, Ogée J, Royles J, Gibon Y, West JB, Burlett R, et al. Bryophyte gas-exchange dynamics along varying hydration status reveal a significant carbonyl sulphide (COS) sink in the dark and COS source in the light. New Phytologist. 2017; 215. https://doi.org/10.1111/nph.14584 PMID: 28467665

11. Rastogi B, Berkelhammer M, Wharton S, Whelan ME, Itter MS, Brian Leen J, et al. Large uptake of atmospheric OCS observed at a moist Old Growth Forest: controls and implications for Carbon Cycle Applications. Journal of Geophysical Research: Biogeosciences. 2018;123. https://doi.org/10.1029/2018JG004430

12. Ma J, Kooijmans LMJ, Cho A, Montzka SA, Glatthor N, Worden JR, et al. Inverse modelling of carbonyl sulfide: implementation, evaluation and implications for the global budget. Atmospheric Chemistry and Physics. 2021; 21. https://doi.org/10.5194/acp-21-3507-2021

13. Remaud M, Chevallier F, Maignan F, Belviso S, Berchet A, Parouffe A, et al. Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling. 2021. https://doi.org/10.5194/acp-2021-326

14. Belviso S, Lebegue B, Ramonet M, Kazan V, Pison I, Berchet A, et al. A top-down approach of sources and non-photosynthetic sinks of carbonyl sulfide from atmospheric measurements over multiple years in the Paris region (France). PLoS ONE. 2020; 15(2). https://doi.org/10.1371/journal.pone.0228419 PMID: 32040521

15. Ramonet M, Ciais P, Apadula F, Bartyzel J, Bastos A, Bergamaschi P, et al. The fingerprint of the summer 2018 drought in Europe on ground-based atmospheric CO2 measurements. Philosophical Transactions B. 2020; 375. http://dx.doi.org/10.1098/rstb.2019.0513

16. Belviso S, Reiter IM, Loubet B, Gros V, Lathière J, Montagne D, et al. A top–down approach of surface carbonyl sulfide exchange by a Mediterranean oak forest ecosystem in Southern France. Atmospheric Chemistry and Physics. 2016; 16. https://doi.org/10.5194/acp-16-14909-2016

17. Yver C., Schmidt M, Bousquet P, Zahorowski W, Ramonet M. Estimation of the molecular hydrogen soil uptake and traffic emissions at a suburban site near Paris through hydrogen, carbon monoxide, and radon-222 semi continuous measurements. Journal Geophysical Research: Atmospheres. 2009; 114, D18304. https://doi.org/10.1029/2009JD012122

18. Lin X, Indira NK, Ramonet M, Delmotte M, Ciais P, Bhatt BC, et al. Long-lived atmospheric trace gases measurements in flask samples from three stations in India, Atmospheric Chemistry and Physics. 2015; 15. https://doi.og/10.5194/acp-15-9819-2015

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

20. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. Quaterly Journal of the Royal Meteorological Society. 2020; 146, https://doi.org/10.1002/qj.3803

21. Sitch S, Friedlingstein P, Gruber N, Jones SD, Murray-Tortarolo G, Ahlström A, et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. Biogeosciences. 2015; 12. https://doi.org/10.5194/bg-12-653-2015

22. Lardy R., Bellocchi G., and Soussana JF. A new method to determine soil organic carbon equilibrium. Environmental Modelling & Software. 2011; 26. https://doi.org/10.1016/j.envsoft.2011.05.016

23. Pal S, Lopez M, Schmidt M, Ramonet M, Gibert F, Xueref-Remy I, et al. Investigation of the atmospheric boundary layer depth variability and its impact on the ^{222}Rn concentration at a rural site in France, Journal Geophysical Research: Atmospheres. 2015; 120, 623–643, https://doi.org/10.1002/2014JD022322

24. Hannigan JW, Ortega I, Bahramvash Shams S, Blumenstock T, Campbell JE, Conway S, et al. Global Atmospheric OCS Trend Analysis from 22 NDACC Stations. Journal Geophysical Research: Atmospheres. 2022; 127. https://doi.org/10.1029/2021JD035764

25. Kesselmeier J, Merk L. Exchange of carbonyl sulfide (COS) between agricultural plants and the atmosphere: studies on the deposition of COS to peas, corn and rapeseed. Biogeochemistry. 1993; 23.

26. Gonzaga Gomez L, Loubet B, Lafouge F, Ciuraru R, Buysse P, Duranda B, et al. Comparative study of biogenic volatile organic compounds fluxes by wheat, maize and rapeseed with dynamic chambers

over a short period in northern. Atmospheric Environment. 2019; 214. https://doi.org/10.1016/j.atmosenv.2019.116855

27. Gonzaga Gomez L, Loubet B, Lafouge F, Ciuraru R, Bsaibes S, Kammer J, et al. Effect of senescence on biogenic volatile organic compound fluxes in wheat plants. Atmospheric Environment. 2021; 266. https://doi.org/10.1016/j.atmosenv.2021.118665

28. Berkelhammer M, Alsip BM, Matamala R, Cook DR, Whelan ME, Joo E, et al. 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. 2019; 47. https://doi.org/10.1029/2019GL085652

29. Blaževića I, Montaut S, Burčul F, Olsen CE, Burow M, Rollin P, et al. Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. Phytochemistry. 2020; 169. https://doi.org/10.1016/j.phytochem.2019.112100

30. Fahey JW, Zalcmann AT, Talalay P. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry. 2001; 56. https://doi.org/10.1016/s0031-9422(00)00316-2 PMID: 11198818

31. Bonjean AP, Dequidt C, Sang T—Groupe Limagrain. Rapeseed in China. Oilseeds & fats Crops and Lipids. 2016; 23(6) D605. https://doi.org/10.1051/ocl/2016045

9.3 Evaluation of CMIP/TRENDY model gross primary productivity using atmospheric COS and CO_2 data

Evaluation of CMIP/TRENDY model gross primary productivity and terrestrial ecosystem respiration using atmospheric COS and CO_2 data

Peylin Philippe[1], Abadie Camille[1], Haslehner Mylene[2], Belviso Sauveur[1], Cadule Patricia[3], Remaud Marine[1,4], Maignan Fabienne[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement (LSCE Saclay), IPSL, CEA, CNRS, UVSQ, CE Saclay, Bât 701 L'Orme des Merisiers, 91191 Gif-sur-Yvette, France
[2]Somic Verpackungsmaschinen GmbH & CO KG, 82123 Amerang, Germany
[3]Institut Pierre-Simon Laplace, Sorbonne Université/CNRS, Paris, France
[4]Amsterdam Institute for Life and Environment (A-LIFE), Section Systems Ecology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Corresponding author: philippe.peylin@lsce.ipsl.fr

Abstract

Vegetation carbon uptake through photosynthesis (referred to as Gross Primary Production, GPP) is the most uncertain flux of the carbon cycle. Simulation of GPP and the associated ecosystem respiration flux (TER) by global land surface models (LSMs) still differ substantially between models, hampering robust prediction of the terrestrial carbon budget. Reducing GPP uncertainty is thus a key challenge of the climate modeling community. Recently, carbonyl sulfide (COS) has been proposed as a new tracer to constrain GPP at large spatial scales. COS is absorbed by plants like CO_2, following the same diffusion pathway but there is no emission flux like plant respiration for CO_2. In this study, we propose to use the atmospheric measurements of COS and CO_2 to provide new constraints on both GPP and TER of global LSMs. Specifically, we evaluate model simulations from three ensembles, the last two climate modeling intercomparison projects CMIP5 and CMIP6 ensembles and the TRENDY-v7 LSM intercomparison used for the Global Carbon Budget (GCB). For each model, COS plant uptake is estimated using the Leaf Relative Uptake (LRU) empirical relationship, which links COS to CO_2 leaf uptake, with CO_2 uptake represented by the simulated GPP. The LMDZ transport model is used to simulate atmospheric concentrations derived from surface fluxes. We focus the evaluation on the mean seasonal cycle of GPP and TER using the mean seasonal cycle of atmospheric COS and CO_2 concentrations at two stations of the Northern hemisphere, Mauna Loa (MLO) and Alert (ALT). The evaluation of the simulated COS and CO_2 concentrations allows us to identify specific biases in the simulated GPP seasonal cycle amplitude of individual LSMs from the three ensembles. The CMIP5 ensemble leads to a large spread of the seasonal amplitude at both MLO and ALT stations that includes the observation but with clear model outliers (either too low or too large amplitude). The spread of the amplitude is largely reduced in the CMIP6 ensemble, indicating substantial improvements of the modeled GPP compared to CMIP5. Despite using similar climate forcing, the LSMs of the TRENDY-v7 collection show also a large amplitude spread at MLO and ALT, revealing specific model biases for temperate and high latitude ecosystems. In addition, the analysis of the phase of the mean COS and CO_2 concentration seasonal cycles highlights phase issues with respect to GPP and/or TER for specific models of each ensemble. This study demonstrates the potential of combining atmospheric COS and CO_2 concentration measurements to evaluate GPP and TER simulated by LSMs. The proposed

diagnostics are however still dependent on the overall COS flux set up, not only the LRU values linking plant COS uptake to GPP, but also the soil COS uptake and the ocean source. Further quantifying the seasonality of these fluxes will help refine the proposed diagnostics.

Key words: COS, GPP, CO_2

1. Introduction

Projections of terrestrial vegetation functioning in response to climate changes (precipitation, temperature, light and humidity), atmospheric CO_2 increase and nutrients availability and its capacity to mitigate climate change heavily rely on our ability to accurately represent vegetation carbon exchanges in land surface models (LSMs) (Anav et al., 2015). Although vegetation carbon uptake through photosynthesis, known as gross primary production (GPP), is the largest flux in the terrestrial carbon cycle over the globe, it is also the most uncertain (Friedlingstein et al., 2023). To quantify GPP at large scales (Beer et al., 2010), identify long-term trends (Sitch et al., 2008, 2024), and study the drivers of changes in plant carbon fluxes (Sitch et al., 2015), various LSMs have been developed by the land surface modeling community, following similar principles but differing in the implementation of the different processes regulating the water, carbon, and energy transfers in the soil - plant - atmosphere continuum. Although these models are usually calibrated using different observational datasets related to either plant physiological states or water, carbon and energy fluxes (i.e., eddy-covariance measurements), large differences exist between the simulated gross and net carbon fluxes as well as the simulated carbon stocks (both spatially and temporally; Anav et al., 2015). These differences can be seen as uncertainties arising from the model structure related to process representation, the parameter values used in these models (Zaehle et al., 2005), and the uncertainty in the data used to drive the models (Wu et al., 2017).
For over 25 years, intercomparison exercises between Earth System Models (ESMs) including a LSM as the surface component, have been conducted to compare and evaluate various climate-related model outputs (Eyring et al., 2016; Taylor et al., 2012; Meehl et al., 2000), provide climate projections, and contribute to the Intergovernmental Panel on Climate Change (IPCC) reports. These regular intercomparison experiments allow for the assessment of LSM performances compared to previous versions and other new models. For instance, GPP was found to be overestimated in the models included in the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Anav et al., 2013). Gier et al. (2024) found that the issue of overestimating GPP in CMIP5 was mostly addressed in CMIP6 for models with an interactive nitrogen cycle, but persisted in models that lacked this feature. In addition, most LSMs participate each year in the global carbon budget (GCB; Friedlingstein et al., 2023) through the TRENDY intercomparison (Trend in the Land Carbon Cycle). Despite similar climates, land use and nutrient forcings, there is still a large spread in terms of gross carbon fluxes (Sitch et al., 2024). For temperate and high latitude ecosystems differences occur primarily on the amplitude of the seasonal cycle while for tropical ecosystems phase differences dominate.
Accurately simulating plant GPP and plant respiration fluxes in LSMs is challenging due to the lack of direct measurements beyond the leaf scale (Damm et al., 2010). The net ecosystem carbon exchange includes contributions from both vegetation uptake and soil and vegetation respiration (terrestrial ecosystem respiration, TER). Consequently, partitioning these fluxes at the ecosystem scale depends on models that incorporate various assumptions about the drivers of these gross fluxes and uncertainties (Tramontana et al., 2020). Despite the emergence of several proxies to constrain vegetation activity, such as the near-infrared reflectance of vegetation (Badgley et al., 2017), the normalized difference vegetation index (Bhandari et al., 2012), or solar-induced fluorescence (Berry et al., 2018), LSMs exhibit significant differences in their simulated carbon

fluxes, which can be of the same order of magnitude as the range of estimates from observational data (Seiler et al., 2022).

In the last decade, carbonyl sulfide (COS), a trace gas in the atmosphere that can be measured at ground stations (Montzka et al., 2007), has sparked increasing interest for informing GPP and stomatal conductance in LSMs (Abadie et al., 2023; Berry et al., 2013; Cho et al., 2023; Kooijmans et al., 2021; Maignan et al., 2021; Remaud et al., 2022; Chen et al., 2023; Zhu et al., 2024a, 2024b). COS exhibits a seasonal cycle of atmospheric concentrations relatively similar to that of CO_2, with a drawdown in spring and summer in the Northern hemisphere due to vegetation uptake (although with some lag), the primary land sink for both gases (Montzka et al., 2007; Whelan et al., 2018). Indeed, both COS and CO_2 are absorbed by plant leaves through a common diffusion pathway. However, COS benefits from the absence of a return flux analogous to that of plant respiration for CO_2 (Stimler et al., 2010). Recently, several studies have used biospheric COS flux measurements to directly constrain GPP-related parameters in LSMs through data assimilation (Abadie et al., 2023; Cho et al., 2023; Chen et al. 2023; Zhu et al., 2024a, 2024b). Ecosystem COS flux observations provide valuable information on GPP at the ecosystem scale, as these fluxes are predominantly influenced by vegetation COS uptake (Commane et al., 2015). However, such observations are limited to a few sites and rarely span multiple years (Berkelhammer et al., 2014; Commane et al., 2015; Kooijmans et al., 2019; Wehr et al., 2017).

In contrast, atmospheric COS concentrations have been recorded at numerous stations since 2000 (Montzka et al., 2007). Consequently, previous studies have used temporal and spatial gradients of COS concentration measurements to evaluate and reduce uncertainty in the components of the global COS budget (Remaud et al., 2023, Berry et al., 2013; Kettle et al., 2002; Suntharalingam et al., 2008), aiming to exploit COS data to constrain GPP. Launois et al., (2015) used atmospheric COS data to assess the annual, seasonal, and latitudinal variations of GPP from three LSMs (LPJ, CLM4 and ORCHIDEE), with a first attempt to scale the different COS flux components including the GPP-induced COS uptake. More recently, atmospheric inversion studies have utilized atmospheric COS concentration measurements to optimize the spatial and temporal gradients of COS surface fluxes (Ma et al., 2021) and constrain simulated GPP and respiration fluxes from LSMs (Remaud et al., 2022).

In this study, we aim to utilize the dual information provided by the temporal gradients of atmospheric CO_2 concentrations, indicative of both GPP and respiration fluxes, while using atmospheric COS concentrations to provide a more direct constraint on GPP. The link between the surface fluxes and the atmospheric concentrations for CO_2 and COS will be made with the atmospheric transport model, LMDz (Hourdin et al., 2006). Hence, the joint analysis of COS and CO_2 concentrations will intend to identify potential biases in GPP and/or TER fluxes simulated by LSMs from three model intercomparison ensembles (CMIP5, CMIP6 and a version of TRENDY). The proposed framework focuses on the mean seasonal cycle with the objective to identify potential seasonal amplitude and phase biases of the individual models, while also evaluating systematic biases within and between the three ensembles.

The first section outlines our approach to simulate atmospheric COS and CO_2 concentrations from the various surface flux components for each model of the three intercomparison experiments, alongside a description of the observed COS and CO_2 concentrations used for the evaluation. We then detail the framework behind the joint analysis of COS and CO_2 concentrations to detect potential biases in the seasonal amplitude and/or phase of simulated GPP and respiration fluxes. The results are then detailed for two atmospheric stations and the three model ensembles, including sensitivity analysis of the results to various choices for the COS fluxes set up. Finally, we discuss the overall benefit of atmospheric COS concentrations in evaluating simulated GPP and respiration in LSMs, along with the limitations and anticipated improvements for future model evaluations using this framework.

2. Observations, models and methodology

2.1. Atmospheric measurements of COS and CO2

The atmospheric COS and CO_2 concentration measurements are taken from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA/ESRL). COS pair flask samples have been collected one to five times a month since 2000 between 11 and 17h local time and analyzed using gas chromatography and mass spectrometry detection at the NOAA/GML's Boulder laboratories. The maximum difference allowed between the pair flasks is 6.3 ppt to retain COS measurements. The atmospheric COS concentration measurements can be downloaded at ftp://ftp.cmdl.noaa.gov/hats/carbonyl_sulfide/ (last access: 7 February 2022) (extension of the data first published in Montzka et al., 2007) and the atmospheric CO_2 concentration data come from the NOAA's GlobalView Plus Observation Package (ObsPack; Cooperative Global Atmospheric Data Integration Project, 2018).
We have used 10 stations with available COS and CO_2 measurements (Table 1) but we mainly focus on 2 stations in the core of the manuscript: Mauna Loa (MLO) situated in the Hawaiian region and Alert (ALT) in Canada. While MLO represents the background atmospheric concentrations of the whole Northern hemisphere, ALT represents more specifically the influence of Northern high latitude ecosystems. Note that both MLO and ALT stations are located relatively far from human influences. Therefore, their seasonal CO_2 and COS cycles are expected to undergo fewer anthropogenic contributions. The results for the other stations are provided in the Annexes.

Table 1: List of air sampling stations selected for the joint analysis of COS and CO_2 data. The two stations of interest for this study (MLO and ALT) are in bold.

Code	Stations	Coordinates	Elevation above sea level (m)
SPO	South Pole, Antarctica (United States)	90.0°S, 24.8°E	2810
CGO	Kennaook / Cape Grim, Australia	40.68°S, 144.69°E	164
SMO	Tutuila, American Samoa	14.25°S, 170.56°W	77
MLO	**Mauna Loa, United States**	**19.54°N, 155.58°W**	**3397**
KUM	Cape Kumukahi, United States	19.74°N, 155.01°W	3
NWR	Niwot Ridge, United States	40.04°N, 105.54°W	3475
LEF	Wisconsin, United States	45.95°N, 90.28°W	868
MHD	Mace Head, Ireland	53.33°N, 9.9°W	18
BRW	Utqiagvik (formerly Barrow), United States	71.32°N, 155.61°W	8
ALT	**Alert, Canada**	**82.45°N, 62.51°W**	**195**

2.2. Land surface model simulations

For the COS and CO_2 analysis, we are using simulated GPP and net ecosystem productivity (NEE) fluxes from process-based global land surface models. Three ensembles of simulations

corresponding to three different model intercomparison projects (MIPs) are used: the Coupled Model Intercomparison Projects Phase 5 (CMIP5, (Taylor et al., 2012)) with 24 models and Phase 6 (CMIP6, (Eyring et al., 2016)) with 25 models, and the Trend in the Land Carbon Cycle Project Version 7 (TRENDY-v7; Le Quéré et al., 2018) with 15 models. The complete list of the models for each ensemble is reported in the Appendix, Table A1.

CMIP5 and CMIP6 simulations:
We used the simulated land carbon fluxes of the historical climate simulations made with different Earth System Models (ESMs) that were used in these two exercises (2000 - 2005 for CMIP5 and 2000 - 2014 for CMIP6). The different model simulations thus largely differ in terms of climate forcing that is seen by the land surface component of each ESM to compute the gross and net carbon fluxes. Note that several variants exist for a given ESM, mainly with differences linked to the spatial resolution of the main model components (i.e., Ocean, Atmosphere) and that several ESMs use the same land surface model, although with different atmosphere or ocean components.

TRENDY-v7 simulations:
The TRENDY simulations are performed each year by most global LSMs in order to contribute to the Global Carbon Budget (GCB) synthesis of the Global Carbon Project (GCP). We have used the version v7 with flux simulations covering the period 1960 - 2017 (see Le Quéré et al., 2018); this version corresponds approximately to the LSMs version that were used in the CMIP6 exercise. Contrary to CMIP simulations, the TRENDY experiment uses a prescribed climate forcing for all land surface models: the 6-hourly climate reanalysis, JRA-55, with a monthly bias correction based on the Climate Research Unit (CRU) data (Harris et al., 2019). Like for the CMIP simulations, land cover changes make use of the HYDE land-use change data set (Klein Goldewijk et al., 2017), which provides annual half-degree fractional data on cropland and pasture.
From all model simulations, we extracted the monthly mean GPP and NBP to simulate the atmospheric CO_2 and COS concentrations as described in the following sections. All fluxes have been re-gridded from the native resolution (see Appendix, Table A1) to the spatial resolution of the atmospheric transport model (see below).

