

```

%% Copernicus Publications Manuscript Preparatio Template for LaTeX
Submissions
%% -----
%% This template should be used for copernicus.cls
%% The class file and some style files are bundled in the Copernicus
Latex Package, which can be downloaded from the different journal
webpages.
% * <marine.remaud@gmail.com> 2018-05-15T08:56:53.865Z:
%
% ^.
%% For further assistance please contact Copernicus Publications at:
production@copernicus.org
%%
https://publications.copernicus.org/for_authors/manuscript_preparatio.htm
1

%% Please use the following documentclass and journal abbreviations for
discussion papers and final revised papers.

%% 2-column papers and discussion papers
\newcommand{\tcr}[1]{\textcolor{red}{#1}}

\documentclass[journal abbreviation, manuscript]{copernicus}

%% Journal abbreviations (please use the same for discussion papers and
final revised papers)

% Advances in Geosciences (adgeo)
% Advances in Radio Science (ars)
% Advances in Science and Research (asr)
% Advances in Statistical Climatology, Meteorology and Oceanography
(ascmo)
% Annales Geophysicae
(angeo)https://www.overleaf.com/project/5e380def95e1690001735a39
% Archives Animal Breeding (aab)
% ASTRA Proceedings (ap)
% Atmospheric Chemistry and Physics (acp)
% Atmospheric Measurement Techniques (amt)
% Biogeosciences (bg)
% Climate of the Past (cp)
% Drinking Water Engineering and Science (dwes)
% Earth Surface Dynamics (esurf)
% Earth System Dynamics (esd)
% Earth System Science Data (essd)
% E&G Quaternary Science Journal (egqsj)
% Fossil Record (fr)
% Geographica Helvetica (gh)
% Geoscience Communication (gc)
% Geoscientific Instrumentation, Methods and Data Systems (gi)
% Geoscientific Model Development (gmd)
% History of Geo- and Space Sciences (hgss)
% Hydrology and Earth System Sciences (hess)
% Journal of Micropalaeontology (jm)
% Journal of Sensors and Sensor Systems (jsss)

```

a mis en forme : Français

```

% Mechanical Sciences (ms)
% Natural Hazards and Earth System Sciences (nhess)
% Nonlinear Processes in Geophysics (npg)
% Ocean Science (os)
% Primate Biology (pb)
% Proceedings of the International Association of Hydrological Sciences
(piahs)
% Scientific Drilling (sd)
% SOIL (soil)
% Solid Earth (se)
% The Cryosphere (tc)
% Web Ecology (we)
% Wind Energy Science (wes)

```

```

%% \usepackage commands included in the copernicus.cls:

```

```

%\usepackage[german, english]{babel}
%\usepackage{tabularx}
%\usepackage{cancel}
%\usepackage{multirow}
%\usepackage{supertabular}
%\usepackage{algorithmic}
%\usepackage{algorithm}
%\usepackage{amsthm}
%\usepackage{float}
%\usepackage{subfig}
%\usepackage{rotating}

```

```

\begin{document}
\title{Plant gross primary production, plant respiration and carbonyl
sulfide emissions over the globe inferred by atmospheric inverse
modelling}

```

```

%\title{Joint assimilation of Carbon Dioxide and Carbonyl Sulphide
measurements into the LMDz transport model: implication for the global
 $\text{COS}$  budget and its potential to estimate the GPP}
%\title{Optimizing the Gross Primary Production: assimilation of Carbonyl
Sulphide surface measurements into the LMDz transport model}

```

```

% \Author[affil]{given_name}{surname}

```

```

\Author[1]{Marine}{Remaud}
\Author[1]{Frédéric}{Chevallier}
\Author[1]{Fabienne}{Maignan}
\Author[1]{Sauveur}{Belviso}
\Author[1]{Antoine}{Berchet}
\Author[1]{Alexandra}{Parouffe}
\Author[1]{Camille}{Abadie}
\Author[1]{Cédric}{Bacour}
\Author[2]{Sinikka}{Lennartz}
\Author[1]{Philippe}{Peylin}

```

```

%\Author[Laboratoire des Sciences du Climat et de
l'Environnement]{et}{al}

```

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: %\

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement

a supprimé: %\

a supprimé: Laboratoire des Sciences du Climat et de l'Environnement}{Vladislav}{Bastrikov}
%\Author[Laboratoire des Sciences du Climat et de l'Environnement]{Philippe}{Bousquet}
%\Author[Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research]{Michael}{Kiefer}
\Author[Laboratoire des Sciences du Climat et de l'Environnement]{Cédric}{Bacour}
\Author[Institute and Chemistry and Biology of the Marine Environment, University of Oldenburg..

a mis en forme : Français

a supprimé: %\Author[NOAA Global Monitoring Laboratory]{Steve}{Montzka}
%\Author[Weizmann Institute of Science]{Dan}{Yakir}
\Author[Laboratoire des Sciences du Climat et de l'Environnement

```
\affil[1]{Laboratoire des Sciences du Climat et de l'Environnement, CEA-  
CNRS-UVSQ, UMR8212, IPSL, Saint-Aubin, France}  
\affil[2]{Institute and Chemistry and Biology of the Marine Environment,  
University of Oldenburg, Oldenburg, Germany}
```

a supprimé:](Orme des Merisiers, 91190

a supprimé: }

```
%\affil[] {ADDRESS}
```

a mis en forme : Français

```
%% The [] brackets identify the author with the corresponding  
affiliation. 1, 2, 3, etc. should be inserted.
```

```
\runningtitle{TEXT}
```

```
\runningauthor{TEXT}
```

```
\correspondence{Marine Remaud (mremaud@lsce.ipsl.fr)}
```

```
\received{}
```

```
\pubdiscuss{} %% only important for two-stage journals
```

```
\revised{}
```

```
\accepted{}
```

```
\published{}
```

```
%% These dates will be inserted by Copernicus Publications during the  
typesetting process.
```

```
\firstpage{1}
```

```
\maketitle
```

```
\begin{abstract}
```