2.3. Modeled COS flux components

The modeled COS fluxes considered in this study are vegetation uptake, exchanges by oxic soils that can be both uptake or emission, anoxic soil production, oceanic emissions, biomass burning emissions, anthropogenic sources, and oxidation by OH in the troposphere. COS photolysis in the stratosphere is neglected, which would account for -50 $\pm$ 15 GgS/y (Whelan et al., 2018). Emissions of COS from volcanoes are also neglected as they are largely dominated by eruptive and post-eruptive emissions (Belviso et al., 1986; Whelan et al., 2018), the intermittent nature of which makes the estimation highly uncertain.

2.3.1. Vegetation COS uptake

Plant uptake is the largest COS sink (Asaf et al., 2013; Berry et al., 2013; Campbell et al., 2008; Montzka et al., 2007). Inside the leaf, COS follows the same diffusional pathway as CO_2 from the ambient air to the chloroplast, where it is irreversibly hydrolysed by the carbonic anhydrase (CA) enzyme following the exergonic (spontaneous) reaction: $COS + H_2O \rightarrow CO_2 + H_2S$ (Protoschill-Krebs et al., 2016; Stimler et al., 2010).
A mechanistic model of vegetation COS uptake has been developed by Berry et al. (2013) to represent COS diffusion and hydrolysis inside the leaves. This mechanistic approach has only been implemented in a few LSMs for now, as in SiB (Berry et al., 2013; Kooijmans et al., 2021) and ORCHIDEE (Maignan et al., 2021). Therefore, in this study we use an empirical concept, the Leaf

Relative Uptake (LRU) (Sandoval-Soto et al., 2005), to compute vegetation COS uptake based on GPP (µmol COS/m^2/s) simulated by each model,

$$F_{COS,vegetation} = LRU \cdot GPP \cdot \frac{[COS]}{[CO_2]} \quad (1)$$

where $F_{COS,vegetation}$ is COS vegetation uptake (pmol COS/m^2/s) and $[COS]$ and $[CO_2]$ are the atmospheric mixing ratios of COS (ppt) and CO_2 (ppm). We use hemispheric mean mixing ratios estimated with weighted monthly means from NOAA sampling stations for $[COS]$, and a global mean mixing ratio for $[CO_2]$.

A constant (in time) LRU value is usually defined for each vegetation type (Maignan et al., 2021; Seibt et al., 2010; Whelan et al., 2018). Maignan et al. (2021) compared the use of the LRU empirical approach and the mechanistic vegetation COS uptake model from Berry et al. (2013) using the ORCHIDEE LSM and the LMDz atmospheric transport model. Their study shows that the two approaches give similar results in terms of simulated atmospheric COS concentrations when transporting monthly mean vegetation COS fluxes obtained with the empirical or the mechanistic approach.

Several sets of LRU value per vegetation type have been estimated experimentally (Seibt et al., 2010; Whelan et al., 2018). In this study, we aim at testing the sensitivity of the simulated atmospheric COS concentrations to different sets of LRU values. First, we use as our reference the LRU values from Seibt et al., (2010) (LRU_{ref}), estimated for various plant types assuming that the internal conductance is an intermediate limitation in the COS diffusional pathway (the stomatal to mesophyll conductances ratio equals to 0.1). In addition, we use the LRU set from Whelan et al. (2018) that defines one LRU value for all C3 plant types and one for all C4 plant types (LRU_{Whelan}). Finally, we consider a set of LRU values computed from the ORCHIDEE LSM as described in Maignan et al., (2021) (LRU_{ORC}). We define the different vegetation categories using the 13 plant functional types described in Table 7 of Poulter et al., (2011). Then, we assign a LRU value per PFT for each LRU set as presented in Table 2. Note that the values of LRU_{Whelan} and LRU_{ORC} are both lower than LRU_{ref}, by 30% for LRU_{Whelan} and 40% for LRU_{ORC} on average across all PFTs. Considering lower sets of LRU values reduces vegetation COS uptake according to equation 1, as presented in Table 3 with a mean vegetation COS sink ranging from -1324 to -1070 GgS/y using LRU_{ref}, compared to estimates between -757 and -615 GgS/y for LRU_{Whelan} and between -709 and -573 GgS/y for LRU_{ORC} across all model experiments.

Table 2: Sets of LRU per PFT tested in this study.

PFTs	Seibt Ref: LRU_{ref}	Whelan: LRU_{Whelan}	ORCHIDEE: LRU_{ORC} *
1 Tropical Evergreen	3.09	1.68	1.79
2 Tropical Raingreen	3.38	1.68	1.69
3 Temperate Needleleaf Evergreen	1.89	1.68	1.41
4 Temperate Broadleaf Evergreen	3.60	1.68	1.11
5 Temperate Broadleaf Summergreen	3.60	1.68	1.40
6 Boreal Needleleaf Evergreen	1.89	1.68	0.97
7 Boreal Needleleaf Deciduous	1.89	1.68	0.92
8 Boreal Broadleaf Summergreen	1.94	1.68	1.07
9 Natural Grassland (C3)	2.53	1.68	1.19
10 Natural Grassland (C4)	2.53	1.21	1.33
11 Managed Grassland (C3)	2.26	1.68	1.38
12 Managed Grassland (C4)	2.26	1.21	1.59
13 Non-vegetated (Barren, Water, Urban)	0.00	0.00	0.00

** The LRU values obtained with the ORCHIDEE LSM differ from the ones presented in (Maignan et al., 2021) as they are computed using 3-hourly variable atmospheric COS and CO_2 concentrations in this study, while Maignan et al. (2021) used a constant value of 500 ppt for COS and monthly means for CO_2.*

2.3.2. Soil COS exchanges

At the global scale, soils have been estimated to be a net COS sink, with a budget ranging from −30 to −409 GgS yr^{-1} (Abadie et al., 2022; Berry et al., 2013; Kettle et al., 2002; Kooijmans et al., 2021; Launois et al., 2015). COS can be both taken up or emitted by oxic soils, whereas it was found to be only emitted by anoxic soils (Whelan et al., 2018). In oxic soils, COS can diffuse from the atmosphere into the soil matrix and be hydrolyzed mainly by CA contained in soil microorganisms (Smith et al., 1999), and be produced through thermal or photo degradation processes (Whelan et al., 2018). On the other hand, COS production by anoxic soils has been related to the activity of sulfate reduction metabolisms (Whelan et al., 2013).

Several empirical (Berry et al., 2013; Launois et al., 2015) and mechanistic (Ogée et al., 2016; Sun et al., 2015) models of soil COS exchanges have been developed and implemented in LSMs (Abadie et al., 2022; Berry et al., 2013; Kooijmans et al., 2021; Launois et al., 2015). In this study, we use a mechanistic representation of soil COS fluxes (our reference) and test an empirical approach to evaluate the sensitivity of simulated atmospheric COS concentrations to soil COS flux. The mechanistic approach is based on the Ogée et al., (2016) model implemented in ORCHIDEE as described in Abadie et al., (2022). It resolves COS diffusion into the soil matrix, abiotic and biotic COS hydrolysis, and COS production by oxic soils (Ogée et al., 2016). Anoxic soil grid cells are

represented using a map of wetlands from Tootchi et al., (2019), and anoxic soil COS flux is expressed as a function of soil temperature following Ogée et al., (2016). The more empirical approach is based on a relationship between H_2 deposition and soil COS uptake developed by Launois et al., (2015). A distinction is also made for anoxic soils, which are represented by a map of seasonal methane emissions from Wania et al., (2010) with anoxic soil COS fluxes derived from Whelan et al. (2013).

2.3.3. Ocean COS emissions

The ocean is a major source of COS with a global contribution estimated between 120 and 600 GgS yr^{-1} (Lennartz et al., 2020). COS can be directly emitted by the ocean or indirectly emitted through atmospheric oxidation of dimethyl sulfide (DMS) and carbon disulfide (CS_2) produced in seawater. Large uncertainties remain on oceanic COS production with previous inversion studies supporting an oceanic missing source (Berry et al., 2013; Kuai et al., 2015; Launois et al., 2015; Suntharalingam et al., 2008). Top-down studies also lead to higher COS emissions from the ocean, estimated between 400 and 800 GgS yr^{-1} (Kuai et al., 2015; Remaud et al., 2022).
In this study, we test two oceanic COS flux estimates to evaluate the impact of ocean COS flux representation on simulated atmospheric COS concentrations. We first use the optimized ocean COS fluxes computed with the atmospheric inversion system described in Remaud et al., (2022) as our reference. In this inversion study, Remaud et al. (2022) assimilated both COS and CO_2 concentration measurements using the LMDz atmospheric transport model, aiming at better constraining GPP. The optimization led to an optimized ocean COS budget of 526 GgS yr^{-1} with the largest emissions located in the tropics. As a second ocean COS estimate, we consider the direct oceanic COS emissions and indirect emissions from CS_2 estimated by Lennartz et al., (2017, 2020), with indirect emissions from DMS simulated by the NEMO-PISCES model as described in Launois et al. (2015a).

2.3.4. Anthropogenic COS emissions

A second major source of COS comes from anthropogenic activities. Anthropogenic emissions include direct COS emissions and indirect emissions from CS_2 oxidation in the atmosphere. The largest anthropogenic COS production is related to the rayon (staple and yarn) industry, followed by residential and industrial coal, pigments, aluminum melting, agricultural chemicals and tire wear (Campbell et al., 2015). The recent gridded inventory by Zumkehr et al., (2018) has highlighted an underestimation of anthropogenic COS emissions in the previous estimate from Kettle et al., (2002), with a revised budget ranging from 223 to 586 GgS yr^{-1}. We choose the Zumkehr et al., (2018) budget that provides emissions for the period 1980–2012. The largest emissions are concentrated in Asia and 45% of the global source is attributed to China. Note however that there are still large uncertainties associated with anthropogenic emissions, as pointed out for Europe in Belviso et al. (2023). Note also that soil COS emissions from managed lands are not included in the anthropogenic COS source inventory as they are taken into account in the soil COS exchanges.

2.3.5. Biomass burning COS emissions

The last source of COS considered in this study is biomass burning, which accounts for over 10% of COS global budget with 50 to 100 GgS yr^{-1} (Glatthor et al., 2015; Stinecipher et al., 2019). We use the recent inventory of Stinecipher et al., (2019) that is based on a global map of fire emissions provided by the Global Fire Emissions Database (GFED). GFED considers six biomass burning categories, which are savanna and grassland, boreal forests, temperate forests, tropical

deforestation and degradation, peatland fires, and agricultural waste burning. An emission factor and an emission ratio for COS is defined for each biomass burning category.

2.4. Modeled CO_2 flux components

In order to perform a joint COS - CO_2 analysis using the atmospheric data at the 10 selected sites (Table 1), we also need to transport the different components of the global atmospheric CO_2 budget. The objective is to use the terrestrial ecosystem fluxes from each LSM simulation and to add the other surface flux components, namely the anthropogenic emissions and the ocean fluxes, using recent estimates used for the GCP.

Land ecosystem CO_2 fluxes:
For each CMIP5/CMIP6/TRENDY-v7 model simulation, we use the simulated monthly mean NBP. Note that NBP corresponds to GPP minus all ecosystem respiration fluxes and that it includes potentially additional disturbance fluxes that may be considered in a given model, such as biomass burning. In addition, we also use for one model (ORCHIDEE from the TRENDY-v7 ensemble) the GPP and respiration fluxes separately, in order to illustrate their individual contribution to the simulated CO_2 concentration seasonal cycle.

Anthropogenic emissions of CO_2:
We use global fossil fuel emissions based on the GCP's Gridded Fossil Emissions Dataset version 2021.2 (GCP-GridFEDv2021.2; Jones et al., 2021). GCP-GridFEDv2021.2 scales gridded estimates of CO_2 emissions from EDGARv4.3.2 (Janssens-Maenhout et al., 2019) within national territories to match national emission estimates provided by GCB for the years 1959–2020.

Ocean CO_2 fluxes:
We use the optimized gridded global air-sea fluxes (monthly mean) from the CAMS atmospheric CO_2 inversion (version v18r2) provided to the 2019 GCB (Friedlingstein et al., 2019). The inversion is using the same atmospheric transport model (LMDz) than the one used in this study, although based on a more recent model version (V5, Remaud et al., 2018).

2.5. Atmospheric transport model and concentration simulations

We use an atmospheric transport model to simulate the atmospheric CO_2 and COS concentrations from the different set of fluxes described above. We use the version 3 of the atmospheric circulation model of the Laboratoire de Météorologie Dynamique (referred to as LMDZ-v3 (Hourdin and Armengaud, 1999) with a horizontal resolution of 3.75° (longitude) x 2.5° (latitude) and 19 sigma-pressure layers up to 3 hPa. The calculated winds are nudged to the European Centre for MediumRange Weather Forecasts (ECMWF) reanalysis, ERA-40, meteorological data (Uppala et al., 2005) with a relaxation time of 2.5h in order to realistically account for large-scale advection (Hourdin et al., 2006). For this study, we use pre-calculated transport fields (as in Peylin et al., 2016) that correspond to the sensitivity of the concentration at each atmospheric site to the surface flux of each model grid cell (often referred to as influence functions). Such an approach allows to reduce the computing time (compared to the use of the full LMDz model), once the influence functions have been calculated; the transport is then replaced by a matrix multiplication.
To simulate the atmospheric CO_2 concentrations, we transport the NBP flux of all models of the CMIP5, CMIP6 and TRENDY-v7 ensembles, as well as a common air-sea CO_2 exchange and fossil fuel CO_2 emission (see section 2.4). For the COS concentrations, we transport the surface COS fluxes described in section 2.3: a vegetation uptake derived from each modeled GPP flux and common soil, biomass burning, anthropogenic and ocean fluxes. The overall procedure is illustrated in Figure

1. Monthly atmospheric CO_2 and COS concentrations are simulated for 5 years (2005 to 2009) in order to compute a mean seasonal cycle, which is the focus of this study.

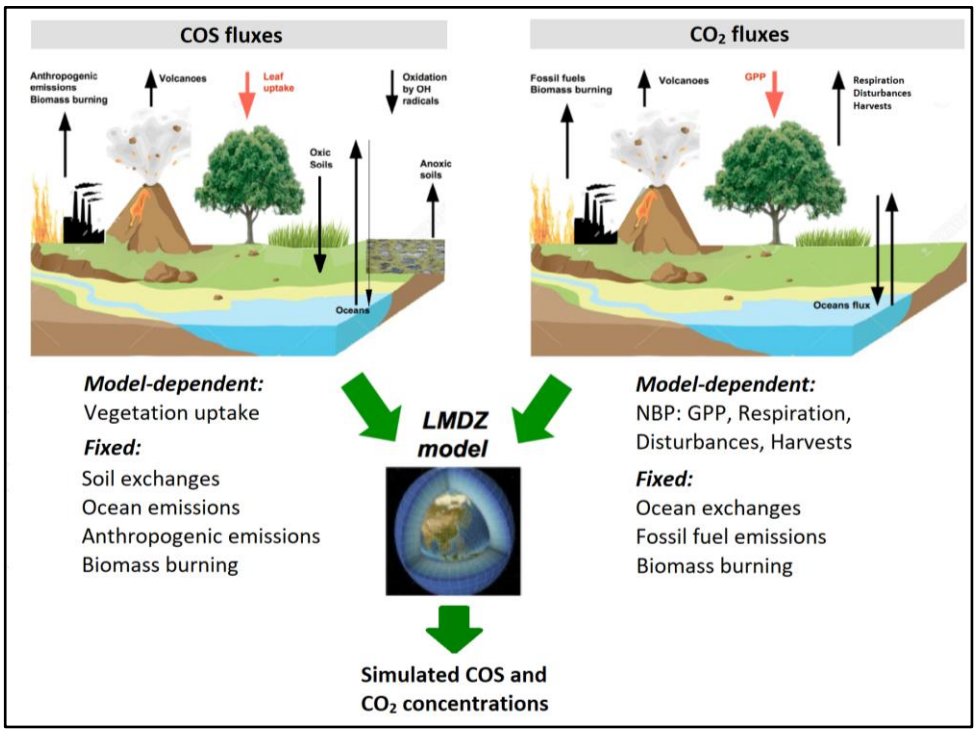

Figure 1: Flowchart of atmospheric CO_2 and COS concentration computation from surface CO_2 and COS fluxes using LMDz transport model.

2.6. Method: joint COS - CO_2 analysis and experimental design

2.6.1. Data processing - Inferring model biases from a joint COS and CO_2 analysis

As described in the introduction, we focus in this paper on the mean seasonal cycle amplitude and phase of COS and CO_2 simulated concentrations to highlight possible model-observation misfits that could be attributed to biases in GPP or respiration fluxes. In order to compute the mean seasonal cycle of simulated and observed concentrations, we use the curve fitting procedure developed by Thoning et al. (1989) (CCGVU; http://www.esrl.noaa.gov/gmd/ccgg/mbl/crvfit/crvfit.html). This procedure fits to the raw data a smoothed-function defined by a second-order polynomial equation (for the growth rate), combined with a four-harmonic function (for the annual cycle) and a low-pass filter of the residuals (with a cut-off defined at 120 days). The mean seasonal cycle is then defined as the sum of the harmonics and the short-term residuals. Figure 2 provides an illustration of the proposed seasonal cycle analysis.

For the amplitude analysis (defined as the maximum minus the minimum of the mean seasonal cycle), we normalize the simulated amplitude with the observed one. For each model, if the seasonal amplitude of the COS and CO_2 concentrations are equal to the observed ones, the model will be located on the red cross in Figure 2.a, which represents the non-biased situation. The red vertical / horizontal dotted lines are used to distinguish four regions in Figure 2.a, according to the observational target. If a model falls into the upper-right / lower-left region in Figure 2.a, it indicates that both COS and CO_2 seasonal amplitudes are overestimated / underestimated compared to the observations, respectively. Since GPP flux directly impacts CO_2 and COS simulated concentrations, it indicates that GPP likely plays a key role in the model-observation misfit. For instance, increasing

(or decreasing) GPP seasonal flux amplitude would increase (or decrease) the atmospheric CO_2 seasonal cycle amplitude but also the COS amplitude given the linear dependency of COS plant uptake to GPP (Eq. 1). Therefore, a change in GPP could help to improve both COS and CO_2 simulated seasonal amplitudes. On the contrary, if a model is located into the upper-left / lower-right box of Figure 2.a, it indicates that the simulated COS amplitude is overestimated / underestimated while the simulated CO_2 amplitude is underestimated / overestimated compared to the observation. These two configurations suggest that a change in the GPP flux only cannot improve the simulated COS and CO_2 amplitudes simultaneously. In this case, simulated concentration biases indicate that the seasonal amplitude and/or phase of both GPP and respiration fluxes are likely biased.

For the phase analysis, we use the correlation between simulated and observed monthly concentrations as a proxy for phase biases. Similarly to the amplitude, the non-biased situation for COS and CO_2 concentrations corresponds to a perfect correlation, as represented by the red cross in Figure 2.b. If a model is located along the diagonal axis (red dotted line in fig. 2b) but with low correlations, it may indicate the predominance of GPP flux in the simulated COS and CO_2 phase biases. If a model falls close to the upper horizontal red dotted line, the model phase bias likely comes from a respiration bias since the COS correlation tends to support an accurate GPP flux. On the contrary, when a model lies close to the vertical red dotted line, it may highlight both GPP and respiration phase biases (i.e., a GPP bias for COS phase and a respiration bias to compensate for any change in GPP).

Note that the proposed analysis framework assumes that all the other components of the COS and CO_2 budgets as well as the transport model are perfect.

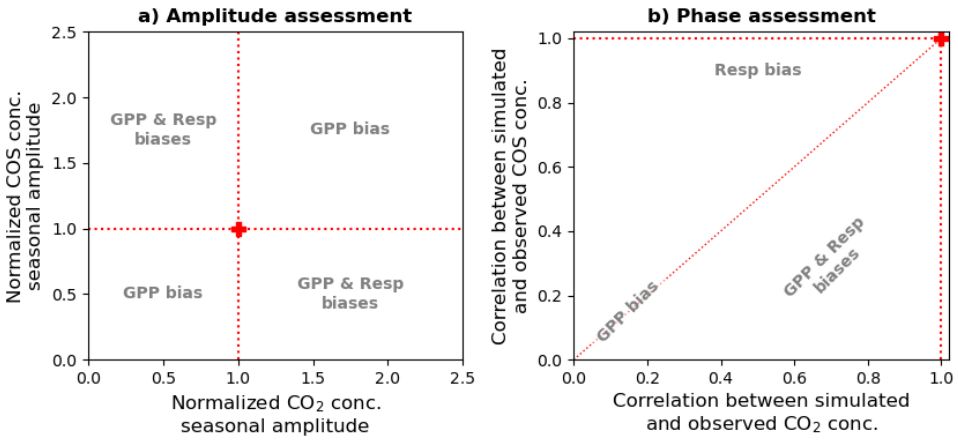

Figure 2: Schemes of model biases inferred from joint COS and CO_2 concentrations analysis with a) CO_2 vs COS simulated seasonal amplitudes normalized by the observed seasonal amplitudes; and b) CO_2 vs COS correlations (proxy of the phase) between simulated and observed concentrations. In both panels, different regions of the graph indicate model biases from GPP and/or respiration flux.

2.6.2. Experimental set-up

As mentioned in section 2.3, we propose to test several variants for different COS flux components in order to evaluate the impact of the most uncertain COS components on the simulated concentrations. Our reference simulations correspond to the use of LRU_{ref} proposed in Seibt et al. (2010), the mechanistic soil fluxes and the optimized ocean fluxes. But we also assess two other LRU cases, LRU_{whelan} and LRU_{ORC}, and one variant for the soil fluxes and the ocean fluxes (see section

2.3). These scenarios and the associated total COS flux are reported in Table 3. Note that for the CO_2 analysis, we do not test different variants for the ocean and anthropogenic fluxes.

Table 3: COS sinks and sources considered in this study, averaged between 2010 and 2014. The COS fluxes selected for the reference scenario to compute atmospheric COS concentrations are represented in bold.

COS flux	Representation	References	Budget (GgS/y) mean [min,max]		
			CMIP5	CMIP6	TRENDY-v7
Net sinks					
Vegetation uptake LRU approach applied to each model GPP using:	LRU_{ref}	Seibt et al. (2010)	-1324 [-2312,-982]	-1070 [-1342,-710]	-1206 [-1517,-984]
	LRU_{Whelan}	Whelan et al. (2018)	-757 [-1327, -566]	-615 [-766,-398]	-694 [-867,-573]
	LRU_{ORC}	Maignan et al. (2021) updated in this study	-709 [-1234,-525]	-573 [-724,-380]	-648 [-817,-520]
Soil exchanges	Mechanistic	Ogée et al. (2016); Abadie et al. (2022)	-39		
	Proxy-based	Launois et al. (2015); Whelan et al. (2013)	-340		
Oxidation by OH in the troposphere	-	Whelan et al. (2018)	-100		
Net sources					
Ocean emissions	Optimized ocean fluxes	Remaud et al. (2022)	+483		
	Box model for COS and CS_2, NEMO PISCES model for DMS	Lennartz et al. (2020); Launois et al. (2015)	+311		

Anthropogenic emissions	Inventory of spatially scaled sources	Zumkehr et al. (2018)	+397		
Biomass burning	Inventory using CO emissions as a reference gas	Stinecipher et al. (2019)	+53		
Net budget					
Total for the reference scenario			-530	-276	-412

3. Results

3.1. Seasonal cycle of atmospheric COS and CO_2 concentrations

3.1.1. Contribution of each flux component to the seasonal cycle

Before analyzing the simulated concentrations obtained from the different LSM fluxes (GPP and NBP), we first investigate the respective contribution of each flux component for both COS and CO_2 concentrations at two stations of the Northern hemisphere, MLO and ALT, using one model, ORCHIDEE, taken from the TRENDY-v7 ensemble. These contributions (see Figure 3, different color lines) correspond to the detrended mean seasonal cycle following the procedure described in §2.6.1.

For COS, we should first notice that the net simulated concentration (sum of all components, Figure 3 - black line) follows relatively closely the observed mean seasonal cycle at the two stations with a peak in late spring (May - June) and a minimum in September - October. The contribution from the vegetation (i.e., uptake of COS related to GPP) dominates and explains most of the seasonal variations of the net concentrations at both MLO and ALT. The other components, ocean, soil, and the combined anthropogenic plus biomass burning and OH sink, have a much smaller contribution, although not negligible. At MLO, the seasonal amplitude of the vegetation contribution is around 50 ppt, while it does not exceed 15 ppt for the three other components. At ALT, the vegetation contribution rises to nearly 60 ppt while the others contributions remain around 20-30 ppt. In terms of phase of the mean seasonal cycle, the vegetation contribution dominates the net signal and imposes the overall shape of the seasonal cycle. However, we notice that for instance at MLO the simulated net minimum occurs too early in September (compared to an observed minimum in November) because of the non-vegetation components and in particular the anthropogenic, biomass burning and OH sink contributions.

For CO_2, we also notice that the net simulated mean seasonal cycle is close to the observations at both stations. However, in this case the net signal is nearly equally influenced by the concentrations resulting from the vegetation uptake of CO_2 (GPP) and from the ecosystem respiration flux (release of CO_2). Both contributions lead to concentration variations in opposite phases with the GPP contribution dominating and thus imposing the phase of the net signal. The GPP-induced seasonal cycle amplitude is 18 and 40 ppm at MLO and ALT, respectively, while the respiration-induced amplitude is 12 and 30 ppm at MLO and ALT, respectively. The contribution from the ocean and fossil fluxes to the mean seasonal cycle is much smaller and nearly negligible at these two stations (less than 2 ppm), which is not the case for stations closer to large anthropogenic emissions in industrial or highly populated areas and in the Southern hemisphere where the ocean flux becomes a key component (not shown).

Overall, this preliminary analysis reveals the dominant contribution of GPP and its associated COS uptake for the seasonal cycle of COS and CO_2 atmospheric concentrations. But while for COS GPP is the main component of the mean seasonal cycle, for CO_2 the contribution of the ecosystem respiration is nearly as important, as already highlighted in (Launois et al., 2015). Therefore, this joint analysis evidences the potential information brought by atmospheric COS concentrations to constrain more specifically the simulated GPP of LSMs.

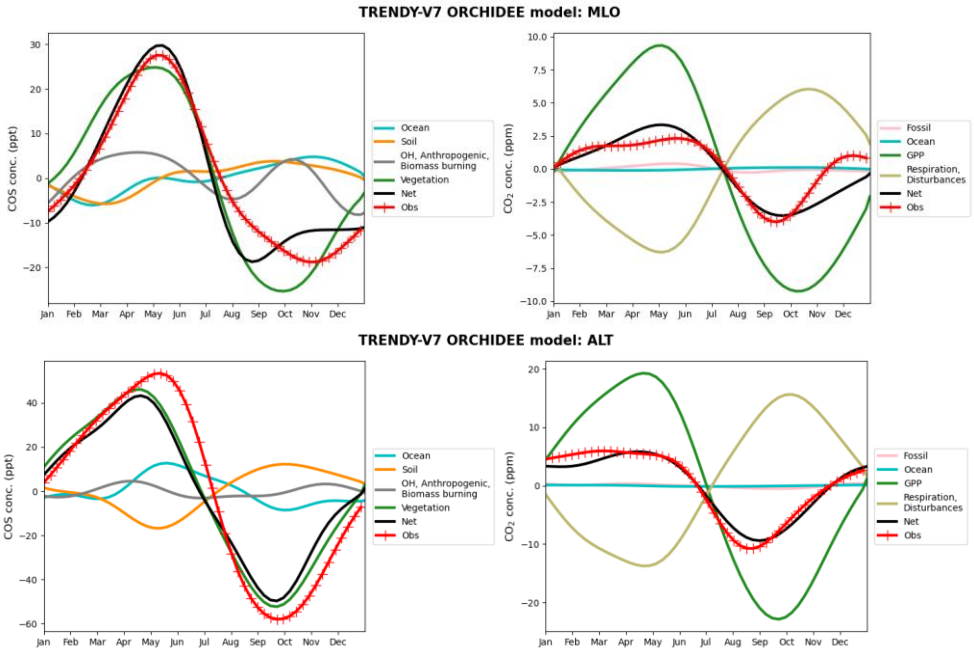

Figure 3: Smoothed seasonal cycles of atmospheric COS (left) and CO_2 (right) concentrations simulated at MLO (top) and ALT (bottom) for the ORCHIDEE model of TRENDY-v7. Color lines correspond to the transport of the individual flux components of the COS or CO2 budget. The net total considering all components of the COS or CO2 budget is represented in black. For each component the detrended mean seasonal cycle has been computed following the method defined in § 2.6.1.

3.1.2. Observed and simulated mean seasonal cycle

Mauna Loa station:

Figure 4 shows the mean seasonal cycle (average over 2010-2014) of modeled and observed monthly mean concentrations at the Mauna Loa (MLO) station for COS and CO2 (fig. 3 left and right), for the three ensembles of model simulations (24 CMIP5, 25 CMIP6 and 15 TRENDY models). Table A1 in the Appendix also provides the COS and CO_2 seasonal amplitude and the correlation with the observation, for each model. The concentrations were simulated from the GPP for COS and the NBP for CO_2 of the different models, following the methodology described in section 2, using the reference setup (see Table 1).

For COS, the modeled concentrations qualitatively follow the observations for the mean seasonal cycle, as noticed above, with a maximum late spring and a minimum in September - October. The amplitude of the mean seasonal cycle varies substantially between the different models of a given ensemble (Table A1). For CMIP5 it varies by a factor of three, between 29 ppt (CanESM2) and 94 ppt (MRI-ESM1), while the observations have an amplitude of 48 ppt. The range of model amplitudes is much more reduced in CMIP6 with amplitudes that are on average slightly lower than the observations, between 32 ppt (TaiESM1) and 57 ppt (MPI-ESM1-2LR) with the exception of NorCPM1 that has a much lower amplitude (21 ppt). For the TRENDY-v7 models, although they use the same

meteorological forcing, the amplitude also substantially varies, between 30 ppt (ORCHIDEE-CNP) and 74 ppt (LPX). The phase of the seasonal cycle shows for the three ensembles a relatively good agreement although with a too early summer draw down of the COS concentration at MLO and a minimum occurring 1 to 2 months earlier than in the observations. No significant differences in terms of phase (measured as the correlation between the models and the observation) occur between the 3 ensemble of simulations.

For CO_2, the modeled concentrations show also similar features, with maximum concentrations in late spring and minimum concentrations in autumn. Like for COS, the amplitude of the mean seasonal cycle substantially varies between the CMIP5 models (Table A1), from 3.1 ppm (INM-CM4) to 13 ppm (IPSL-CM5A-LR), while the observed amplitude is 6.8 ppm. As expected the models with a low/high COS amplitude also have a low/high CO2 amplitude. For the CMIP6 ensemble, the spread is also largely reduced compared to CMIP5 (between 3.2 and 9.0 ppm) with the exception (as for COS) of NorCPM1 (2.2 ppm). For the TRENDY-v7 models, the CO2 amplitudes span like for COS also a large range, between 1.9 ppm (CABLE-POP) and 7.7 ppm (SDGVM). For the phase of the mean seasonal cycle (defined as the correlation with the observations, see Table A1), the spread between the different models of each ensemble is relatively large and similar to that for COS.

Overall, the simulated concentrations at MLO reveal similar skills and biases for the COS and CO2 mean seasonal cycles simulated by a given model. Given that MLO integrates the signal of the fluxes from the whole Northern hemisphere (Buermann et al., 2007), these results point toward specific model GPP biases for the Northern hemisphere ecosystem fluxes.

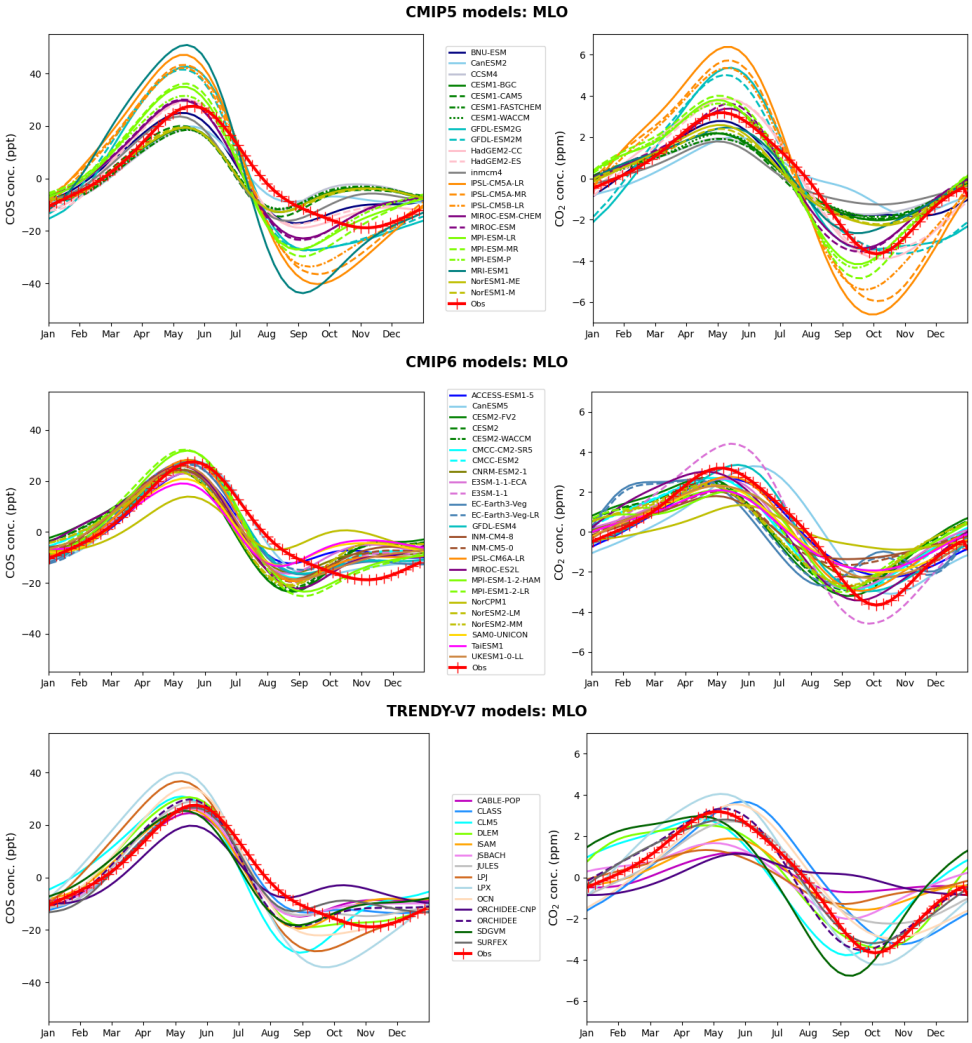

Figure 4: Detrended mean seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at Mauna Loa station (MLO) for all models of CMIP5, CMIP6 and TRENDY-v7 ensembles. The observations are represented by red crosses. The detrended mean seasonal cycles have been computed following the method defined in § 2.6.1.

Alert station:

We performed the same analysis at the Alert station in Canada (ALT) to focus on high latitude ecosystem fluxes (see figure 5 and table A1 for the amplitude and phase diagnostics).

Like for the MLO station, the modeled COS mean seasonal cycles qualitatively follow the observations. However, the maximum COS concentration is simulated one month earlier in spring for most models compared to the observations, while the minimum concentration is simulated around the observed one (September - October). The amplitude of the mean seasonal cycle varies drastically between the different models of a given ensemble (Table A1) and more than for the MLO station. For CMIP5 it varies by over a factor of four, between 43 ppt (CESM1-Fastchem, CCSM4) and 195 ppt (IPSL-CM5A-LR), while the observations have an amplitude of 114 ppt. The range of model amplitudes is again highly reduced in CMIP6 with amplitudes that are on average lower than the observations, between 45 ppt (TaiESM1) and 121 ppt (MPI-ESM1-2LR, MIP-ESM-1-2-Ham) with the exception of NorCPM1 (26 ppt) like for MLO. For the TRENDY-v7 models, the amplitude brackets the observed amplitude with variations between 68 ppt (ORCHIDEE-CNP) and 157 ppt (LPX). No significant differences in terms of correlation between the models and the observation occur between the 3 ensemble of simulations.

For CO2, the modeled mean seasonal cycles at ALT show a larger spread than at MLO for each model ensemble with several models having different seasonal variations than the observations. The amplitude of the mean seasonal cycles (Table A1) varies more across the CMIP5 models (between 9.2 ppm (INM-CM4) and 39.6 ppm (IPSL-CM5A-LR)) than the CMIP6 models (between 8.7 and 23.7 ppm with the exception, as for COS, of NorCPM1 (3.4 ppm)) and the TRENDY-v7 models (between 5.2 and 23.8 ppm). Compared to the MLO station, the spread of the simulated CO2 mean seasonal cycles is larger with the timing of maximum or minimum concentrations that can vary by more than two months between models.

Overall, the phase is on average less well captured for CO2 than it is for COS, at both MLO and ALT stations. The biases observed for CO2 (amplitude and or phase) correspond to a certain extent to similar biases for COS, as expected from the methodology used to derive the COS uptake from the vegetation (proportional to the GPP; see method in section 2).

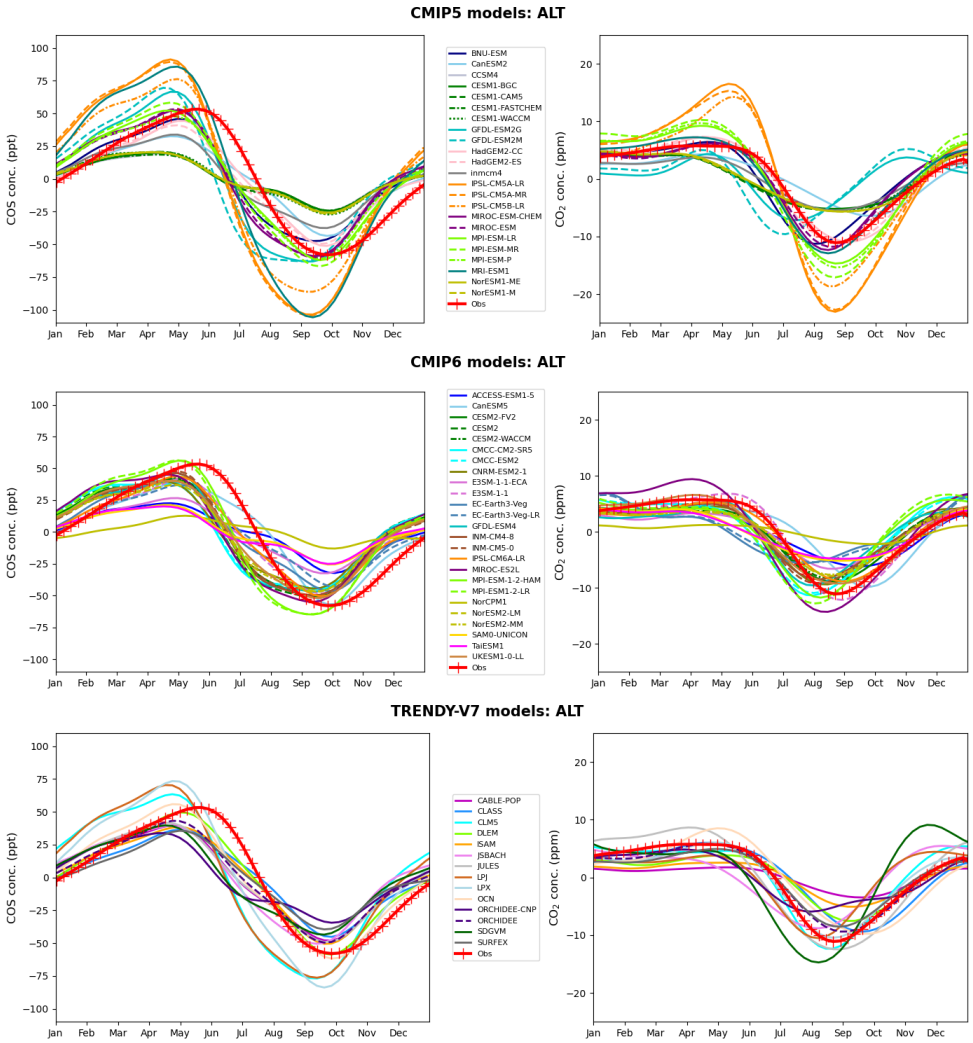

Figure 5: Detrended mean seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at Alert station (ALT) for all models of CMIP5, CMIP6 and TRENDY-v7 ensembles. The observations are represented by red crosses. The detrended mean seasonal cycles have been computed following the method defined in § 2.6.1.

3.2. Joint analysis of COS and CO_2 seasonal amplitudes and phases

3.2.1. Simulations with reference scenario

The potential GPP and respiration biases of the CMIP5/6 and TRENDY-v7 models are assessed by comparing the amplitude and phase of the normalized simulated CO_2 versus COS concentrations (see method in section 2.6.1).