Carbonyl Sulphide (COS), a trace gas showing striking similarity to CO_2 in terms of biochemical diffusion pathway into leaves, has been recognized as a promising indicator of the plant gross primary production (GPP), the amount of carbon dioxide that is absorbed through photosynthesis by terrestrial ecosystems. However, large uncertainties about the other components of its atmospheric budget prevent us from directly relating the atmospheric COS measurements to GPP. The largest uncertainty comes from the closure of its atmospheric budget, with a source component missing. Here, we explore the benefit of assimilating both COS and CO_2 measurements into the LMDz atmospheric transport model to obtain consistent information on GPP, plant respiration and COS budget. To this end, we develop an analytical inverse system that optimizes biospheric fluxes for the 15 plant functional types (PFTs) defined in the ORCHIDEE global land surface model. Plant uptake of COS is parameterized as a linear function of GPP and of the leaf relative uptake (LRU), which is the ratio of COS to CO_2 deposition velocities in plants. A possible scenario for the period 2008-2019 leads to a global biospheric sink of 800 GgS.yr^{-1} , with higher absorption in the high latitudes and higher oceanic emissions between 400 and 600

$\mathrm{GgS.yr^{-1}}$ most of which is located in the tropics. As for the $\mathrm{CO_2}$ budget, the inverse system increases GPP in the high latitudes by a few $\mathrm{GtC.yr^{-1}}$ without modifying the respiration compared to the ORCHIDEE fluxes used as a prior. In contrast, in the tropics the system tends to weaken both respiration and GPP. The optimized components of the COS and $\mathrm{CO_2}$ budgets have been evaluated against independent measurements over Northern America, the Pacific Ocean, at three sites in Japan and at one site in France. Overall, the posterior COS concentrations are in better agreement with the COS retrievals at 250 hPa from the MIPAS satellite and with airborne measurements made over North America and the Pacific Ocean. The system seems to have rightly corrected the underestimated GPP over the high latitudes. However, the change in seasonality of GPP in the tropics disagrees with Solar Induced Fluorescence (SIF) data. The decline in biospheric sink in the Amazon driven by the inversion also disagrees with MIPAS COS retrievals at 250 hPa, highlighting the lack of observational constraints in this region. Moreover, the comparison with the surface measurements in Japan and France suggests misplaced sources in the prior anthropogenic inventory, emphasizing the need for an improved inventory to better partition oceanic and continental sources in Asia and Europe.

a mis en forme : Français

`\end{abstract}`

`%\copyrightstatement{TEXT}`

`\introduction`

Globally, the amount of carbon assimilated by plant photosynthesis, known as Gross Primary Productivity (GPP), exceeds plant respiration by a few $\mathrm{GtC.yr^{-1}}$, which allows terrestrial ecosystems to be a global sink for $\mathrm{CO_2}$ in the atmosphere. By absorbing a quarter of the atmospheric carbon dioxide ($\mathrm{CO_2}$) emitted by human activities, terrestrial ecosystems help to mitigate the increasing $\mathrm{CO_2}$ concentration in the atmosphere, the main driver of climate change [\cite{friedlingstein_global_2020}](#). The spatial distribution of this carbon sink remains uncertain and a subject of intensive research. This is obviously also the case for its components, GPP and respiration. ~~For these~~ gross fluxes, the uncertainty on the seasonal variations and the overall magnitude are also very large [\cite{anav_spatiotemporal_2015}](#).

a supprimé : , and for those

a mis en forme : Français

The two most common methods for estimating ecosystem-wide GPP and respiration are based on eddy-covariance measurements ~~and~~ land surface models (LSMs), ~~respectively~~. While eddy-covariance measurements, on one hand, can be used to routinely estimate GPP and respiration at local scale, their extrapolation to a whole biome is not straightforward due to their small footprint [\cite{jung_scaling_2020}](#). Land Surface Models (LSMs), on the other hand, have global coverage but represent processes that are not well described and are therefore heavily tuned [\cite{kuppel_constraining_2012}](#). For instance, ~~LSMs~~ disagree on the representation of the large spatial and temporal variability of the $\mathrm{CO_2}$ gross and net fluxes [\cite{anav_spatiotemporal_2015}](#). Satellite retrievals of, e.g., solar-induced fluorescence (SIF) or

a supprimé : or

a mis en forme : Français

a supprimé :) .

a mis en forme : Français

a supprimé : they

a mis en forme : Français

normalized difference vegetation index (NDVI) \citep{joiner_new_2016} are also used to constrain GPP. However, remote sensing methods rely on a number of assumptions to convert satellite-measured photons to on-the-ground photosynthesis \citep{sun_overview_2018}. Therefore, there is a need for new information about GPP or respiration to ensure a better partitioning between these components of the CO_2 atmospheric budget.