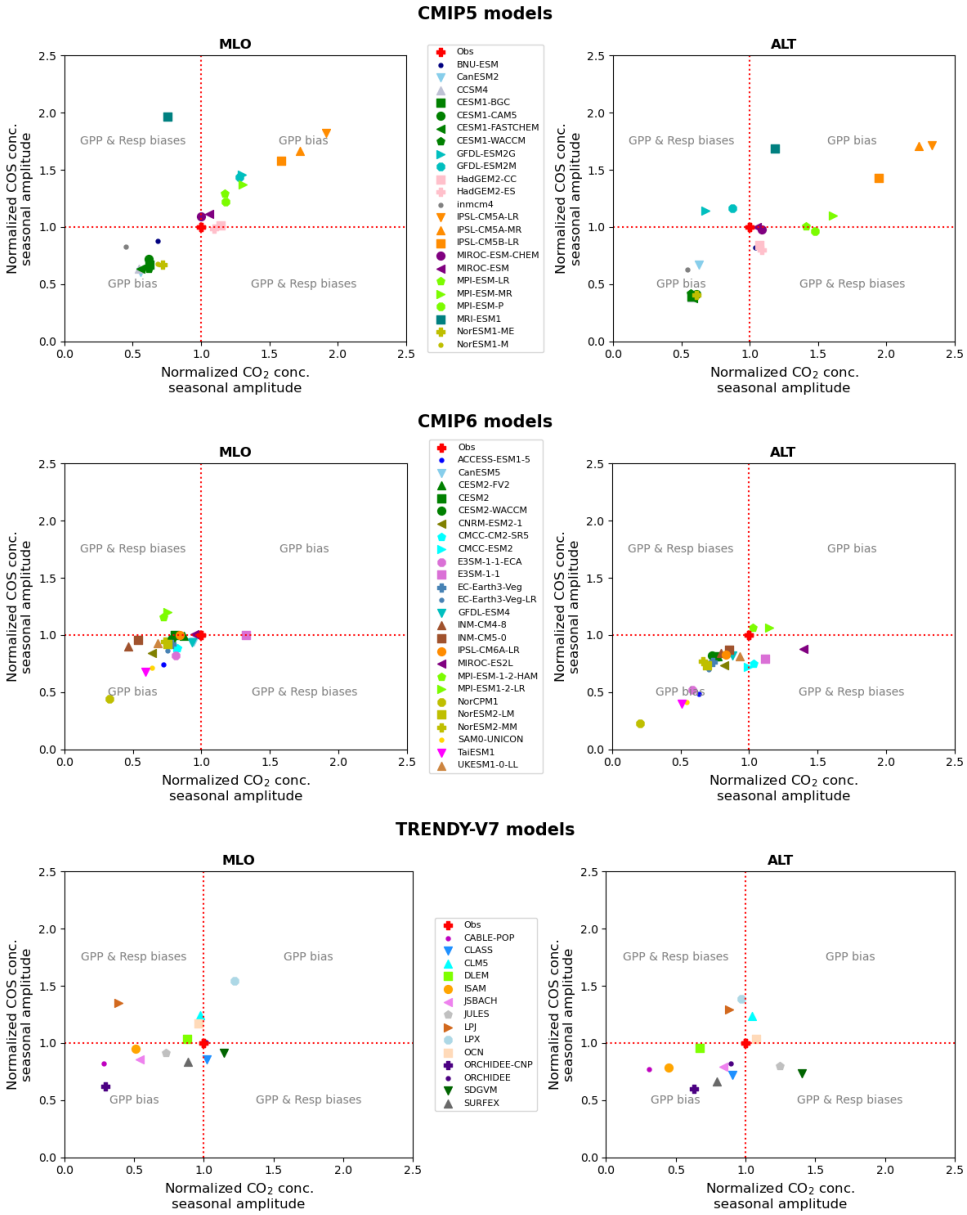

Figure 6: Scatter plots of CO_2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO and ALT for CMIP5, CMIP6 and TRENDY models using the reference scenario. The observed amplitude is identified by the red cross.

Figure 6 shows the evaluation of simulated CO_2 versus COS concentration seasonal amplitudes at MLO and ALT across the three model ensembles, normalized by the observed amplitudes. CMIP5 models exhibit a positive linear relationship between the normalized simulated CO_2 and COS concentration seasonal amplitudes at the two stations. This alignment indicates that an underestimation or an overestimation of CO_2 concentration seasonal amplitudes (compared to the observation) is accompanied by a similar deviation for COS. Because the simulated COS concentration amplitudes are driven by the amplitude of GPP through the LRU relationship (see section 3.1.1 for the contribution of each flux component), this highlights that the main biases arise from the simulated GPP. Notably, all CMIP5 models of the IPSL (IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR) consistently overestimate both CO_2 and COS concentration seasonal amplitudes at the two stations, while the CESM1, NorESM1, CCSM4, CanESM2, and inmcm4 models consistently underestimate them. Importantly, note however that this positive linear relationship passes

through the observational target (red cross), demonstrating that certain models succeed in representing NOAA CO_2 and COS concentration seasonal amplitudes, with the chosen set up for the vegetation COS uptake and the other COS components. Indeed, the seasonal amplitudes of CO_2 and COS concentrations for the MIROC-ESM and HadGEM2 models match the ones of the observations, supporting an accurate simulation of the mean seasonal cycle of GPP over the Northern hemisphere. Note that one model (MRI-ESM1) diverges from the positive linear distribution at MLO in particular, with a normalized COS amplitude significantly too large (1.98) and a normalized CO_2 amplitude too low (0.75). This suggests a bias in both GPP and respiration fluxes as an increase in GPP seasonal amplitude only to better match the observed CO_2 concentration amplitude would lead to a further degradation of the one of COS. At ALT, the two GFDL models also slightly deviate from the linear trend with an overestimation of COS concentration amplitudes but not of CO_2, possibly highlighting both GPP and respiration biases.

In contrast to CMIP5, CMIP6 models show a more restricted range of normalized seasonal amplitudes with reduced values, as explained in the previous section. At ALT station, we still see a positive linear distribution between the normalized seasonal amplitudes of CO_2 and COS concentrations, while at MLO the relationship appears much weaker (partly because of the reduced amplitude spread). Notably, no models show a strong overestimation of CO_2 or COS concentration seasonal amplitudes, except MIROC-ES2L and E3SM-1-1 for CO2 at ALT and MLO, respectively. This indicates significant improvements in CMIP6 compared to CMIP5, with respect to large positive GPP biases found in CMIP5, as clearly illustrated with the IPSL and MPI-ESM models. On the other hand, most CMIP6 models now tend to display an underestimation of both COS and CO_2 concentration seasonal amplitudes, especially for NorCPM1, TaiESM1, SAM0-UNICON, and ACCES-ESM-1. For these models, a GPP seasonal cycle bias could explain the similar underestimation of CO_2 and COS concentration seasonal amplitudes.

While CMIP6 models show an overall improvement over CMIP5 for both tracers, the TRENDY-v7 experiment reveals increased model spread around the observation target and a lower correlation between the two normalized seasonal amplitudes (especially at ALT), despite being driven by the same forcing data. Most TRENDY-v7 models (8 out of 14 at MLO and 10 out of 14 at ALT) underestimate both CO_2 and COS concentration seasonal amplitudes, as noticed for ORCHIDEE, CABLE-POP, JSBACH, or ISAM, likely due to potential low-GPP seasonal amplitude biases for the Northern extra tropics ecosystems. Three models noticeably differ with an overestimation of the COS normalized seasonal amplitude and an underestimation of the CO_2 one, for LPJ at MLO, and the opposite pattern for SDGVM and JULES at ALT, highlighting potential biases in both GPP and respiration fluxes.

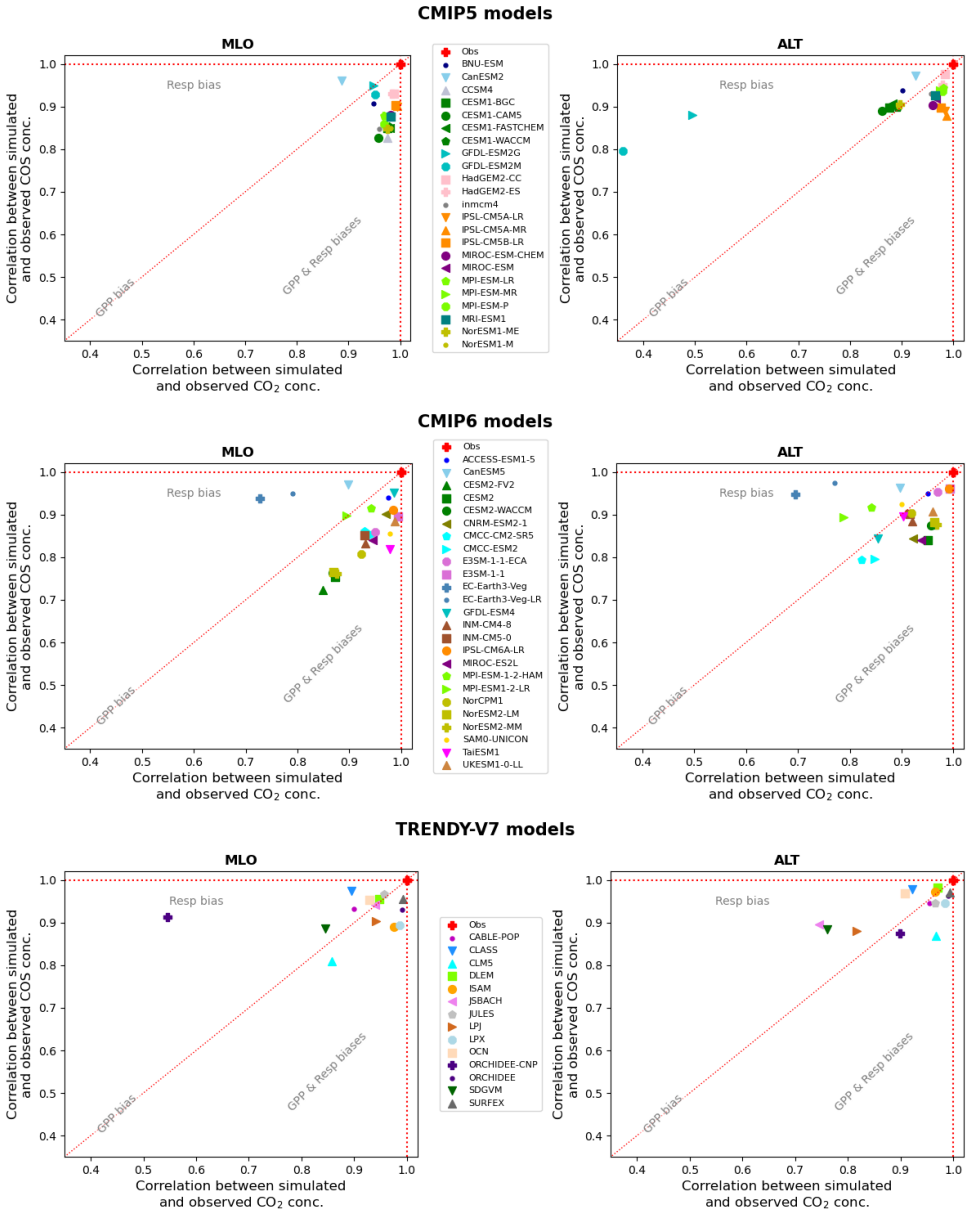

Figure 7: Scatter plots of CO_2 vs COS phase correlations at MLO and ALT simulated by the CMIP5, CMIP6 and TRENDY models. The correlations for each constituent was computed between modeled and observed smoothed seasonal concentrations. A correlation of 1 shows a good representation of reality. The observed phase is identified by the red cross.

Figure 7 illustrates the evaluation of the phase (i.e., correlation analysis) of the simulated concentration mean seasonal cycles for CMIP5/6 and TRENDY-v7 experiments at MLO and ALT. CMIP5 models exhibit strong correlations between simulated and observed concentrations for both COS and CO_2, indicating an accurate representation of the observed mean seasonal cycle phase. Notably, excluding the two GFDL models, the lowest correlations stand at 0.86 and 0.80 for CO_2 and COS, respectively, across the two stations (Table A1). Despite demonstrating relatively high correlations for COS at ALT, the two GFDL models show a poor correlation with observed CO_2 concentrations, with values of 0.50 for GFDL-ESM2G and 0.36 for GFDL-ESM2M (Table A1). This low performance at ALT is related to a too early decrease of the simulated CO_2 concentrations compared to NOAA observations, followed by an increase in concentrations during August and September not seen in the observations (Figure 5). This discrepancy in the seasonal cycle phase for

CO_2 concentrations points toward a potential bias in the simulated respiration, given the accurate representation of the GPP phase as suggested by the agreement in the COS concentration seasonal cycle.

Compared to CMIP5, CMIP6 models display an overall slightly lower performance in capturing the observed CO_2 and COS concentration seasonal cycle phases, with correlations deviating further from the target (red cross). Therefore, the improvement in simulated seasonal amplitudes (Figure 6) between CMIP5 and CMIP6 may have been counterbalanced by a slight decline in simulated phase accuracy or this may highlight issues with the other flux components. Among CMIP6 models, correlation values between simulated and observed CO_2 concentrations range between 0.70 and 0.99, and between 0.72 and 0.98 for COS (Table A1). In particular at MLO, the CESM2 and NorESM models exhibit a degradation of both simulated COS and CO_2 concentration seasonal cycle phases. However, there is a notable improvement in the simulated CO_2 concentration seasonal cycle phase at ALT for the GFDL model included in the CMIP6 experiment compared to CMIP5. Note that the EC-Earth3-Veg models show significantly lower correlations for CO_2 than COS concentrations at both stations, suggesting potential biases in simulated respiration.

The correlation distribution among TRENDY-v7 models appears narrower compared to CMIP6, yet it remains broader than CMIP5 (if we exclude the two GFDL models), despite all TRENDY-v7 models being forced with the same climate data. Excluding ORCHIDEE-CNP at MLO, correlations in the TRENDY-v7 experiment range from 0.75 to 0.99 for CO_2 concentrations and from 0.81 to 0.98 for COS across the two stations (Table A1). The particularly poor performance of ORCHIDEE-CNP in simulating CO_2 concentration seasonal cycle phase at MLO suggests a respiration bias. Finally, note that at MLO in particular, while most CMIP5 and CMIP6 models exhibit slightly lower correlations for COS compared to CO_2, this contrast is not found in the TRENDY-v7 model ensemble.

3.2.2. Sensitivity analysis

Sensitivity to LRU:

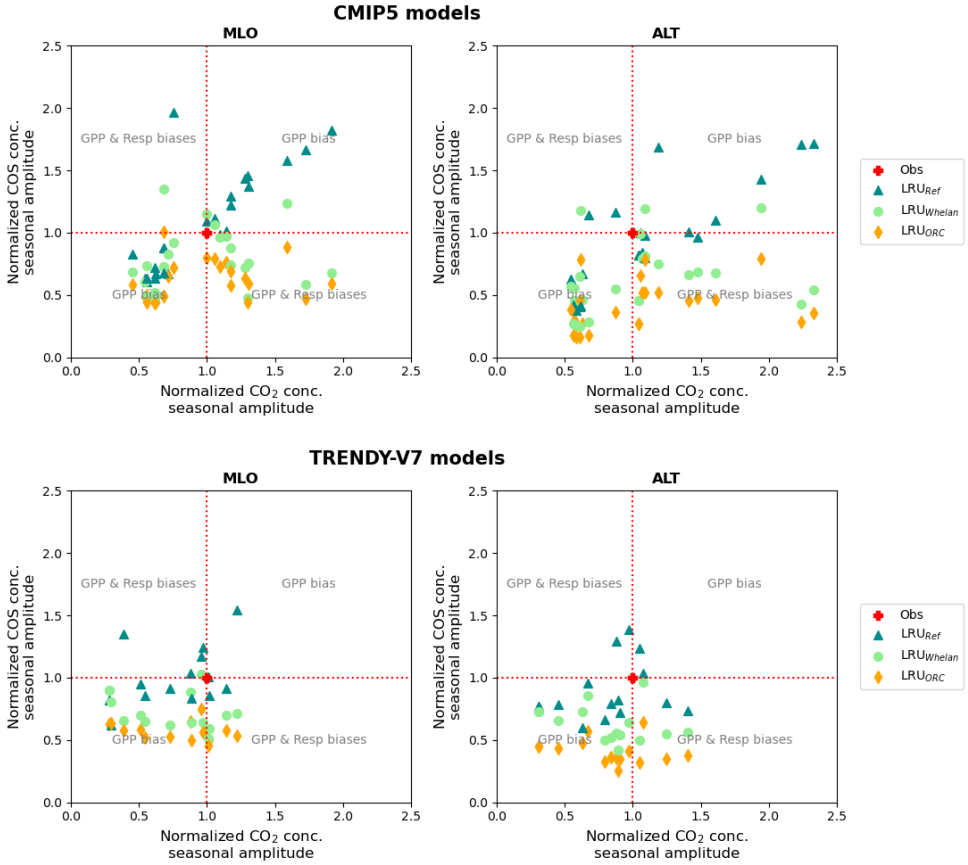

Figure 8: Scatter plots of CO_2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO and ALT for CMIP5 and TRENDY models. The different LRU scenarios are Seibt Reference LRU, ORCHIDEE LRU and Whelan LRU. The observed phase is identified by the red cross.

Figure 8 shows the sensitivity of the normalized simulated CO_2 and COS concentration seasonal amplitudes to the set of LRU values considered to compute vegetation COS flux from GPP, for the CMIP5 and TRENDY-v7 experiments at MLO and ALT (see Appendix D for CMIP6). For CMIP5 models, using LRU_{Whelan} or LRU_{ORC} breaks the strong linear positive relationship between normalized simulated CO_2 and COS concentration seasonal amplitudes obtained with LRU_{ref} at the two stations. Furthermore across all CMIP5 and TRENDY-v7 models, using the much lower LRU_{Whelan} or LRU_{ORC} values (see Table 2) decreases COS concentration seasonal amplitudes compared to the reference case (LRU_{ref}); the greatest reduction occurs for models that simulate the strongest CO_2 concentration seasonal amplitude, as expected. Indeed, using lower LRU values results in a reduced vegetation COS uptake according to equation 1, and the seasonal amplitude of the vegetation COS flux drives primarily the COS concentration seasonal amplitude as illustrated in Figure 3. Therefore in both experiments, when using sets of lower LRU values than LRU_{ref}, the majority of models now underestimate the COS concentration seasonal amplitude, with only 4 models in CMIP5 at the two stations and 1 model in TRENDY-v7 at MLO overestimating it. On average across all models, adopting lower LRU values than the reference fails to replicate the NOAA COS seasonal amplitude, with the lowest values from LRU_{ORC} resulting in the largest deviation from the target. This

underestimation of COS concentration seasonal amplitude with LRU_{Whelan} and LRU_{ORC} would support the use of higher LRU values as in LRU_{ref} to compute vegetation COS uptake from GPP, if we assume that the other COS components are well simulated).

Sensitivity to soil fluxes:

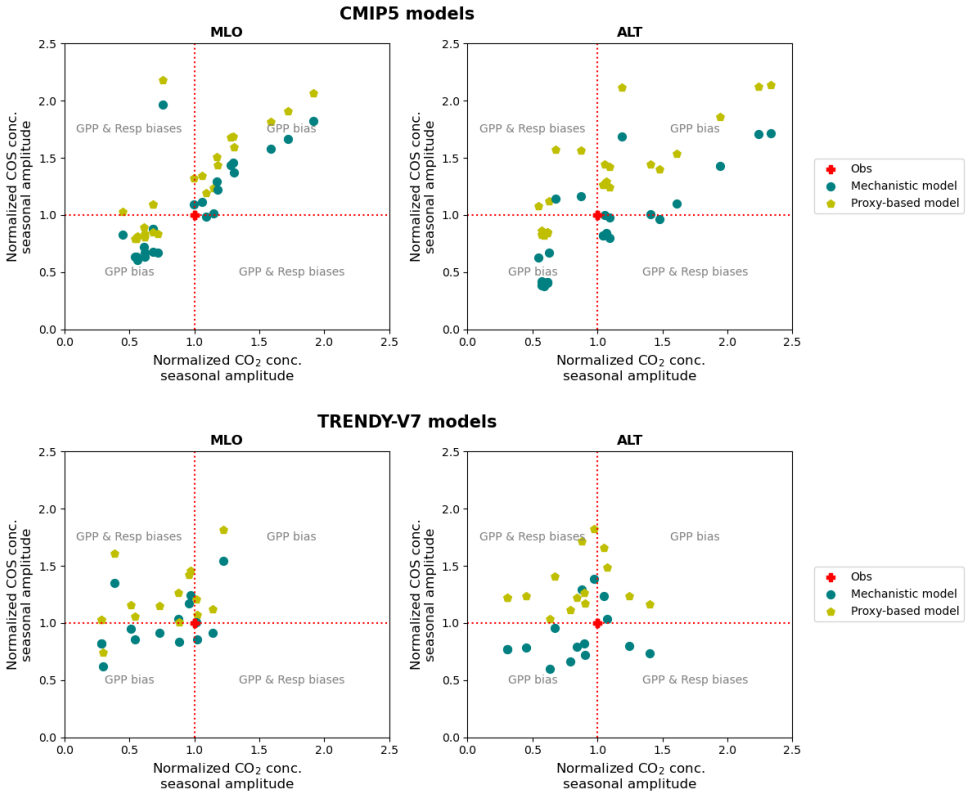

Figure 9: Scatter plots of CO_2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at ALT and MLO for CMIP5 and TRENDY models. The simulated COS concentrations use either the mechanistic model or the proxy-based approach for soil COS fluxes. The observed amplitude is identified by the red cross.

Figure 9 presents the impacts of the approach chosen to compute soil COS fluxes on the normalized simulated CO_2 and COS concentration seasonal amplitudes for the CMIP5 and TRENDY-v7 experiments at the two stations. The proxy-based approach (see method) leads to a net annual soil COS sink more than 8 times higher than the one obtained with the mechanistic approach used as the reference (Table 3). Given the seasonal variations of soil COS fluxes, partly controlled by temperature (see figure 3 for the mechanistic approach), using the proxy-based approach leads also to an increase of the amplitude of the total COS concentration seasonal cycle, compared to the use of the mechanistic approach. We see an upward shift of all points in figure 9 when going from the mechanistic to proxy-based approach, leading to an overestimation of COS seasonal amplitude with the proxy-based model, for at least 15 out of the 23 CMIP5 models, and 13 out of the 14 TRENDY-v7 models across the two stations. This may suggest that the mechanistic approach leads to more accurate soil COS fluxes, but more importantly it reveals that our combined COS - CO_2 analysis depends partly on the uncertainty in the simulated soil COS flux, not the annual mean flux but the amplitude of the seasonal cycle.

Sensitivity to ocean fluxes:

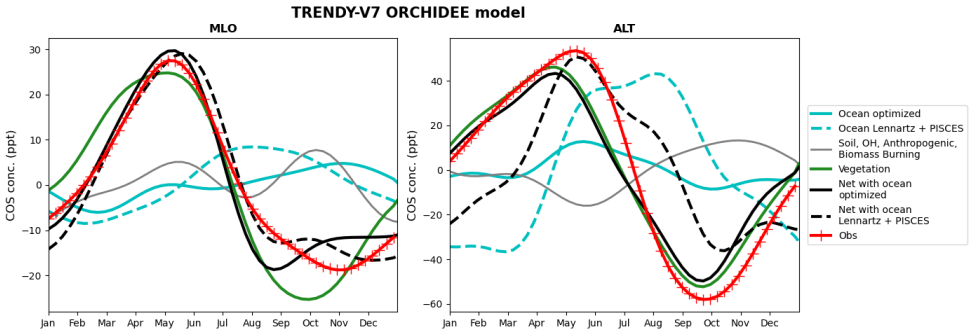

Figure 10: Smoothed seasonal cycles of atmospheric COS concentrations simulated at MLO (left) and ALT (right) for the ORCHIDEE model of TRENDY-v7. The concentrations in colors are obtained by transporting separately the different flux components of the COS budget. The solid colored lines represent the components of the standard scenario, while the dashed blue line represents the ocean component estimated following Lennartz et al. (2017, 2020) for direct COS emissions and indirect emissions from CS_2, and the NEMO-PISCES model estimate for indirect emissions from DMS (Launois et al., 2015) (Lennartz + PISCES). The net total considering all components of the COS budget is represented in black with a solid line for the standard scenario and with a dashed line for the scenario with the ocean Lennartz + PISCES.

Figure 10 shows the COS concentration seasonal cycles following the transport of the individual components of the COS budget, considering either the reference ocean flux ("Ocean optimization" based on Remaud et al. 2022) or the Lennartz-PISCES approach ("Ocean Lennartz + PISCES"), as well as the resulting net seasonal cycle from the transport of all components for these two scenarios at MLO and ALT. At both stations, the COS concentration seasonal cycle associated with the Lennartz-PISCES ocean flux depicts a larger seasonal cycle especially at ALT station, with lower COS concentration in winter (from October to May) and higher concentration in summer compared to the reference ocean contribution. At ALT, the transport of the Lennartz-PISCES ocean flux results in a COS concentration seasonal amplitude of approximately 75 ppt, compared to 20 ppt for the one of the reference ocean flux. While the reference ocean flux has a larger annual COS source (+483 GgS/y) compared to the Lennartz-PISCES approach (+311 GgS/y) (Table 3), the latter method induces a more pronounced seasonality in oceanic COS emissions. The larger COS concentration seasonal amplitude derived from the Lennartz-PISCES ocean flux thus competes with the one induced by the vegetation component (amplitude of about 100 ppt with the ORCHIDEE model from TRENDY-v7) given that they are out of phase by 2-3 months. Consequently, changing the ocean flux component significantly modifies the phase and amplitude of the net seasonal cycle of COS concentrations, especially at ALT and to a small extent at MLO. Figure E3 in the appendix illustrates the change in amplitude of the mean seasonal cycle for COS due to changing the ocean flux component. From the results at ALT station, we clearly see that the reference ocean component ("ocean optimisation") leads to a COS atmospheric seasonal cycle more in line with the observation than the "Lennartz + PISCES" ocean flux. This suggests that the large source of COS at high latitude from direct oceanic COS emissions and indirect emissions from CS_2 estimated by Lennartz et al., 2017, 2020 is too large and not compatible with the temporal dynamics of atmospheric COS concentrations at high latitudes. Remaud et al., 2023 did the same analysis but with several atmospheric transport models within the framework of the transcom-COS model intercomparison experiment, aiming at evaluating the transport errors. The analysis led to the same findings.

4. Discussion and conclusion

The joint atmospheric COS and CO_2 analysis presented above in details for two stations of the Northern hemisphere (MLO, ALT) helped us to identify potential biases in the simulated GPP and TER of individual LSM from three ensembles (CMIP5, CMIP6, TRENDY-v7) as well as specific features of the ensemble themselves. In particular, we noticed:

- The CMIP5 model ensemble provides a large spread of the amplitude of the mean seasonal cycle of COS. This spread is directly linked to the large spread of the simulated GPP seasonal cycles across models. With the selected COS set up (i.e., the chosen LRU values to link GPP and plant COS uptake, and the other COS flux components), the simulated COS and CO_2 normalized amplitudes follow a linear relationship that includes the observation. This highlights the strong and direct control of atmospheric COS and CO_2 seasonal amplitudes by GPP, also evidencing that at least some LSMs may accurately represent the seasonal cycle of GPP for the Northern extra-tropics ecosystems. Reversely and more interestingly, this diagnostic points towards specific model biases with too large or too low GPP seasonal amplitudes.
- The CMIP6 ensemble shows a significant reduction of the COS and CO_2 concentration seasonal amplitude spread at MLO and ALT, which indicates that the models have converged between the CMIP5 and CMIP6 exercises. The COS seasonal amplitudes appear to be slightly too low across all models for CMIP6, which may suggest that i) other components of the COS budget are not contributing enough to the overall mean seasonal cycle amplitude and/or ii) the LRU values are too small (eq. 1) and/or iii) the simulated GPP is underestimated in most CMIP6 models as the CO_2 seasonal amplitude is also underestimated on average.
- The TRENDY-v7 ensemble, despite using the same meteorological forcing (climate reanalysis), shows also a large spread of atmospheric COS and CO_2 seasonal amplitudes across models. The linear relationship between the COS and CO_2 normalized amplitude is weaker at ALT compared to CMIP5 and CMIP6 ensembles but still includes the observation.
- The analysis of the phase of the mean seasonal cycle of atmospheric COS and CO_2 also allows to highlight phase issues with respect to GPP and/or TER for specific models of each ensemble.

Through this paper we thus demonstrate the potential of atmospheric COS concentration measurements, combined with atmospheric CO_2 data, to evaluate the GPP and TER simulated by LSMs. The detailed analysis was made at only two atmospheric stations (MLO and ALT), but the same diagnostics are plotted at 8 other stations of the NOAA network (see figures B1-8 in appendix B). BRW station provides similar information on the potential model biases for the arctic ecosystems than ALT, while KUM and MHD stations provide similar features than MLO station. LEF and NWR stations provide CO_2 specific diagnostics link to North American ecosystem fluxes (GPP and TER). The large CO_2 seasonal amplitude observed at LEF is not captured by most models while the COS amplitude is well reproduced (Figure B4), which may indicate too large LRU values combined with too low GPP for North American ecosystems, especially for the great plains that are upwind of this station. For the tropical station SMO, all models of the three ensembles do not capture well the seasonal cycle of COS (Figure B7), which may reflect biases in the ocean COS flux, while for CO_2 we obtain a large model spread that includes the observations. Such spread reflects the large uncertainties associated with GPP and TER fluxes of tropical ecosystems, but also the difficulties of a climate model to simulate the temporal dynamics of the inter-tropical convergence zone (ITCZ) for the CMIP ensembles. In the Southern hemisphere, the CGO station (Figure B2) shows a greater coherence for COS seasonal cycle across the different models of each ensemble than for CO_2. At CGO the CO_2 seasonal cycle depicts a small amplitude not in phase with the COS cycle, and the large spread obtained for CO_2 at this station suggests potential biases in the seasonal variation of the simulated respiration fluxes. At SPO (Figure B8), the diagnostic is relatively similar to CGO.

Overall, at all stations we see a clear improvement of the simulated CO_2 fluxes in the CMIP6 ensemble compared to CMIP5, as also pointed out by Hu et al. (2022) .

However, the COS diagnostic depends on the overall set up chosen for the computation of the different COS fluxes and especially the LRU values linking plant COS uptake to GPP. The values that we used for our reference set up (Seibt et al., 2010) are considered to be too high compared to other estimates (Whelan et al., 2018; Maignan et al., 2021; Wohlfahrt et al., 2023), especially in order to close the overall atmospheric COS budget (not closed with our reference set up; see table 3). Considering lower LRU values, would indeed enable a more balanced COS budget by reducing the vegetation sink component. However, it would substantially reduce the seasonal amplitude of the simulated COS concentrations (see section 3.2.2, Figure 8), with the risk that most models do not capture the observed amplitude. We argue that this is unlikely, and if LRU values are smaller, other COS processes should compensate with a more pronounced seasonal cycle in phase with that of GPP and/or atmospheric mixing would need to be revised. In addition, the method that we used with constant LRU values has some limitations, given that the COS relative uptake varies with atmospheric conditions such as light radiation (Kooijmans et al. 2019, Maignan et al., 2021; Wohlfahrt et al., 2023). However, with the ORCHIDEE model, Maignan et al. (2021) derived monthly LRU estimates that are much lower than our reference values (see Table 3) with only relatively small seasonal variations (around 0.2). Using monthly-variable LRU values for each PFT should thus be considered as an important direction to improve the COS diagnostic.

In addition, we have seen significant impacts of COS components beyond plant uptake on the total COS concentrations, like soil and ocean fluxes. The sensitivity tests that were conducted have a large impact on the total annual atmospheric COS budget for both components as summarized in Table 3 (i.e., considering -39 vs -340 GgS/yr for the two soil variants and 483 vs 311 GgS/yr for the two ocean variants). However, the impact on the seasonal amplitude is relatively low for the soil component. For the ocean, using the “lennard+pisces” flux instead of the “ocean optimisation" flux (our reference) significantly biases the COS seasonal cycle at Northern high latitudes stations (Figure 10). A key priority is thus to further understand and quantify the seasonality of the ocean COS source as well as of the soil COS uptake. Note finally that the annual budget of the ocean flux is still highly uncertain and remains a key component to close the overall atmospheric COS budget (Berry et al., 2013; Remaud et al., 2022, Remaud et al., 2023).

Last but not least, is the uncertainty associated with the atmospheric transport model used to compute COS and CO_2 concentrations. Remaud et al. (2023) compared the impact of different atmospheric transport models on the simulated COS concentrations, using a common set of COS surface fluxes. They show (see Figure 5 of Remaud et al., 2023) significant impact for high latitudes stations like ALT, BRW and MHD with the most recent version of the LMDZ transport model (V6 compared to V3 used in our study) belonging to a group of three atmospheric transport models producing larger amplitudes than the other group of three models. This pushes for additional studies with different atmospheric transport models, having different vertical and horizontal mixing properties. However, at MLO station, all transport models produce a similar COS amplitude, which reinforces the validity of the diagnostics that we made for the different LSMs, at least on average for the Northern ecosystem flux estimates.

Therefore, this study has allowed a first quantitative assessment of the mean seasonal cycle of GPP and TER fluxes simulated by different LSMs (mainly the amplitude and to a lower extent the phase) using atmospheric COS and CO_2 data. Several directions of research to improve the proposed diagnostics have been highlighted, including the need to refine the temporal and spatial variations of LRU, to reduce the uncertainty on COS flux components other than vegetation uptake, and to

use possibly several transport models for a more robust diagnostic of GPP and TER biases. We also encourage LSM modeling teams to implement a mechanistic model for vegetation COS uptake (as in Maignan et al., 2021 and Kooijmans et al., 2021) to better link GPP and COS plant uptake, accounting more directly for the impact of environmental drivers on the COS diffusion pathways into leaves.

Acknowledgement

The concentration measurements from the NOAA network have been performed by scientists affiliated with NOAA (Stephen Montzka).

References

Abadie, C., Maignan, F., Remaud, M., Kohonen, K.-M., Sun, W., Kooijmans, L., Vesala, T., Seibt, U., Raoult, N., Bastrikov, V., Belviso, S., & Peylin, P. (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(7), e2023JG007407.

Abadie, C., Maignan, F., Remaud, M., Ogée, J., Campbell, J. E., Whelan, M. E., Kitz, F., Spielmann, F. M., Wohlfahrt, G., Wehr, R., Sun, W., Raoult, N., Seibt, U., Hauglustaine, D., Lennartz, S. T., Belviso, S., Montagne, D., & Peylin, P. (2022). Global modelling of soil carbonyl sulfide exchanges. *Biogeosciences* , 19(9), 2427–2463.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., & Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics* , 53(3), 785–818.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., & Zhu, Z. (2013). Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models. *Journal of Climate*, 26(18), 6801–6843.

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.

Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*. https://doi.org/10.1126/sciadv.1602244

Bhandari, A. K., Kumar, A., & Singh, G. K. (2012). Feature extraction using Normalized Difference Vegetation Index (NDVI): A case study of Jabalpur city. *Procedia technology*, *6*, 612-621.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., … Papale, D. (2010). Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, 329(5993), 834–838.

Belviso, S., Pison, I., Petit, J. E., Berchet, A., Remaud, M., Simon, L., ... & Lopez, M. (2023). The Z-2018 emissions inventory of COS in Europe: A semiquantitative multi-data-streams evaluation. *Atmospheric Environment*, *300*, 119689.

Belviso, S., Nguyen, B. C., & Allard, P. (1986). Estimate of carbonyl sulfide (OCS) volcanic source strength deduced from OCS/CO2 ratios in volcanic gases. *Geophysical Research Letters*, 13(2), 133–136.

Berkelhammer, M., Asaf, D., Still, C., Montzka, S., Noone, D., Gupta, M., Provencal, R., Chen, H., & Yakir, D. (2014). Constraining surface carbon fluxes using in situ measurements of carbonyl sulfide and carbon dioxide. *Global Biogeochemical Cycles*, 28(2), 161–179.

Berry, J. (2018). Solar induced chlorophyll fluorescence: Origins, relation to photosynthesis and retrieval. *Compr. Remote Sens*, *3*, 143-162.

Berry, J., Wolf, A., Elliott Campbell, J., Baker, I., Blake, N., Blake, D., Scott Denning, A., Randy Kawa, S., Montzka, S. A., Seibt, U., Stimler, K., Yakir, D., & Zhu, Z. (2013). A coupled model of the global cycles of carbonyl sulfide and CO_2: A possible new window on the carbon cycle. *Journal of Geophysical Research: Biogeosciences*, 118(2), 842–852.

Buermann, W., Lintner, B. R., Koven, C. D., Angert, A., Pinzon, J. E., Tucker, C. J., & Fung, I. Y. (2007). The changing carbon cycle at Mauna Loa Observatory. *Proceedings of the National Academy of Sciences*, 104(11), 4249–4254.

Campbell, J. E., Whelan, M. E., Seibt, U., Smith, S. J., Berry, J. A., & Hilton, T. W. (2015). Atmospheric carbonyl sulfide sources from anthropogenic activity: Implications for carbon cycle constraints. *Geophysical research letters*, *42*(8), 3004-3010.

Campbell, J. E., Carmichael, G. R., Chai, T., Mena-Carrasco, M., Tang, Y., Blake, D. R., Blake, N. J., Vay, S. A., Collatz, G. J., Baker, I., Berry, J. A., Montzka, S. A., Sweeney, C., Schnoor, J. L., & Stanier, C. O. (2008). Photosynthetic Control of Atmospheric Carbonyl Sulfide During the Growing Season. *Science*. https://doi.org/10.1126/science.1164015

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.

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). *Biogeosciences* , 20(13), 2573–2594.

Commane, R., Meredith, L. K., Baker, I. T., Berry, J. A., Munger, J. W., Montzka, S. A., Templer, P. H., Juice, S. M., Zahniser, M. S., & Wofsy, S. C. (2015). Seasonal fluxes of carbonyl sulfide in a midlatitude forest. *Proceedings of the National Academy of Sciences*, 112(46), 14162–14167.

Damm, A., Elbers, J., Erler, A., Gioli, B., Hamdi, K., Hutjes, R., Kosvancova, M., Meroni, M., Miglietta, F., Moersch, A., Moreno, J., Schickling, A., Sonnenschein, R., Udelhoven, T., Van Der LINDEN, S., Hostert, P., & Rascher, U. (2010). Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). *Global Change Biology*, 16(1), 171–186.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C., Hauck, J., ... & Zheng, B. (2023). Global carbon budget 2023. *Earth System Science Data*, *15*(12), 5301-5369.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., ... Zaehle, S. (2019). Global Carbon Budget 2019. *Earth System Science Data*, 11(4), 1783–1838.

Gier, B. K., Schlund, M., Friedlingstein, P., Jones, C. D., Jones, C., Zaehle, S., & Eyring, V. (2024). Representation of the Terrestrial Carbon Cycle in CMIP6. *EGUsphere*, 1–63.

Glatthor, N., Höpfner, M., Baker, I. T., Berry, J., Campbell, J. E., Kawa, S. R., Krysztofiak, G., Leyser, A., Sinnhuber, B.-M., Stiller, G. P., Stinecipher, J., & von Clarmann, T. (2015). Tropical sources and sinks of carbonyl sulfide observed from space. *Geophysical Research Letters*, 42(22), 10,082–10,090.

Harris, I. (2019). CRU JRA: Collection of CRU JRA forcing datasets of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data. *Centre for Environmental Data Analysis (CEDA).*

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J. L., ... & Lott, F. (2006). The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, *27*, 787-813.

Hu, Q., Li, T., Deng, X., Wu, T., Zhai, P., Huang, D., ... & Qin, Z. (2022). Intercomparison of global terrestrial carbon fluxes estimated by MODIS and Earth system models. *Science of the Total Environment*, *810*, 152231.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., & Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, *Earth Syst. Sci. Data*, 11, 959–1002, https://doi.org/10.5194/essd-11-959-2019, 2019.

Jones, M. W., Andrew, R. M., Peters, G. P., Janssens-Maenhout, G., De-Gol, A. J., Ciais, P., Patra, P. K., Chevallier, F., & Le Quéré, C.: Gridded fossil CO2 emissions and related O2 combustion consistent with national inventories 1959–2018, *Sci. Data*, 8, 2, https://doi.org/10.1038/s41597-020-00779-6, 2021.

Kettle, A. J., Kuhn, U., von Hobe, M., Kesselmeier, J., & Andreae, M. O. (2002). Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks. *Journal of Geophysical Research, D: Atmospheres*, 107(D22), ACH 25–1 – ACH 25–16.

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene–HYDE 3.2. *Earth System Science Data*, *9*(2), 927-953.

Kooijmans, L. M. J., Cho, A., Ma, J., Kaushik, A., Haynes, K. D., Baker, I., Luijkx, I. T., Groenink, M., Peters, W., Miller, J. B., Berry, J. A., Ogée, J., Meredith, L. K., Sun, W., Kohonen, K.-M., Vesala, T., Mammarella, I., Chen, H., Spielmann, F. M., … Krol, M. (2021). Evaluation of carbonyl sulfide biosphere exchange in the Simple Biosphere Model (SiB4). *Biogeosciences* , 18(24), 6547–6565.