a supprimé: the

a mis en forme : Français

Carbonyl sulfide (COS) is recognized as a promising tracer of GPP at the leaf scale \citep{stimler_relationships_2010,seibt_kinetic_2010} and at large scale \citep{campbell_photosynthetic_2008,blake_carbonyl_2008}. COS follows the same diffusion pathway from the leaf boundary layer to the plant cells where photosynthesis takes place. However, while CO_2 is re-emitted into the atmosphere through respiration, COS is nearly irreversibly hydrolyzed in a reaction catalyzed by the enzyme carbonic anhydrase (CA) \citep{protoschill-krebs_consumption_1996}. Therefore, the atmospheric drawdown of COS reflects the uptake of COS by the plant to a large extent. Despite this property, COS measurements cannot easily be used in inverse modelling to constrain GPP because the other terms of the COS atmospheric budget are also poorly quantified, to the point that the bottom-up COS atmospheric budget is even less closed than the bottom-up CO_2 atmospheric budget. The process description of all components of the COS budget (i.e. bottom-up budget) suggests a decreasing concentration of COS , but the latter has been relatively stable around 500 parts per trillion (ppt, 1 ppt is 10^{-12} $\mathrm{mol.mol}^{-1}$) over the past 30 years \citep{whelan_reviews_2018}. The current notion is that there is a "missing" source in the current atmospheric COS budget, likely in the tropics \citep{montzka_global_2007,glatthor_tropical_2015,berry_coupled_2013}.

a mis en forme : Français

The terrestrial sink induced by both plants and soils has been estimated between 500-1200 $\mathrm{GgS.yr}^{-1}$ consistent with the large COS deficit seen in airborne profiles in the northern hemisphere \citep{campbell_photosynthetic_2008,suntharalingam_global_2008,berry_coupled_2013}. Soil uptake, resulting from the presence of CA in soil microorganisms, is thought to be much smaller in magnitude than vegetation fluxes \citep{whelan_reviews_2018}. In the atmosphere, COS has also two chemical sinks: models indicate that about 100 $\mathrm{GgS.yr}^{-1}$ of COS is oxidized by OH in the low troposphere while 50 \pm 15 $\mathrm{GgS.yr}^{-1}$ is photolysed in the stratosphere \citep{whelan_reviews_2018}. The largest sources of COS are from human activities and the ocean, with minor contributions from biomass burning (50-100 $\mathrm{GgS.yr}^{-1}$), \citep{glatthor_global_2017,stinecipher_biomass_2019}). The oceanic source has been estimated between 200 and 400 $\mathrm{GgS.yr}^{-1}$ \citep{lennartz_direct_2017,lennartz_monthly_2020}. The missing source is unlikely to arise from direct ocean emissions since the ship cruises have recorded a sub-saturation of tropical sea waters with respect to COS \citep{lennartz_direct_2017}. COS production from atmospheric oxidation of dimethyl sulfide (DMS) and carbon disulfide (CS_2) are two other candidates that may

a supprimé: hydrolyzed

a mis en forme : Français

a supprimé: are

a mis en forme : Français

a supprimé: into

a mis en forme : Français

support the missing source, as they have been reported to peak over the tropics. Recently, \cite{lennartz_monthly_2020} developed a mechanistic model to simulate COS emissions via CS_2 and estimated a global source of 70 GgS.yr^{-1} , too low to support the missing source. However, this model still relies on many assumptions and has limitations such as the lack of oceanic horizontal transport. As for the emissions through DMS , the oxidation yield is currently deduced from experiments carried out under conditions which are not representative of the atmospheric environment with high DMS concentrations and without NO_x at 298 K \cite{barnes_ftir_1996}. The recent identification of novel DMS oxidation products \cite{berndt_fast_2019,veres_global_2020} could challenge our current understanding of the mechanistic links between DMS and COS formation into the atmosphere. Regarding the anthropogenic emissions, the inventory from \cite{kettle_global_2002} used by most top-down studies has been demonstrated to be incomplete \cite{blake_carbonyl_2004,du_important_2016}. The anthropogenic inventory has been revised upward from 200 GgS.yr^{-1} to 400 GgS.yr^{-1} , with the largest source shifting from North America to Asia \cite{zumkehr_global_2018}. Yet, firn air sampled in Antarctica and Greenland suggests that anthropogenic emissions are still underestimated and are closer to 600 GgS.yr^{-1} \cite{aydin_anthropogenic_2020}.

a supprimé: \$

a supprimé: \$

a mis en forme : Français

a mis en forme : Français

a supprimé: leading to a high oceanic source

a mis en forme : Français

a supprimé: Groenland

a mis en forme : Français

As an alternative to modelling direct emissions, attempts have been made to constrain the COS budget through inverse or "top-down" approaches. With the help of a transport model and a priori information, these approaches adjust the surface fluxes to better match simulated atmospheric concentrations with observations. Previous top-down assessments of the COS budget identified the missing source as likely being from the ocean, with a total oceanic release between 500 and 1000 GgS.yr^{-1} \cite{suntharalingam_global_2008,berry_coupled_2013,kuai_estimate_2015,laouis_new_2015}. This finding is consistent with the high concentrations of COS observed over tropical waters \cite{montzka_global_2007,glattthor_tropical_2015,kuai_estimate_2015}, but remains preliminary due to the scarcity of observations \cite{ma_inverse_2021}. Top-down approaches have so far followed two computational strategies: the analytical strategy directly computes the closed-form solution to the inverse problem and is in principle reserved for small inverse problems, while the variational strategy can tackle larger problems by iteratively reaching the neighborhood of the closed-form solution. The analytical inverse system used by \cite{berry_coupled_2013} calibrated a single scaling factor for the oceanic source per latitudinal band. \cite{laouis_new_2015} used a similar technique but they optimized each term of the COS budget at an annual scale from COS surface measurements, applying one scaling factor per COS component. When assimilating Tropospheric Emissions Spectrometer (TES) satellite retrievals, \cite{kuai_estimate_2015} divided the tropics into several regions and optimized one scaling coefficient of the oceanic source per region. Recently, \cite{ma_inverse_2021} used a variational inverse system to optimize the COS surface fluxes at each pixel of their model grid using COS surface measurements, but still had to apply a large auto-correlation length to compensate for the sparse

observation network. These systems have assimilated only COS atmospheric measurements.