Kooijmans, L. M. J., Sun, W., Aalto, J., Erkkilä, K.-M., Maseyk, K., Seibt, U., Vesala, T., Mammarella, I., & Chen, H. (2019). Influences of light and humidity on carbonyl sulfide-based estimates of photosynthesis. *Proceedings of the National Academy of Sciences*, 116(7), 2470–2475.

Kuai, L., Worden, J. R., Elliott Campbell, J., Kulawik, S. S., Li, K.-F., Lee, M., Weidner, R. J., Montzka, S. A., Moore, F. L., Berry, J. A., Baker, I., Scott Denning, A., Bian, H., Bowman, K. W., Liu, J., & Yung, Y. L. (2015). Estimate of carbonyl sulfide tropical oceanic surface fluxes using Aura Tropospheric Emission Spectrometer observations. *Journal of Geophysical Research, D: Atmospheres*, 120(20), 11,012–11,023.

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.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., … & Zheng, B. (2018). Global carbon budget 2018. *Earth System Science Data*, *10*(4), 2141-2194.

Lennartz, S. T., Marandino, C. A., von Hobe, M., Andreae, M. O., Aranami, K., Atlas, Elliot, Berkelhammer, M., Bingemer, H., Booge, D., Cutter, G., Cortes, P., Kremser, S., Law, C. S., Marriner, A., Simó, R., Quack, B., Uher, G., Xie, H., & Xu, X. (2020). Marine carbonyl sulfide (OCS) and carbon disulfide (CS_2): a compilation of measurements in seawater and the marine boundary layer. *Earth System Science Data*, 12(1), 591–609.

Lennartz, S. T., Marandino, C. A., von Hobe, M., Cortes, P., Quack, B., Simo, R., Booge, D., Pozzer, A., Steinhoff, T., Arevalo-Martinez, D. L., Kloss, C., Bracher, A., Röttgers, R., Atlas, Elliot, & Krüger, K. (2017). Direct oceanic emissions unlikely to account for the missing source of atmospheric carbonyl sulfide. *Atmospheric Chemistry and Physics*, 17(1), 385–402.

Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K.-M., Commane, R., Wehr, R., Campbell, J. E., Belviso, S., Montzka, S. A., Raoult, N., Seibt, U., Shiga, Y. P., Vuichard, N., Whelan, M. E., & Peylin, P. (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.

Ma, J., Kooijmans, L. M. J., Cho, A., Montzka, S. A., Glatthor, N., Worden, J. R., Kuai, L., Atlas, Elliot L, & Krol, M. C. (2021). Inverse modelling of carbonyl sulfide: implementation, evaluation and implications for the global budget. *Atmospheric Chemistry and Physics*, 21(5), 3507–3529.

Meehl, G. A., G. J. Boer, C. Covey, M. Latif, & R. J. Stouffer. (2000). The Coupled Model Intercomparison Project (CMIP). *Bull. Amer. Meteor. Soc*., 81, 313-318.

Montzka, S. A., Calvert, P., Hall, B. D., Elkins, J. W., Conway, T. J., Tans, P. P., & Sweeney, C. (2007). On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO_2. *Journal of Geophysical Research, D: Atmospheres*, 112(D9). https://doi.org/10.1029/2006JD007665

Ogée, 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.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., & Zimmermann, N. E. (2011). Plant functional type mapping for earth system models. *Geoscientific Model Development*, 4(4), 993–1010.

Protoschill-Krebs, G., Wilhelm, C., & Kesselmeier, J. (1996). Consumption of carbonyl sulphide (COS) by higher plant carbonic anhydrase (CA). *Atmospheric Environment*, *30*(18), 3151-3156.

Remaud, M., Ma, J., Krol, M., Abadie, C., Cartwright, M. P., Patra, P., ... & Peylin, P. (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.

Remaud, M., Chevallier, F., Maignan, F., Belviso, S., Berchet, A., Parouffe, A., Abadie, C., Bacour, C., Lennartz, S., & Peylin, P. (2022). Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling. *Atmospheric Chemistry and Physics*, 22(4), 2525–2552.

Remaud, M., Chevallier, F., Cozic, A., Lin, X., & Bousquet, P. (2018). On the impact of recent developments of the LMDz atmospheric general circulation model on the simulation of CO_2 transport. *Geoscientific Model Development*, *11*(11), 4489-4513.

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 (CO_2). *Biogeosciences* , 2(2), 125–132.

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.

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., Goll, D., Jain, A. K., Joetzjer, E., Lienert, S., Lombardozzi, D., Luyssaert, S., Nabel, J. E. M., Tian, H., Vuichard, N., Walker, A. P., Yuan, W., & Zaehle, S. (2022). Are Terrestrial Biosphere Models Fit for Simulating the Global Land Carbon Sink? *Journal of Advances in Modeling Earth Systems*, 14(5), e2021MS002946.

Sitch, S., O'Sullivan, M., Robertson, E., Friedlingstein, P., Albergel, C., Anthoni, P., ... & Zaehle, S. (2024). Trends and drivers of terrestrial sources and sinks of carbon dioxide: an overview of the TRENDY Project. *Global Biogeochemical Cycles*.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., ... Myneni, R. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* , 12(3), 653–679.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., & Woodward, F. I. (2008). Evaluation of the terrestrial

carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, 14(9), 2015–2039.

Smith, K. S., Jakubzick, C., Whittam, T. S., & Ferry, J. G. (1999). Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. *Proceedings of the National Academy of Sciences*, 96(26), 15184–15189.

Stimler, K., Montzka, S. A., Berry, J. A., Rudich, Y., & Yakir, D. (2010). Relationships between carbonyl sulfide (COS) and CO2 during leaf gas exchange. *The New Phytologist*, 186(4), 869–878.

Stinecipher, J. R., Cameron-Smith, P. J., Blake, N. J., Kuai, L., Lejeune, B., Mahieu, E., Simpson, I. J., & Campbell, J. E. (2019). Biomass Burning Unlikely to Account for Missing Source of Carbonyl Sulfide. *Geophysical Research Letters*, 46(24), 14912–14920.

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). https://doi.org/10.1029/2008GL034332

Sun, W., Maseyk, K., Lett, C., & Seibt, U. (2015). A soil diffusion–reaction model for surface COS flux: COSSM v1. *Geoscientific Model Development*, 8(10), 3055–3070.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *In Bulletin of the American Meteorological Society* (Vol. 93, Issue 4, pp. 485–498). https://doi.org/10.1175/bams-d-11-00094.1

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.

Tramontana, G., Migliavacca, M., Jung, M., Reichstein, M., Keenan, T. F., Camps-Valls, G., Ogee, J., Verrelst, J., & Papale, D. (2020). Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks. *Global Change Biology*, 26(9), 5235–5253.

Wania, R., Ross, I., & Prentice, I. C. (2010). Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1. *Geoscientific Model Development*, 3(2), 565–584.

Wehr, R., Commane, R., Munger, J. W., McManus, J. B., Nelson, D. D., Zahniser, M. S., Saleska, S. R., & Wofsy, S. C. (2017). Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake. *Biogeosciences* , 14(2), 389–401.

Whelan, M. E., Min, D. H., & Rhew, R. C. (2013). Salt marsh vegetation as a carbonyl sulfide (COS) source to the atmosphere. *Atmospheric Environment*, *73*, 131-137.

Whelan, M. E., Lennartz, S. T., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L. M. J., Hilton, T. W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K.-F., Yakir, D., Zumkehr, A., Katayama, Y., Ogée, J., Spielmann, F. M., Kitz, F., Rastogi, B., Kesselmeier, J., Marshall, J., Erkkilä, K.-M., Wingate, L., Meredith, L. K., He, W., Bunk, R., Launois, T., Vesala, T., Schmidt, J. A., Fichot, C. G., Seibt, U., Saleska, S., Saltzman, E. S., Montzka, S. A., Berry, J. A., and Campbell, J. E.: Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer

for carbon and water cycles. *Biogeosciences*, 15, 3625–3657, https://doi.org/10.5194/bg-15-3625-2018, 2018.

Wohlfahrt, G., Hammerle, A., Spielmann, F. M., Kitz, F., & Yi, C. (2023). Novel estimates of the leaf relative uptake rate of carbonyl sulfide from optimality theory. *Biogeosciences*, *20*(3), 589-596.
Wu, Z., Ahlström, A., Smith, B., Ardö, J., Eklundh, L., Fensholt, R., & Lehsten, V. (2017). Climate data induced uncertainty in model-based estimations of terrestrial primary productivity. *Environmental Research Letters: ERL* [Web Site], 12(6), 064013.

Zaehle, S., Sitch, S., Smith, B., & Hatterman, F. (2005). Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. Global Biogeochemical Cycles, 19(3). https://doi.org/10.1029/2004GB002395

Zhu, H., Wu, M., Jiang, F., Vossbeck, M., Kaminski, T., Xing, X., Wang, J., Ju, W., & Chen, J. M. (2024a). Assimilation of carbonyl sulfide (COS) fluxes within the adjoint-based data assimilation system – Nanjing University Carbon Assimilation System (NUCAS v1.0). *Geoscientific Model Development*, *17*(16), 6337–6363. https://doi.org/10.5194/gmd-17-6337-2024

Zhu, H., Xing, X., Wu, M., Ju, W., & Jiang, F. (2024b). Optimizing the terrestrial ecosystem gross primary productivity using carbonyl sulfide (COS) within a two-leaf modeling framework. *Biogeosciences*, *21*(16), 3735–3760. https://doi.org/10.5194/bg-21-3735-2024

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.

Appendix A: COS and CO_2 seasonal amplitude and phase analysis per model

Table A1: COS and CO_2 seasonal amplitude (in ppt for COS and ppm for CO_2) and phase correlation (amplitude/correlation) for each of the CMIP5, CMIP6 and TRENDY-v7 models.

	Model	MLO		ALT	
		COS	CO_2	COS	CO_2
Obs		48.0/1.0	6.8/1.0	113.8/1.0	17.0/1.0
CMIP5	BNU-ESM	42.1/0.91	4.6/0.95	93.5/0.94	17.7/0.90
	CanESM2	29.0/0.96	3.8/0.89	76.3/0.97	10.7/0.93
	CCSM4	30.6/0.83	3.7/0.98	43.7/0.90	9.7/0.89
	CESM1-BGC	32.2/0.85	4.2/0.97	44.5/0.90	9.7/0.88
	CESM1-CAM5	34.6/0.83	4.2/0.96	46.8/0.89	10.4/0.86
	CESM1-FASTCHEM	30.4/0.85	3.8/0.98	42.9/0.91	10.0/0.88
	CESM1-WACCM	30.5/0.85	4.2/0.98	47.8/0.90	9.7/0.89
	GFDL-ESM2G	69.9/0.95	8.8/0.95	130.1/0.88	11.5/0.50
	GFDL-ESM2M	69.1/0.93	8.7/0.95	132.7/0.80	14.8/0.36
	HadGEM2-CC	48.7/0.93	7.8/0.99	95.7/0.98	18.2/0.98
	HadGEM2-ES	47.3/0.93	7.4/0.98	90.9/0.95	18.5/0.98
	INM-CM4	39.8/0.85	3.1/0.96	71.6/0.93	9.2/0.96
	IPSL-CM5A-LR	87.4/0.90	13.0/0.99	195.1/0.89	39.6/0.98
	IPSL-CM5A-MR	79.8/0.90	11.7/0.99	194.1/0.88	38.1/0.99
	IPSL-CM5B-LR	75.8/0.90	10.8/0.99	162.7/0.90	33.0/0.98
	MIROC-ESM-CHEM	52.6/0.88	6.8/0.98	111.5/0.90	18.5/0.96
	MIROC-ESM	53.6/0.88	7.2/0.97	113.7/0.92	17.9/0.97

	MPI-ESM-LR	62.1/0.88	8.0/0.97	114.5/0.94	24.0/0.98
	MPI-ESM-MR	65.9/0.88	8.9/0.97	125.1/0.94	27.4/0.97
	MPI-ESM-P	58.6/0.86	8.0/0.97	109.3/0.94	25.1/0.98
	MRI-ESM1	94.5/0.88	5.1/0.98	191.6/0.93	20.2/0.97
	NorESM1-ME	32.2/0.85	4.9/0.98	46.2/0.91	10.4/0.90
	NorESM1-M	32.4/0.85	4.6/0.98	47.2/0.90	10.4/0.89
CMIP6	ACCESS-ESM1-5	35.8/0.94	4.9/0.98	54.8/0.95	10.7/0.95
	CanESM5	45.2/0.97	5.9/0.90	96.0/0.96	14.5/0.90
	CESM2-FV2	47.7/0.72	5.9/0.85	92.4/0.88	13.1/0.96
	CESM2	47.9/0.75	5.5/0.87	92.6/0.84	12.8/0.95
	CESM2-WACCM	46.3/0.76	5.3/0.87	93.4/0.87	12.4/0.96
	CMCC-CM2-SR5	42.7/0.86	5.6/0.93	84.8/0.79	17.6/0.82
	CMCC-ESM2	42.7/0.85	5.6/0.95	82.0/0.80	16.8/0.85
	CNRM-ESM2-1	40.6/0.90	4.3/0.97	83.8/0.84	14.0/0.92
	E3SM-1-1-ECA	39.5/0.86	5.5/0.95	59.3/0.95	10.0/0.97
	E3SM-1-1	47.9/0.90	9.0/0.99	90.2/0.96	19.0/0.99
	EC-Earth3-Veg	44.1/0.94	5.3/0.73	86.6/0.95	12.5/0.70
	EC-Earth3-Veg-LR	41.6/0.95	5.1/0.79	79.9/0.98	12.0/0.77
	GFDL-ESM4	44.8/0.95	6.3/0.99	93.2/0.84	14.9/0.85
	INM-CM4-8	43.2/0.83	3.2/0.93	95.9/0.88	13.5/0.92
	INM-CM5-0	45.8/0.85	3.7/0.93	99.4/0.90	14.5/0.92
	IPSL-CM6A-LR	47.8/0.91	5.7/0.98	94.3/0.96	14.1/0.99

	MIROC-ES2L	48.3/0.84	6.4/0.95	100.0/0.84	23.7/0.94
	MPI-ESM-1-2-HAM	55.5/0.92	4.9/0.94	121.1/0.92	17.5/.0.84
	MPI-ESM1-2-LR	57.6/0.90	5.1/0.89	121.0/0.89	19.5/0.79
	NorCPM1	21.2/0.81	2.2/0.92	25.8/0.90	3.4/0.92
	NorESM2-LM	44.3/0.76	5.1/0.87	83.6/0.88	11.7/0.96
	NorESM2-MM	45.1/0.76	5.0/0.88	87.5/0.88	11.3/0.97
	SAM0-UNICON	34.1/0.86	4.3/0.98	47.0/0.92	9.3/0.90
	TaiESM1	32.6/0.82	4.0/0.98	45.3/0.90	8.7/0.90
	UKESM1-0-LL	44.6/0.88	4.6/0.99	92.2/0.91	15.9/0.96
TRENDY-v7	CABLE-POP	39.5/0.93	1.9/0.90	87.8/0.95	5.2/0.96
	CLASS	41.0/0.98	6.9/0.89	81.7/0.98	15.4/0.92
	CLM5	59.5/0.81	6.6/0.86	140.4/0.87	17.8/0.97
	DLEM	49.5/0.95	6.0/0.95	108.4/0.98	11.4/0.97
	ISAM	45.6/0.89	3.5/0.98	89.1/0.97	7.6/0.97
	JSBACH	41.1/0.94	3.7/0.94	89.8/0.90	14.3/0.75
	JULES	43.7/0.97	5.0/0.96	91.0/0.95	21.1/0.97
	LPJ	64.8/0.90	2.6/0.94	147.0/0.88	15.0/0.82
	LPX	74.2/0.89	8.3/0.99	157.5/0.95	16.5/0.98
	OCN	56.4/0.95	6.5/0.93	118.0/0.97	18.3/0.91
	ORCHIDEE-CNP	29.7/0.91	2.0/0.55	68.4/0.88	10.7/0.90
	ORCHIDEE	48.5/0.93	6.9/0.99	93.0/0.96	15.2/0.99
	SDGVM	43.7/0.89	7.7/0.84	83.3/0.88	23.8/0.76

	SURFEX	40.0/0.96	6.0/0.99	75.3/0.97	13.5/0.99

Appendix B: Mean seasonal cycles of COS and CO_2 atmospheric concentrations at 8 stations

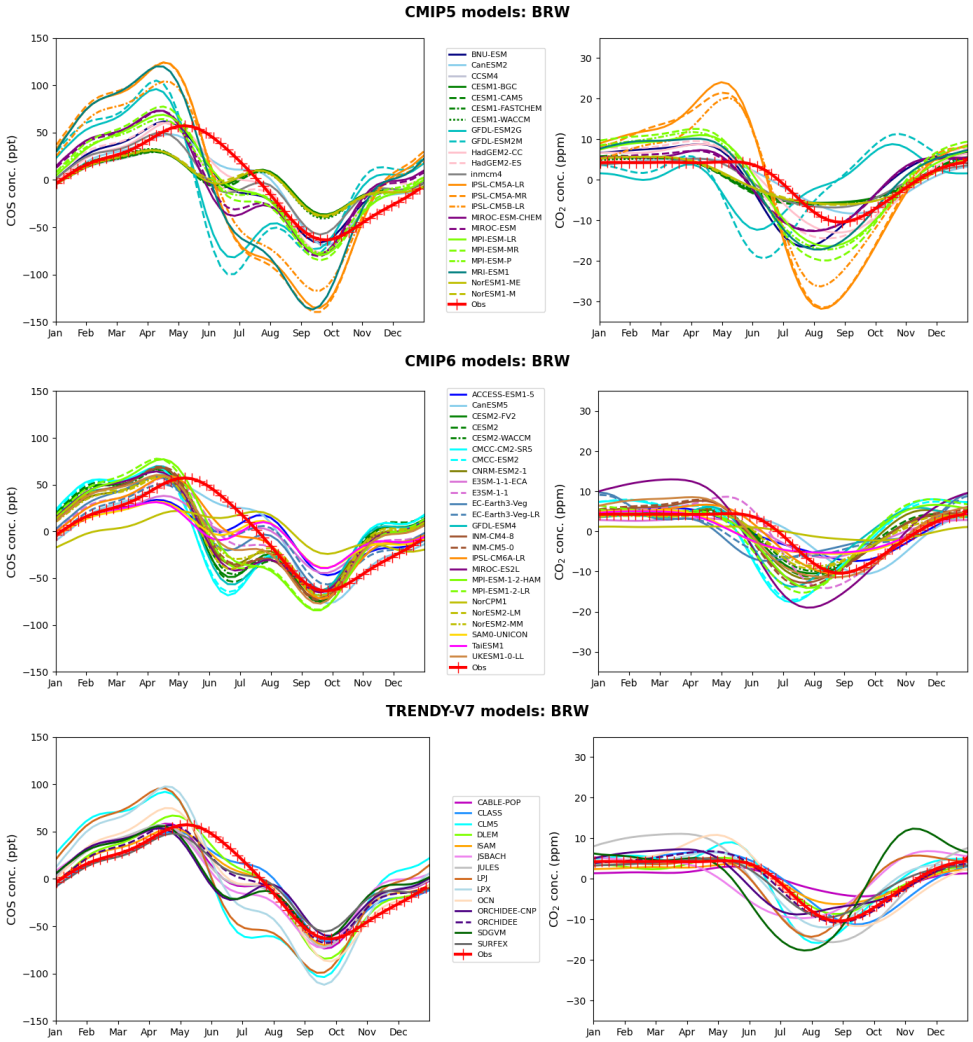

Figure B1: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at BRW for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

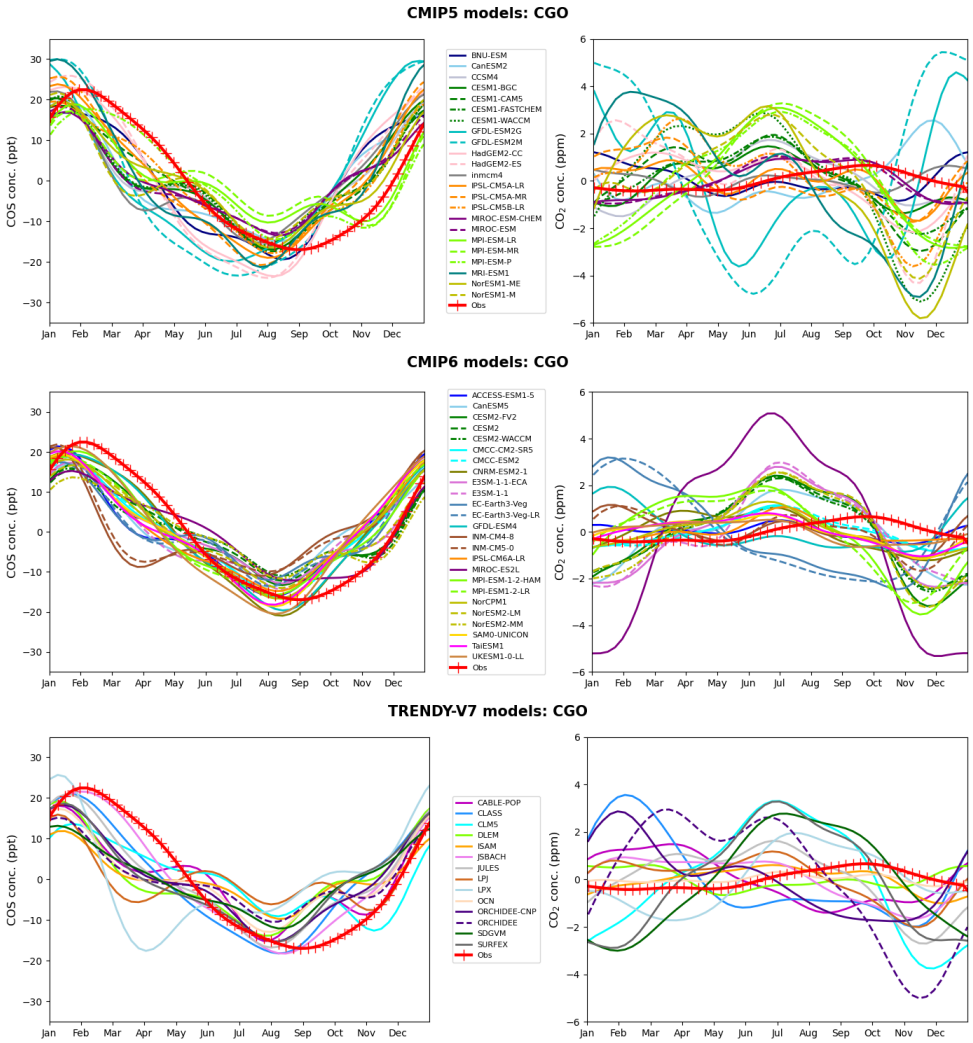

Figure B2: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at CGO for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

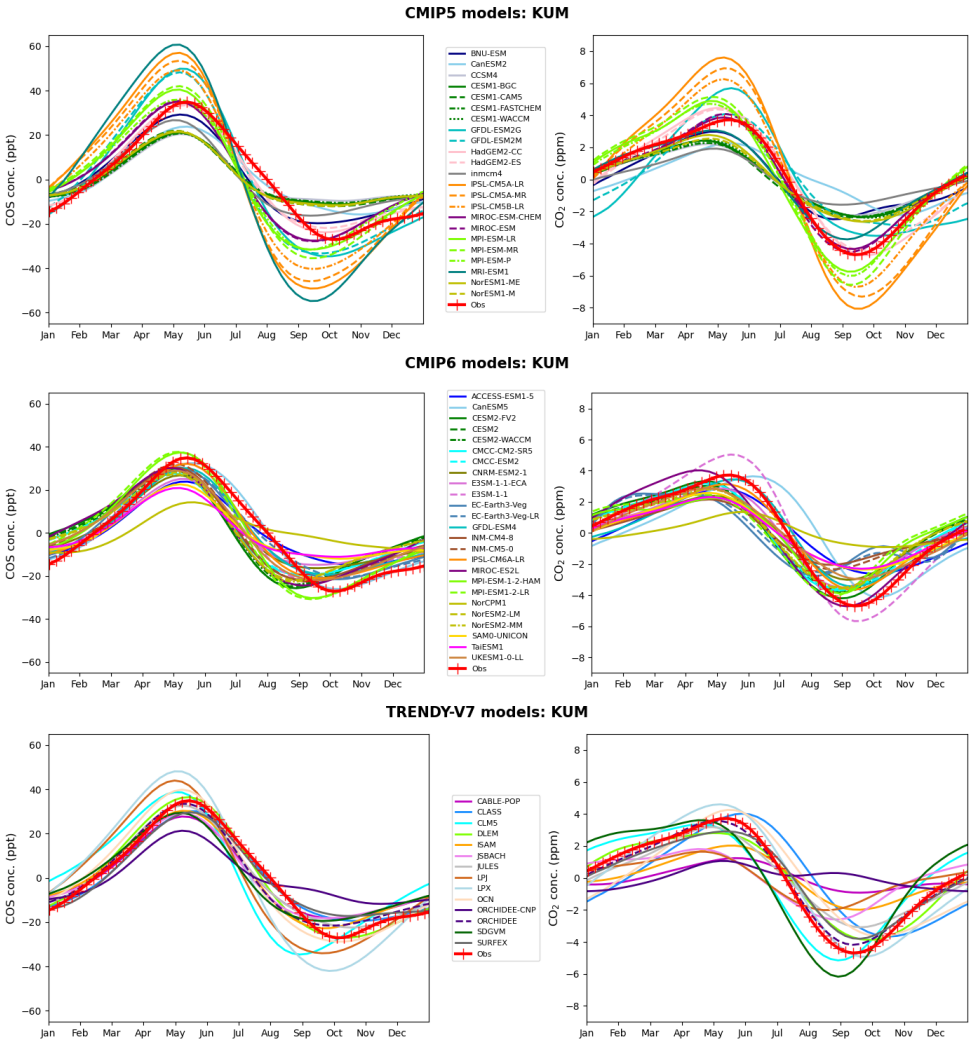

Figure B3: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at KUM for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

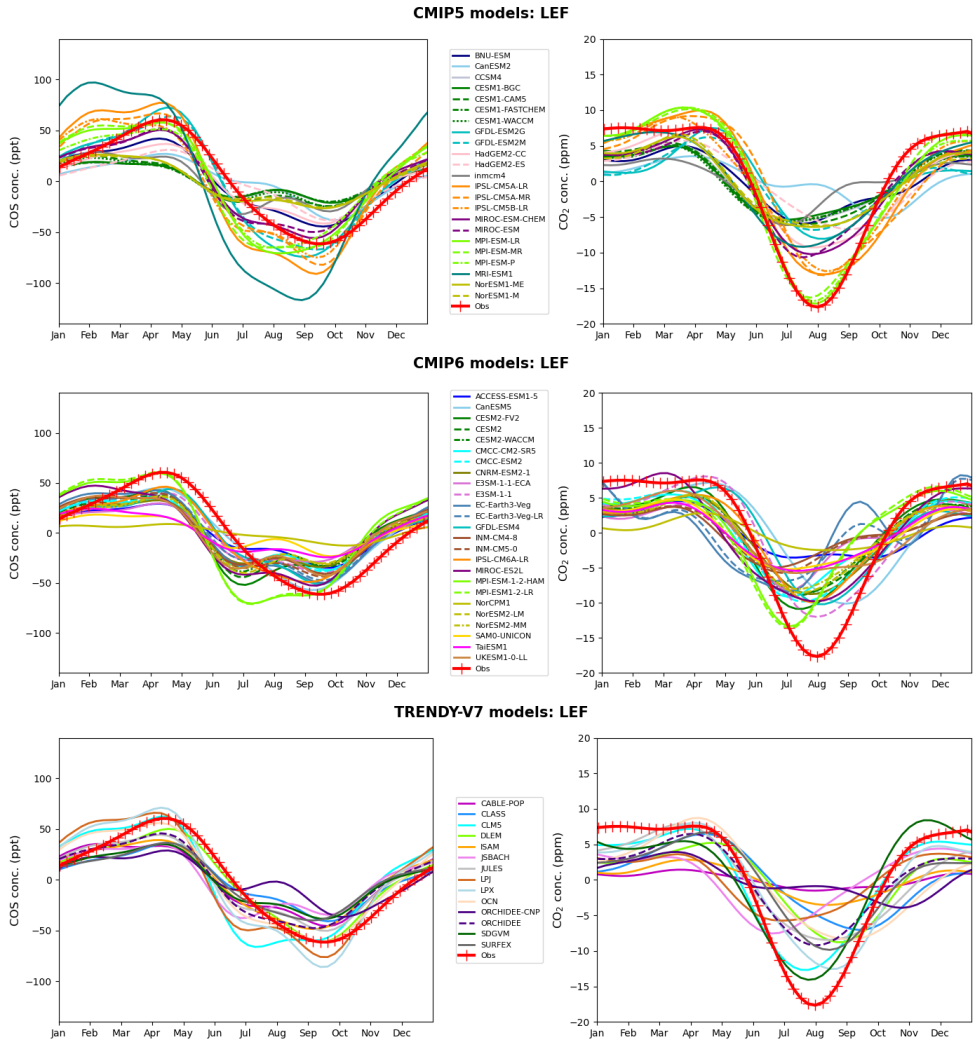

Figure B4: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at LEF for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

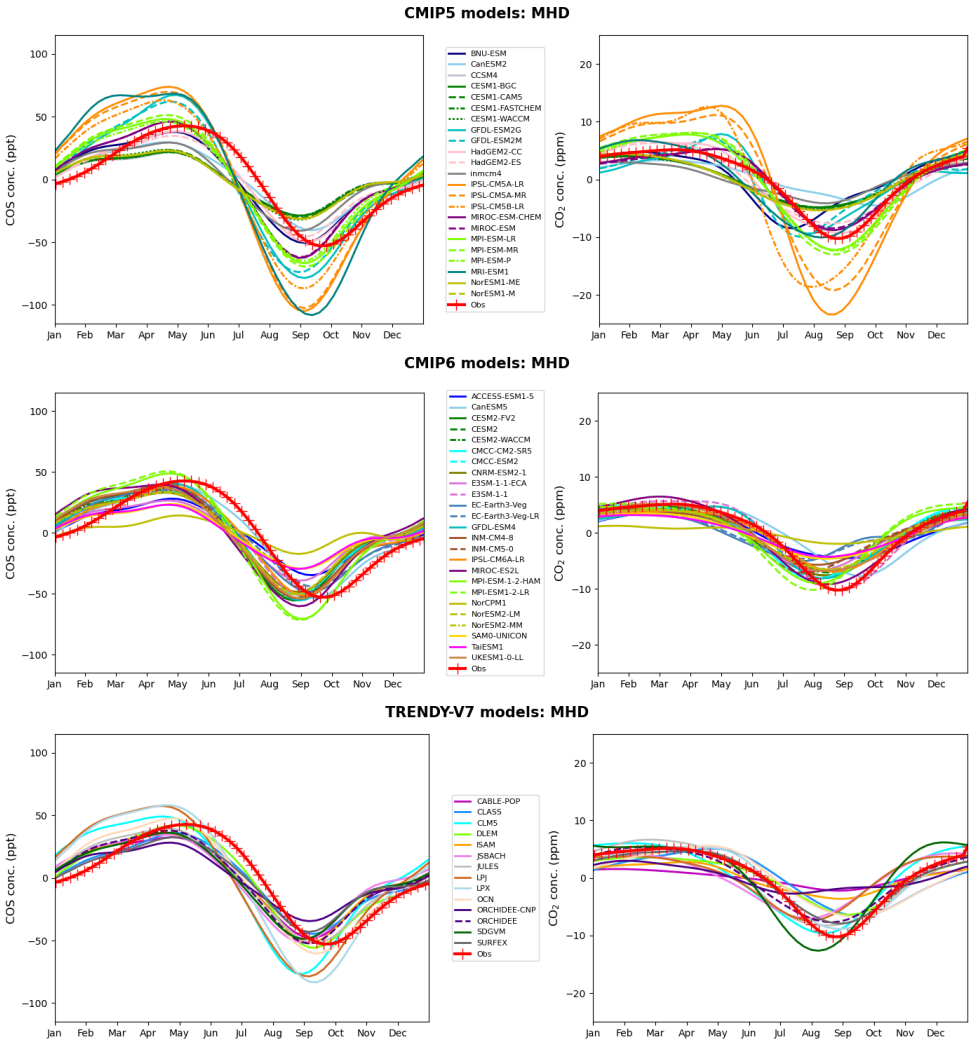

Figure B5: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at MHD for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

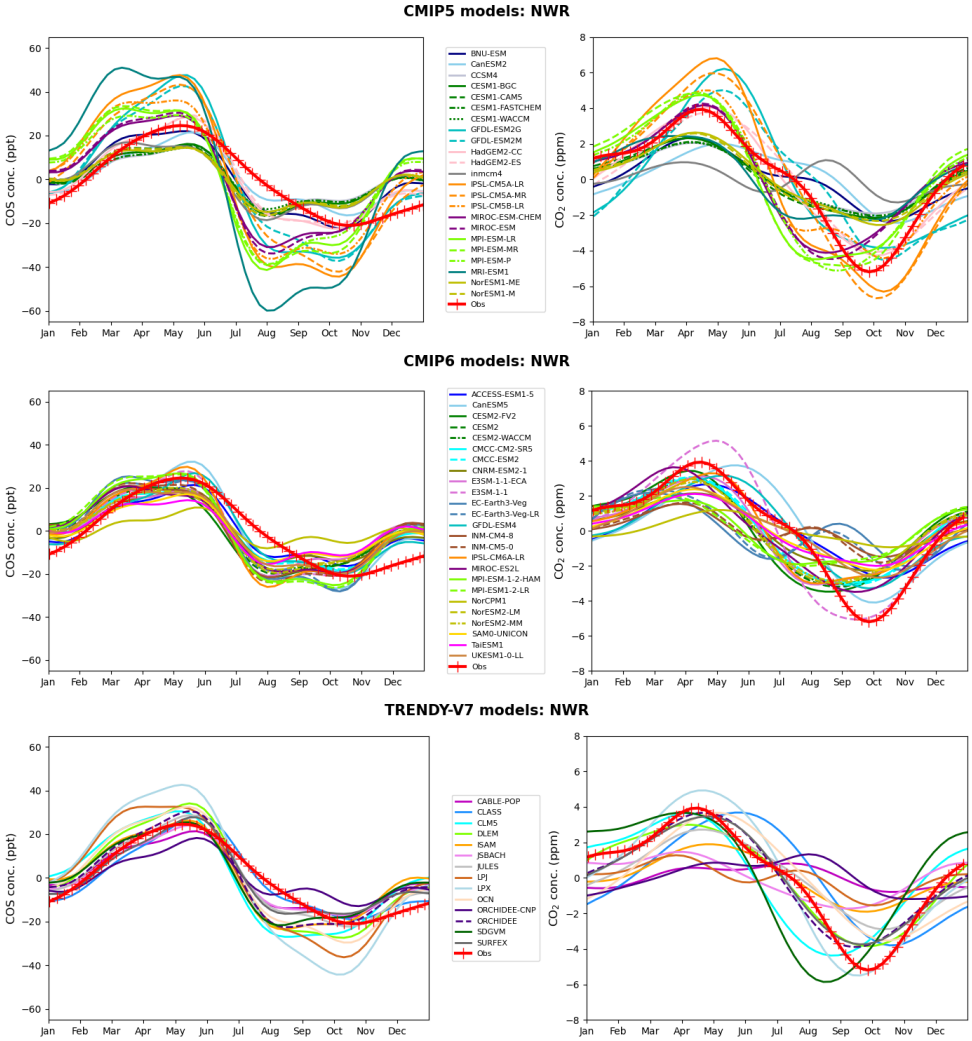

Figure B6: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at NWR for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

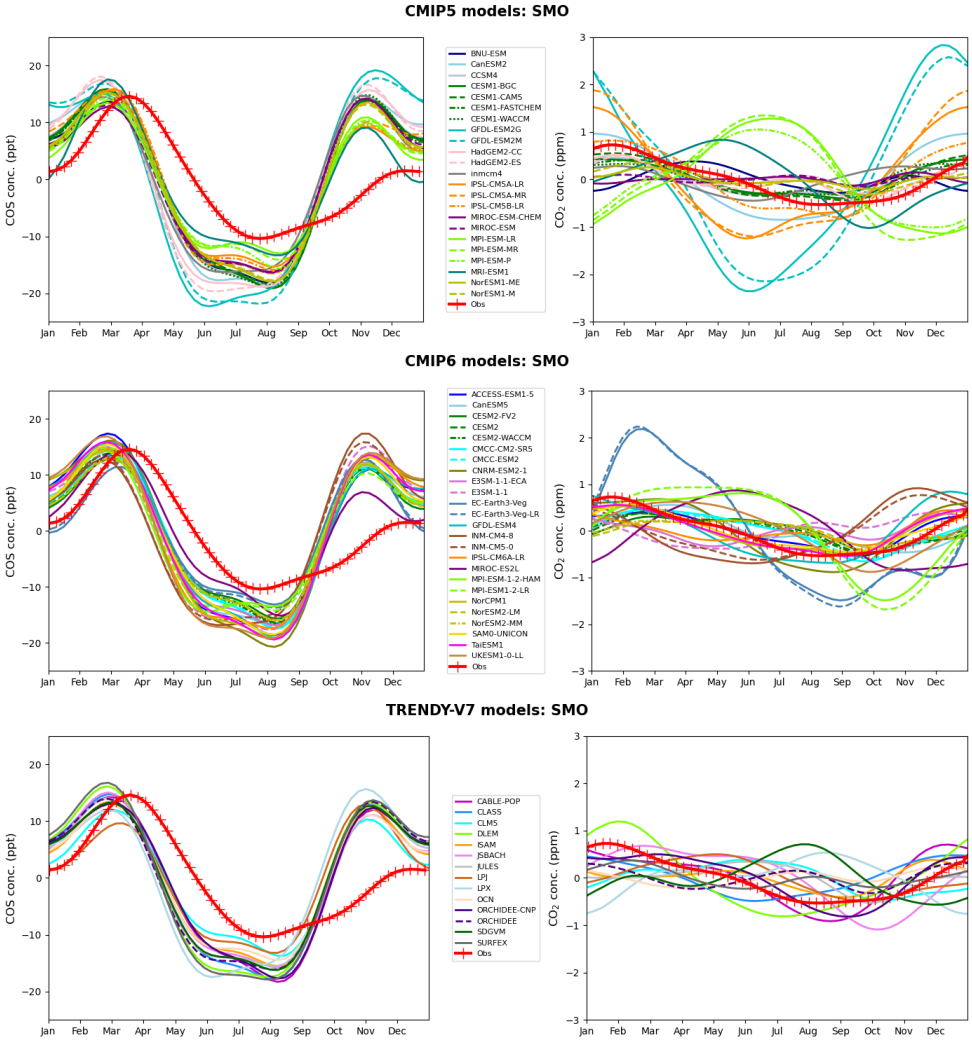

Figure B7: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at SMO for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

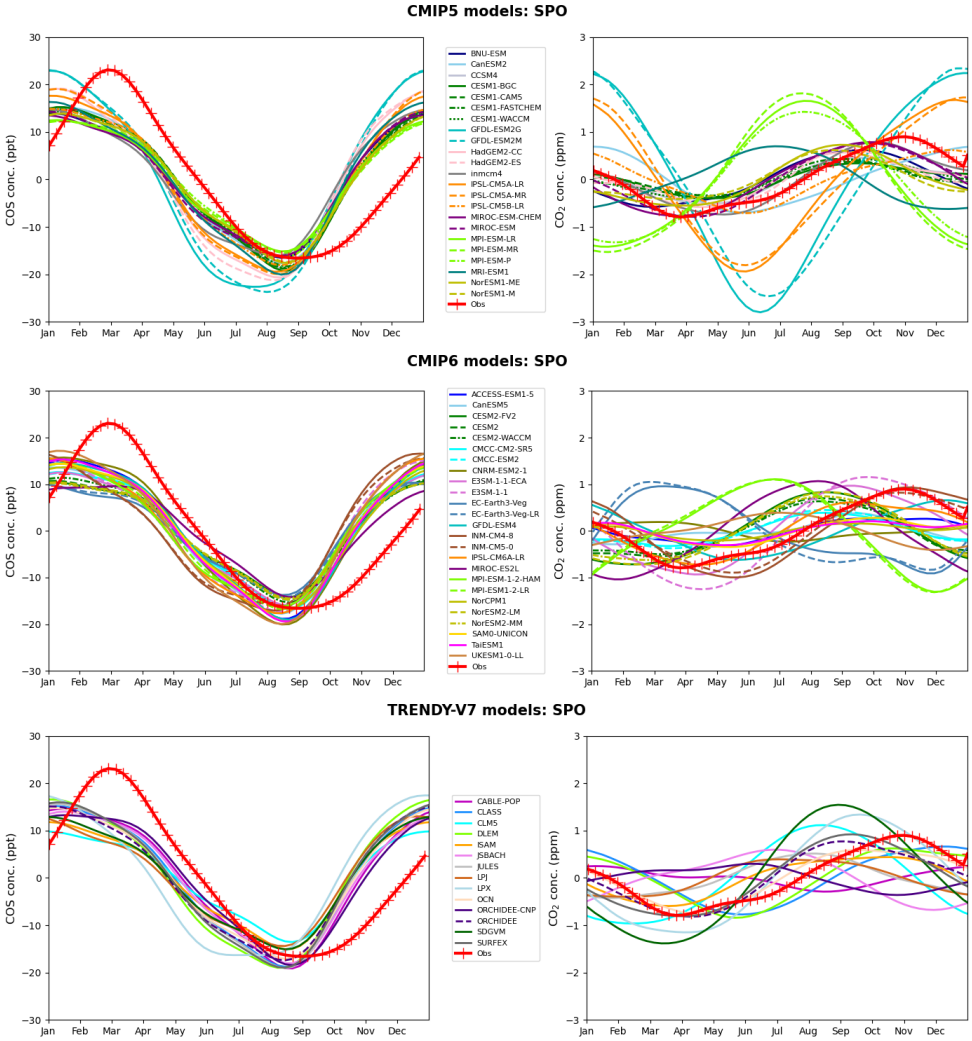

Figure B8: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at SPO for CMIP5, CMIP6 and TRENDY-v7 models. The observations are represented by red crosses.