Here, we present an update of the [\citet{launois_new_2015}](#) analytical system that jointly assimilates COS and CO_2 measurements, using recent prior fluxes and many more degrees of freedom given to the inversion. The new system makes it possible to optimize each process by region and by month and in particular, the GPP for each of the 15 Plant Functional Types (PFT) of the ORCHIDEE (ORganizing Carbon and Hydrology In Dynamic Ecosystems, [\citet{krinner_dynamic_2005}](#)) terrestrial model.

We assume a linear relationship between GPP and biospheric COS uptake under a leaf relative uptake (LRU) approach. We also take advantage of the additional sophistication of the inversion system to assimilate COS measurements together with CO_2 measurements, in order to constrain both GPP and respiration fluxes. Our study period spans 12 years, from 2008 to 2019.

The objectives of our study are threefold:

- \begin{enumerate}
- \item{Evaluating the analytical inverse system applied for the first time to the joint assimilation of COS and CO_2 measurements from a technical point of view,}
- \item{Providing an improved COS budget estimate,}
- \item{Providing improved estimates of GPP and respiration based on the joint assimilation of COS and CO_2 measurements.}
- \end{enumerate}

After a description of the inverse system and its setup in Section 2, inverse results will be shown in Section 3 with an emphasis on the global budget and on the seasonal cycle of the optimized fluxes. In Section 4, the fluxes will be prescribed to the LMDz atmospheric transport model and the resulting concentrations will be evaluated against independent observations over North America, the Pacific Ocean, Japan and France. We will also compare the simulated concentrations against Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) [\citep{fischer_mipas_2008}](#) retrievals over the tropics. Finally, we will discuss the potential and limitations of this inverse system to constrain the GPP with COS observations.

\section{Data and method}

\subsection{Atmospheric transport}

We simulate the global atmospheric transport at spatial resolution $3.75^\circ \times 1.9^\circ$ (longitude times latitude) with 39 layers in the vertical, based on the general circulation model of the Laboratoire de Météorologie Dynamique, LMDz [\citep{hourdin_lmdz6a_2020}](#). LMDz6A is our reference version: it was prepared for the 6th Climate Intercomparison Project (CMIP6) as part of the Institut Pierre-Simon Laplace Earth system model. [\citet{remaud_impact_2018}](#) evaluated more specifically the skill of the model to represent the transport of passive tracers. We use the offline version of the LMDz code, which was created by [\citet{hourdin_use_1999}](#) and adapted by [\citet{chevallier_contribution_2005}](#) for atmospheric inversions. It is driven by air mass fluxes calculated by the complete general circulation model, run at the same resolution and nudged here towards winds from the fifth generation of meteorological analyses of the European Centre for

a supprimé: %Sulfur isotope measurements sampled in Japan interpreted by back trajectories suggest important anthropogenic emissions coming from China in winter.¶

a mis en forme : Français

a mis en forme : Français

a supprimé: %In doing so, we have developed an analytical inverse system aimed at optimizing the COS and the CO2 budgets from surface atmospheric measurements and prior knowledge of its components. Given that both gases exhibit similar spatio-temporal and seasonal behaviour over continental areas [\citep{montzka_can_2004,montzka_global_2007,parazoo_covariation_2020}](#), a joint assimilation of COS and CO2 is particularly interesting. ...

a mis en forme : Français

a supprimé: %This studies is a first step toward an optimization of the ORCHIDEE parameters using multi-stream data (atmospheric measurements, fluxNet...).¶

a mis en forme : Français

a supprimé: inversion

a mis en forme : Français

Medium-Range Weather Forecasts (ERA5). The off-line model only solves the mass balance equation for tracers, which significantly reduces the computation time.

For the sake of simplicity, we refer to LMDz as the offline model in the following.

%Analytical versions of the LMDz tangent-linear and adjoint operators have been developed, so that operations $\text{M}\{\mathbf{x}\}$ and $\text{M}^T\{\mathbf{y}\}$, with M the Jacobian matrix of LMDz, \mathbf{x} a vector of input variables of LMDz (i.e. tracer surface fluxes and initial tracer values), and \mathbf{y} a vector of size the number of output variables (i.e. the four-dimensional atmospheric concentrations), can be computed at the machine epsilon despite conditional statements in the LMDz code:

LMDz is weakly non-linear with respect to the surface fluxes, %further to following the use of slope limiters in the \cite{van_leer_towards_1977} advection scheme which ensures monotonicity.

Analytical versions of the LMDz tangent-linear and adjoint operators have been developed.

Those codes respectively perform operations $\text{M}\{\mathbf{x}\}$ and $\text{M}^T\{\mathbf{y}\}$, with M the Jacobian matrix of LMDz, \mathbf{x} a vector of input variables of LMDz (i.e. tracer surface fluxes and initial tracer values), and \mathbf{y} a vector of size the number of output variables (i.e. the atmospheric concentrations at observation location and time), at the machine epsilon despite conditional statements in the LMDz code.

In our study, we assimilate LMDz to one of its Jacobian matrices: we linearized LMDz beforehand around a top-down estimation of the CO_2 surface fluxes from the Copernicus Atmosphere Monitoring Service (<https://atmosphere.copernicus.eu/>). "We checked that this linearization using CO_2 was still valid for COS fluxes and expected COS flux increment λ patterns through a test for the tangent linear model. Specifically, we checked the alignment of the non linear evolution of $\text{M}(x_0 + \lambda)$ with the linear evolution $\text{M}(\lambda)$ for the COS fluxes x_0 (not shown). The archived Jacobian matrix was generated by the adjoint code of LMDz. This way of doing is in principle an improvement over previous COS studies with LMDz \cite{launois_new_2015,peylin_new_2016} which used a rough approximation of the adjoint, $\text{M}^T\{\mathbf{y}\}$, called "retro-transport", in which the direction of time was simply reversed in LMDz without strict inversion of the order of calculations \cite{hourdin_eulerian_2006}. In addition, we use a much more recent version of LMDz here (LMDz6A, \cite{remaud_impact_2018}, vs. LMDz3, \cite{hourdin_lmdz4_2006}), and at higher resolution, in particular in the vertical (39 vs. 19 layers). The adjoint code of LMDz was initially developed for variational inversion, but we use this facility for the first time with LMDz in an analytical framework, to calculate the rows of the Jacobian Matrix M that correspond to the places where, and the times when, we have observations to assimilate. By definition, each value of M is a derivative of an output tracer concentration relative to an input surface flux or initial tracer value. More specifically, we use one adjoint run $\text{M}^T\{\mathbf{y}\}$ for each observation to assimilate, with the elements of \mathbf{y} set to zero or one. We use the Community Inversion Framework (CIF, \cite{berchet_community_2020}) to manage these computations.

a supprimé: patterns

a supprimé: ,

a supprimé: which

In our case, $\mathbf{M}^T \mathbf{y}^*$ represents the sensitivity of an output tracer concentrations to an input surface flux as the impact of the prescribed initial concentration is very small after one month (not shown).

Thus, we link the model concentration \mathbf{y} at a given station and time to a source \mathbf{S} through the relation [\citep{hourdin_eulerian_2006-1}](#):

$$\mathbf{y}^* = (\mathbf{Y} - \mathbf{H}(\mathbf{X})) \mathbf{R}^{-1} \mathbf{S}.$$

[\begin{equation}](#)

$$\mathbf{y} = \mathbf{H}_0^T \mathbf{C}|_{t_0} + \mathbf{H}_s^T \mathbf{S} \approx \mathbf{H}_s^T \mathbf{S}$$
[\label{EqAdjoint}](#)
[\end{equation}](#)

with \mathbf{H}_0^T and \mathbf{H}_s^T being the sensitivity of the model concentration to the initial conditions $\mathbf{C}|_{t_0}$ and to the sources \mathbf{S} , respectively. In our case, the source \mathbf{S} is a monthly source and the first term of the equation is negligible (not shown). Such simplification assumes that the initial air mass is homogeneously distributed within the atmosphere after one month. Note that the equation [\ref{EqAdjoint}](#) applies only in the case of linear transport model.

The relation [\ref{EqAdjoint}](#) results from the time symmetry of the atmospheric transport provided that the mass of a passive tracer is conserved along trajectories and the atmospheric transport is linear [\citep{hourdin_eulerian_2006-1}](#). Note that temporal discretisation of the transport equations and, in particular, the presence of slope limiters in the [\cite{van_leer_towards_1977}](#) advection scheme implemented in LMDz break the linearity and the time symmetry property [\citep{hourdin_eulerian_2006}](#). The linearity approximation has been validated by performing the tangent linear test using an increment and a surface flux equal to the biospheric CO2 flux (not shown). An other method would have been to solve as many forward problems as fluxes within the optimization vector. Though, it is more efficient to solve one adjoint or inverse equation from which the same sensitivity of the concentrations to the the source can be derived (Ending 2002). This allows us to test many configurations of the optimization vectors without running the model each time.

By reducing the slope between two grid points, the slope limiters aims at attenuating numerical diffusion.

In practice, we considered 8-day-average synthetic observations at each selected measurement site (see Section 2.2.1) between 2008 and 2019. The implication is that the atmospheric transport model can not represent the temporal variability within a week. For sites below 1000 m above sea level, only afternoon observations were used as the models do not simulate the accumulation of the tracers in the nocturnal boundary layer well [\citep{locatelli_atmospheric_2015}](#). For elevated stations, both daytime and early nighttime observations were discarded because coarse-resolution models cannot represent the advection of air masses during the day by upslope winds over sunlit mountain slopes in the afternoon [\citep{geels_comparing_2007}](#). After corresponding forward runs that defined the tracer linearization trajectories, the adjoint model was run nine months backward in time from measurement time for each of these synthetic observations (with appropriate \mathbf{y}^*), giving as output the series of integrated sensitivities of the corresponding measurement with respect to the surface fluxes throughout the nine months and to the concentrations at the initial point in time.

For times prior to nine months, we have in fact

a supprimé: $\text{\textbf{MR}}$: Le paragraphe suivant est long et très technique. Au cas où le journal choisi n'est pas gmd, peut-être serait-il mieux dans le supplementary?

a supprimé: for each 8-day period

a supprimé: (forward-counted).

not used the exact adjoint values. Instead, we extended the databases of adjoint outputs for the surface fluxes beyond the nine-month windows with two parts: (i) monthly adjoint outputs between months 9 and 24 taken from computations for the year 2017, and (ii) beyond 24 months, a globally-homogeneous value (i.e. 1 GtC emitted at the surface is translated to an average concentration of 0.38 $\mu\text{mol.mol}^{-1}$, or parts per million, ppm). We have verified that the CO_2 and COS concentrations obtained by the resulting Jacobian matrix ($\text{M} \cdot \mathbf{x}$) match well the one produced by the full LMDz transport model over the period (See Fig. S3 of the Supplementary [materiel](#)).

a supprimé: given

a supprimé: S2

In total, we have computed 15 stations \times 12 years \times 2 weeks \times 12 months adjoint computations of 8 process time hours each on a local parallel cluster. 2 weeks correspond to the typical frequency of the COS measurements.