Appendix C: Sensitivity of simulated COS concentrations to soil fluxes

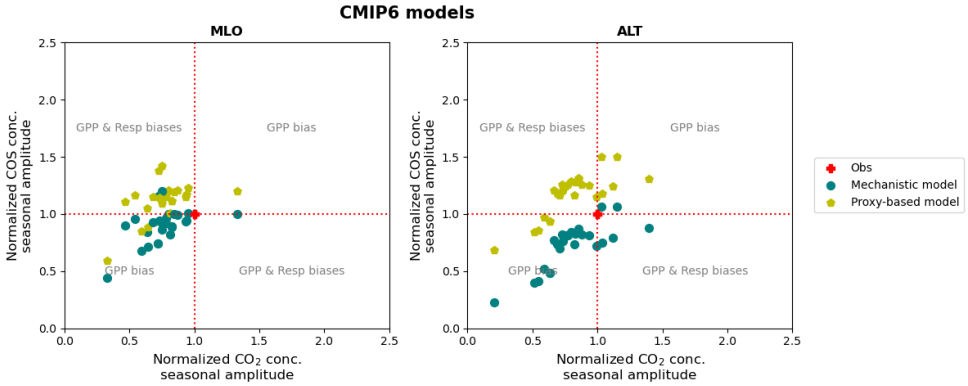

Figure C1: Scatter plots of CO2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO and ALT for CMIP6 models. The simulated COS concentrations use either the mechanistic model or the proxy-based approach for soil COS fluxes. The observed amplitude is identified by the red cross.

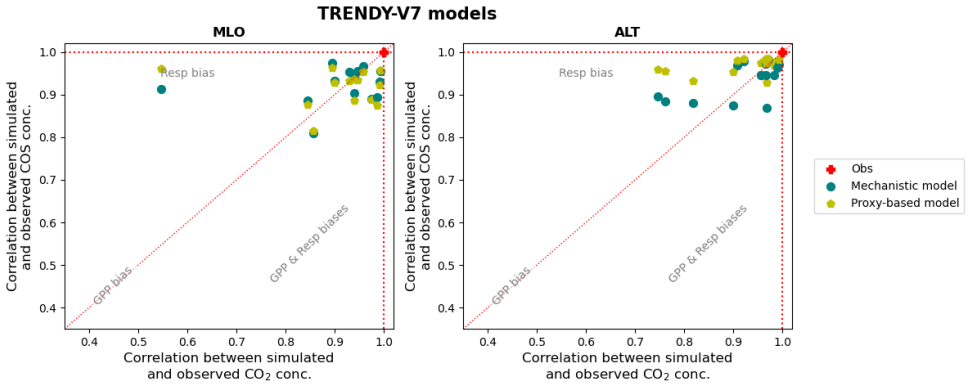

Figure C2: Scatter plots of CO2 vs COS phase correlations between modeled and observed smoothed seasonal concentrations at MLO and ALT simulated by the TRENDY-v7 models. The simulated COS concentrations use either the mechanistic model or the proxy-based approach for soil COS fluxes. The observed amplitude is identified by the red cross.

Appendix D: Sensitivity of simulated COS concentrations to LRU

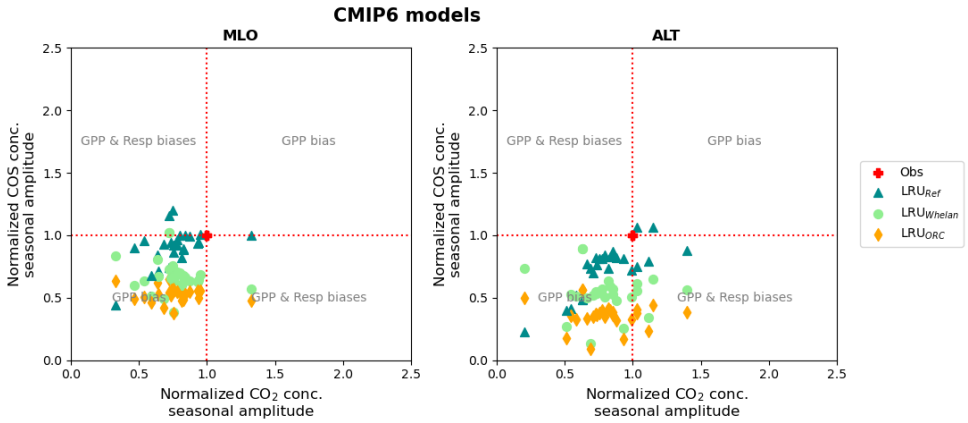

Figure D1: Scatter plots of CO2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO and ALT for CMIP6 models. The different LRU scenarios are Seibt Reference LRU, ORCHIDEE LRU and Whelan LRU. The observed phase is identified by the red cross.

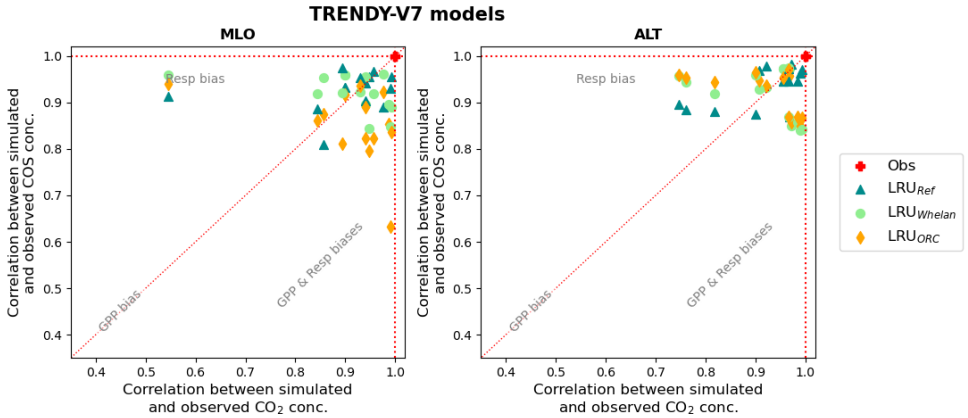

Figure D2: Scatter plots of CO2 vs COS phase correlations between modeled and observed smoothed seasonal concentrations at MLO and ALT simulated by the TRENDY-v7 models. The different LRU scenarios are Seibt Reference LRU, ORCHIDEE LRU and Whelan LRU. The observed phase is identified by the red cross.

Appendix E: Sensitivity of simulated COS concentrations to ocean fluxes

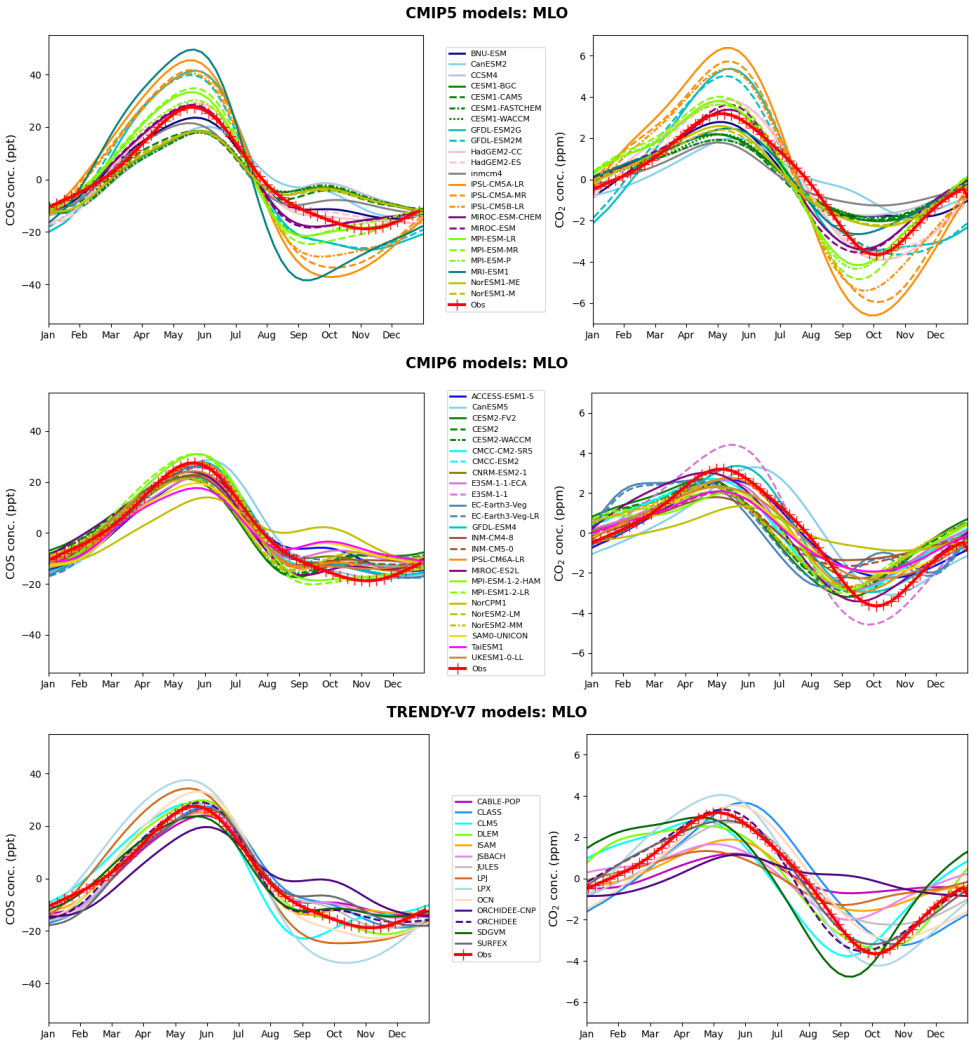

Figure E1: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at MLO for CMIP5, CMIP6 and TRENDY-v7 models. The simulated atmospheric COS concentrations use the oceanic estimates by Lennartz et al. (2017, 2020) for direct COS emissions and indirect emissions from CS_2, and the NEMO-PISCES model estimate for indirect emissions from DMS (Launois et al., 2015). The observations are represented by red crosses.

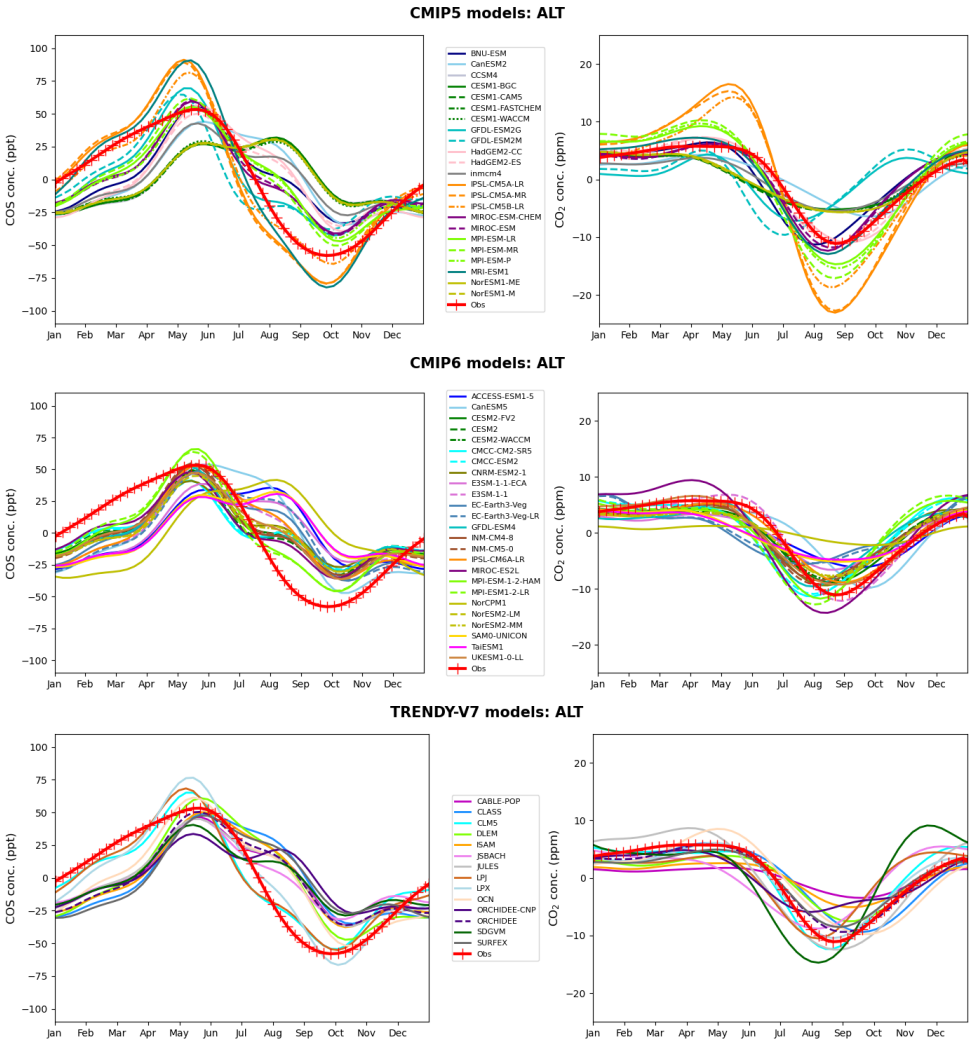

Figure E2: Smoothed seasonal cycles of atmospheric COS (left) and CO2 (right) concentrations simulated at ALT for CMIP5, CMIP6 and TRENDY-v7 models. The simulated atmospheric COS concentrations use the oceanic estimates by Lennartz et al. (2017, 2020) for direct COS emissions and indirect emissions from CS_2, and the NEMO-PISCES model estimate for indirect emissions from DMS (Launois et al., 2015). The observations are represented by red crosses.

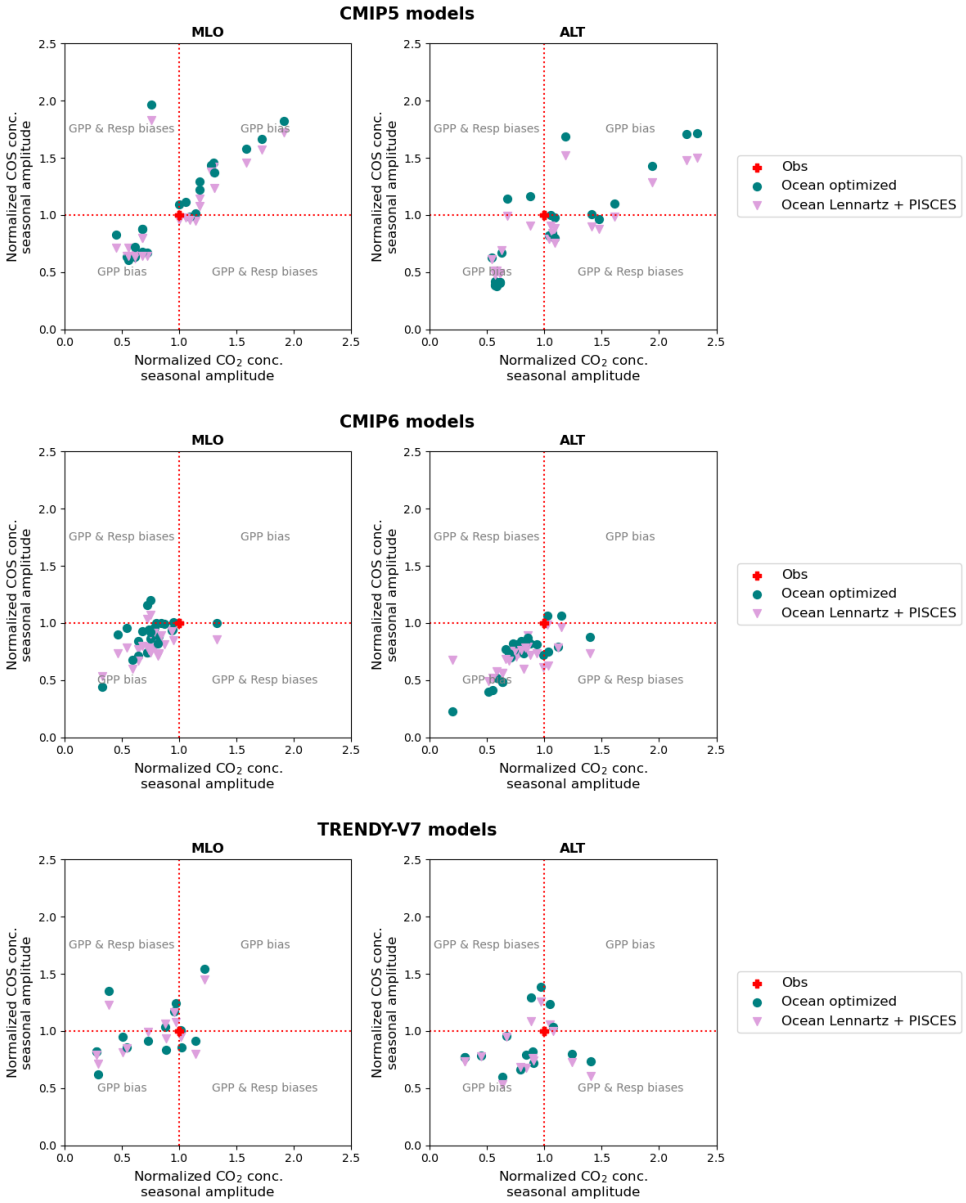

Figure E3: Scatter plots of CO2 vs COS simulated amplitudes normalized by observed amplitudes of smoothed seasonal concentrations at MLO and ALT for CMIP5, CMIP6 and TRENDY-v7 models. The simulated COS concentrations use either the optimized ocean fluxes, or the ocean fluxes estimated by Lennartz et al. (2017, 2020) for direct COS emissions and indirect emissions from CS_2 and the NEMO-PISCES model estimate for indirect emissions from DMS (Launois et al., 2015). The observed amplitude is identified by the red cross.

Appendix F: LRU seasonal variations

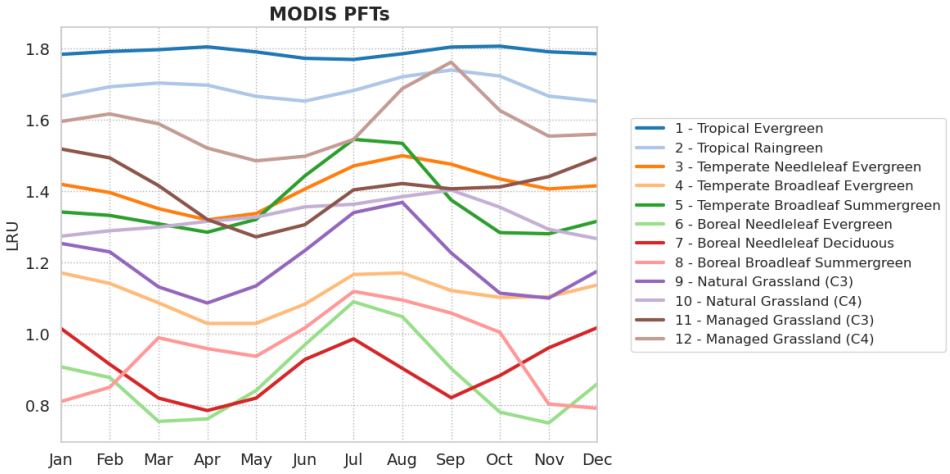

Figure F1: Seasonal variations of mean monthly LRU per PFT over 2000-2009. The LRU values were computed following the same approach as for LRU_{ORC} using ORCHIDEE simulated GPP and vegetation COS fluxes, and 3-hourly variable atmospheric COS and CO_2 concentrations simulated by the LMDZ transport model.