In the past, the matrix \mathbf{H} has been computed using the retro-transport of the LMDz atmospheric model instead of the adjoint transport (Peylin et al., 2005; Launois et al., 2015; Peylin et al., 2016). Even though the adjoint transport is mathematically equivalent to the retro tracer transport (Hourdin et al., 2006), the two quantities are conceptually different and obtained through two different ways. The adjoint transport is computed using systematic techniques reversing the sequence of individual processes of the transport model. The retro transport, on the other hand, describes the reverse temporal evolution of an atmospheric concentration conserved along a trajectory and is calculated with processes following the same order that the forward model by changing the signs of the equations. An other novelty is the use of LMDz6A version of LMDz (Renaud et al., 2018) for the computation of the adjoint, the version used in (Peylin et al., 2005; Launois et al., 2015) was LMDz3 (Hourdin et al., 2006).

As explained below in Section 2.4.2, LMDz is complemented here for the modelling of COS in the atmosphere by a chemical sink, represented by a surface flux.

`\subsection{Observations and data sampling}`

`\subsubsection{Assimilated observations: CO_2 and COS surface sites}`

We used the NOAA/ESRL measurements of both CO_2 and COS between 2008 and 2019 at 15 sites whose location is depicted on Fig. \ref{Obs}: Cape Grim, Australia (CGO, 40.4°S, 144.6°W, 164 m above sea level, asl), American Samoa (SMO, 14.2°S, 170.6°W, 77 m asl), Mauna Loa, United States (MLO, 19.5°N, 155.6°W, 3397 m asl), Cape Kumukahi, United States (KUM, 19.5°N, 154.8°W, 3 m asl), Niwot Ridge, United States (NWR, 40.0°N, 105.54°W, 3475 m asl), Wisconsin, United States (LEF, 45.9°N, 90.3°W, 868 m asl, inlet is 396 m above ground on a tall tower), Harvard Forest, United States (HFM, 42.5°N, 72.2°W, 340 m asl, inlet is 29 m aboveground), Barrow, United States (BRW, 71.3°N, 155.6°W, 8 m asl), Alert, Canada (ALT, 82.5°N, 62.3°W, 195 m asl), Trinidad Head, United States (THD, 41.0°N, 124.1°W, 120 m asl), Mace Head, Ireland (MHD, 53.3°N, 9.9°W, 18 m asl), Weizmann Institute of Science at the Arava Institute, Ketura, Israel (WIS, 29.96°N, 35.06°E, 151 m asl), Palmer Station, Antarctica, United States (PSA, 64.77°S, 64.05°W, 10.0 m asl), South Pole, Antarctica, United

States (SPO, 89.98° S, 24.8° W, 2810.0 asl) and since mid-2004 at Summit, Greenland (SUM, 72.6° N, 38.4° W, 3200 m asl). The COS samples have been collected as pair flasks one to five times a month since 2000 and have then been analysed with gas chromatography and mass spectrometry detection. Most measurements have been performed in the afternoon between 11 and 17h local time when the boundary layer is well mixed. The COS measurements have been kept for this study only if the difference between the pair flasks is less than 6.3 ppt. With the exception of site WIS, most sites have at least one measurement per month for 11 months out of 12 within each year over the years 2008–2019 (see Figure S17). These data represent an extension of the measurements first published in [citet{montzka global 2007}](#).

a supprimé: citep

The Jacobian Matrix M described in the previous section reveals the information content provided by these measurements in terms of tracer surface flux. In particular, it helps to identify to what extent each region of the globe is seen by the observations and therefore, it provides an indication of the details needed or not in the flux variables to be optimized. The transport sensitivities to the sources integrated over two months are represented in Fig. [\ref{Obs}](#) on average for the period 2016–2019. The zonal distribution of sensitivities reflects the zonal atmospheric circulation at mid and high latitudes, with the north (south) stations seeing the entire domain above (under) 30° N. The tropics are not well constrained by the observations: the tropical circulation, mainly vertical, limits the extension of the footprint zone around SMO and MLO, leaving the Indo-Pacific region for the most part unconstrained. However, the tropical areas are slightly constrained by well mixed air masses coming from remote stations (see Fig. S2). We also see that the southern and northern oceans are also more constrained by the observations than the continents, with the exception of North America which is relatively well covered by the measurements. Fig. [\ref{Obs}](#) suggests the need to separate between each latitudinal band (Tropics, northern and southern latitudes) and also between oceans and continents in the inversion.

a supprimé: Tropics

a supprimé: our

a supprimé: inversion will not be able to partition the COS and CO_2 components in this zone. The...

a supprimé: north

a supprimé: south

`\begin{figure}[h!]`

`\centering`

`\includegraphics[scale=0.7]{figs/All_station.pdf}`

`\caption{Annual climatology of Jacobians computed by the adjoint of the LMDz model: map of the partial derivatives, in $\text{ppm}/(\text{kg}/\text{m}^2/\text{s})$, of a monthly mean concentration at all stations from the NOAA network with respect to CO_2 surface fluxes in the two previous months. The yellow dots denote the location of the surface sites. The site KUM is not depicted as it has the same coordinates than MLO but at sea level.}`

a supprimé: weekly

a supprimé: month

`\label{Obs}`

`\end{figure}`

Note that, if computed with respect to the COS fluxes, the annual climatology of Jacobian shown on Fig. [\ref{Obs}](#) would have the same spatial pattern but with a different unit given that the atmospheric transport is linear and there are no atmospheric chemical reactions. In Section 3, even though the model simulations are not compared to measurements, the model sampling still refers to some observation selection (in the afternoon for the zonal-mean profiles, or following a satellite retrieval pattern for the total column), as indicated in the corresponding text.

\subsubsection{Independent observations}

An ensemble of independent observations - i.e. data that are not assimilated in LMDz - is used to evaluate the fluxes retrieved by our inverse system. We focus here on the observations used to evaluate the COS and the GPP fluxes.

(MR: Est ce que je dois introduire les observations de la NOAA si elles sont montrées uniquement dans le Supplément et qu'elles sont mentionnées une fois dans le texte (en renvoyant au Supplément)?)

The first observation program is the HIAPER Pole-to-Pole Observations (HIPPO, \cite{wofsy_hiaper_2011}). HIPPO consisted of five aircraft transects of many trace gas measurements, including for COS and CO_2 , in the troposphere over the Western Pacific: HIPPO 1 (January 2009), HIPPO 2 (November 2009), HIPPO 3 (March-April 2010), HIPPO 4 (June 2011) and HIPPO 5 (August 2011). The HIPPO measurements were made from flask and in-situ measurements by NOAA and the University of Miami. They were rescaled to be consistent with the calibration scale used for the NOAA surface network results.

In order to assess the North-South latitudinal COS gradient over Japan, surface measurements for winter and summer 2019 at three sampling sites in Japan from \cite{hattori_constraining_2020} have been used as well: Miyakojima (24°80'N, 125°27'E), Yokohama (35°51'N, 139°48'E), and Otaru (43°14'N, 141°16'E). In winter, the Miyakojima site samples air masses strongly influenced by anthropogenic emissions from Chinese megacities including Beijing and Shanghai, while Yokohama and Otaru are only influenced by the northern periphery of China. During the summer, all sites mainly sample ocean air masses coming from southeastern Japan \cite{hattori_constraining_2020}.

The French sampling site, GIF (48°42'N - 2°08'E), is located about 20 km to the south west of Paris where ground level COS has been monitored on a hourly basis since August 2014 \cite{belviso_top-down_2020}. According to the recent COS global gridded anthropogenic emission inventory of \cite{zumkehr_global_2018}, the Paris region is an important source of COS (791 MgS/yr, J. Stinecipher personal communication November 2018) where its indirect emissions from the rayon industry largely overpass its direct emissions from the aluminium industry and traffic. These estimates have been challenged by \cite{belviso_top-down_2020}.

The location of the HIPPO data, NOAA airborne profiles, Japanese and GIF sites are depicted in Figure \ref{ObsI}.

The fourth observation program is made of the satellite COS retrievals from MIPAS. The MIPAS spectrometer measured limb-emission spectra for several trace gases in the mid-infrared \cite{fischer_mipas_2008} from the European Space Agency (ESA) Environmental Satellite (ENVISAT) between March 2002 and 2012. The IMK/IAA retrieval processor operated at KIT-IMK was used to calculate the COS profiles of data version \$V5R_OCS_221/222\$ which were used for this work \cite{glatthor_tropical_2015,glatthor_global_2017}. The number of vertical layers of the MIPAS retrievals is 60. Between

a supprimé: measurements have

a supprimé: citep

altitudes 7 and 25 km the accuracy of the COS profiles is around 50 ppt in the absence of clouds (in particular deep-convective ones) \citep{glatthor_tropical_2015}.

The evaluation against airborne measurements \citep{glatthor_tropical_2015} and latter against MkIV and SPIRALE profiles MkIV \citep{glatthor_global_2017} shows .

Last, the SIF satellite retrievals from the Global Ozone Monitoring Experiment-2 (GOME-2) make it possible to evaluate the seasonality of GPP inferred by inverse modelling for each PFT. SIF represents the amount of light reemitted by chlorophyll molecules as a byproduct of photosynthesis. Satellite-based SIF data is considered as a proxy for the GPP of terrestrial ecosystems at large spatial-temporal scales \citep{frankenbergnew_2011,guanter_retrieval_2012,zhang_model-based_2016,yang_solar-induced_2015,li_chlorophyll_2018}. We use release number 28 of the NASA GOME-2 (Global Ozone Monitoring Experiment-2 onboard the MetOp-A satellite) daily corrected SIF product \citep{joiner_global_2013,joiner_new_2016}. The dataset is available at: https://avdc.gsfc.nasa.gov/pub/data/satellite/MetOp/GOME_2F/v28/. We used the monthly level 3 product gridded at a 0.5 $^{\circ}$ resolution between years 2008 and 2019. This GOME-2 SIF product was shown to be very similar in terms of seasonality and magnitude (after spectral scaling) to the reference Orbiting Carbon Observatory (OCO-2, launched in 2014, \citep{sun_overview_2018}) data \citep{bacour_differences_2019}. For each PFT, we average all the grid points within the LMDz grid points that have a fractional cover greater than 0.8. We lower this threshold to 0.3 for PFTs 7 (Boreal Broad-leaved Evergreen Forest), 8 (Boreal Broad-leaved Summergreen Forest), 9 (Boreal Needleleaf Summergreen Forest) and 15 (Boreal C3 grass). The PFTs are further defined in section 2.4.

```
\begin{figure}[h!]  
\centering  
\includegraphics[scale=0.7]{figs/map-independent.pdf}  
\caption{Location of the HIPPO airborne measurements, NOAA airborne  
platforms and surface sites in Japan and France that are used as  
independent observations for evaluating the inverse results. The HIPPO  
measurements have been averaged into bins of 10 $^{\circ}$  each. The NOAA  
airborne measurements are exploited in the Supplement.}  
\label{ObsI}  
\end{figure}
```

\subsubsection{Data sampling}

For each species and each measurement, the simulated concentration fields were sampled at the LMDz 3D grid box nearest to the observation location. As mentioned above, the observations at selected local times are assimilated as 8-day averages. For the independent observations, LMDz is sampled at the closest time from the observations. All observations are dry-air mole fractions calibrated relative to the compound World Meteorological Organization (WMO) mole fraction scale. Satellite retrievals, are dry-mole fractions tuned by the data providers to the compound World Meteorological Organization (WMO) mole fraction scale. For comparison, the corresponding dry-air variables in the model simulations are used.

When comparing MIPAS data with LMDz simulations, the a priori and vertical sensitivity of the retrievals must be taken into account. For each MIPAS retrieval, the modelled COS profiles have been

a supprimé: (or, for the satellite

a supprimé: ,

a supprimé:)

interpolated linearly to the MIPAS vertical resolution ([60 layers](#)) while ensuring the conservation of the column-average mixing ratio [\citep{chevallier_statistical_2015}](#). They were then smoothed with the corresponding MIPAS averaging kernels.

following the equation:

$$X_s = X_a + A[X_m - X_p].$$

where X_s , X_a , and X_m are smoothed, a priori, and model vertical profile, respectively, and A_k is the averaging kernel matrix.

The a priori profile for the COS retrievals is a zero profile [\citep{glatthor_tropical_2015}](#), hence it had not to be taken into account. As done in [\cite{glatthor_tropical_2015}](#), we focus here on the spatial distribution of the COS mixing ratio at the 250 hPa pressure level (still after convolution of the model with the averaging kernels) for the period 2008–2012. In order to dampen the random noise [in the MIPAS observations](#), we aggregate the retrievals into in $5^\circ \times 15^\circ$ latitude-longitude bins.

\subsection{Inverse framework}

Our inverse system seeks to estimate the amplitude of n sources or sinks of CO_2 and COS gathered in a vector \mathbf{x} by reducing the mismatch between the observed concentrations gathered in a vector \mathbf{y}^o and those simulated with the atmospheric transport model M forced by these sources and sinks. Together with an initial disaggregation operator (that converts the low-resolution control vector into gridded fluxes using gridded [\it reference fluxes](#), see section 2.5.1) and a sampling operator (see previous section), the [linearized](#) transport model M is part of the linear observation operator H that relates \mathbf{x} and the model-equivalent CO_2 and COS measurements \mathbf{y} at the sites shown in Fig. [\ref{Obs}](#):

$$\mathbf{y} = \text{H} \mathbf{x}$$

In order to regularize the inverse problem corresponding to Eq. (1), we use a Bayesian framework involving an a priori control vector, \mathbf{x}^b with its associated uncertainty statistics, summarized in covariance matrix B . Within the Gaussian assumption of the prior and [observation](#) errors, the solution of the inverse problem can be simply expressed by the following equation (see for instance [\cite{tarantola_1989}](#)) for the the posterior control vector \mathbf{x}^a and the uncertainty covariance matrix P^a :

$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^b + \text{B} \text{H}^T (\text{H} \text{B} + \text{H}^T \text{R} \text{H} + \text{R})^{-1} (\mathbf{y}^o - \text{H} \mathbf{x}^b) \\ \text{P}^a &= \text{B} - \text{B} \text{H}^T (\text{H} \text{B} + \text{H}^T \text{R} \text{H} + \text{R})^{-1} \text{H} \text{B} \end{aligned}$$

with R the error covariance matrix of the observations, encompassing measurement errors and H errors. Within the

a supprimé: observations

a supprimé: 19897

Gaussian assumption with no bias for all errors, the above solution minimizes the cost function :

```
\begin{align}
J(\mathbf{x}) = \frac{1}{2} [ (\mathbf{y}^o - \text{H} \mathbf{x})^T \text{R}^{-1} (\mathbf{y}^o - \text{H} \mathbf{x}) + (\mathbf{x} - \mathbf{x}^b)^T \text{B}^{-1} (\mathbf{x} - \mathbf{x}^b) ]
\label{CostF}
\end{align}
```

`\subsection{Gridded reference fluxes}`

In the following, we call "reference fluxes" the maps of CO_2 and COS fluxes that are used in the observation operator, the control vector \mathbf{x} being a low-resolution multiplier to these (see Section 2.5.1).

For use at resolution $3.75^\circ \times 1.90^\circ$, the maps of the following components of the CO_2 and COS fluxes have been interpolated from their native resolution. All projections conserved mass.

`\subsubsection{\text{CO}_2 fluxes} \label{FlxCO2}`

Our reference fluxes combine several information sources. Fossil fuel emissions are from the gridded fossil emission dataset GCP-GridFED (version 2019.1) (Jones et al. 2021). Biomass burning fluxes vary inter-annually and are described by the GFED 4.1s database (<https://www.globalfiredata.org/data.html>). Monthly air-sea CO_2 exchange is prescribed from the Copernicus Marine Environment Service database (Denzil-Sommer et al. 2019). The GPP and respiration fluxes have been simulated at a resolution of 0.5° both in longitude and latitude by the ORCHIDEE land surface model (Krinner et al. 2005).

ORCHIDEE explicitly parameterizes the main processes influencing the water, carbon and energy balances at the interface between land surfaces and atmosphere. The vegetation is represented by 15 PFTs with a spatial distribution prescribed from the ESA Climate Change Initiative (CCI) land cover products (Poulter et al. 2015). The plant phenology is prognostic and PFT-specific. We used version 9 tuned for the CMIP6 exercise and forced by the global CRUJRA reanalysis at global scale (<https://sites.uea.ac.uk/cru/data/>) v1-v2, applying land-use change and realistic increase of CO_2 atmospheric concentration. Emissions from the land use and wood harvest have been included beforehand in the respiration term. Biomass burning emissions are not taken into account in this respiration term from ORCHIDEE. The yearly global GPP from ORCHIDEE amounts to 126.7 GtC.y^{-1} during 2008-2019. This value is within the range of the GPP estimates (106-137 GtC.yr^{-1}) based on photosynthesis proxies (see Table S1) (Beer et al. 2009, Beer et al. 2010, Welp et al. 2011, Al-Mohammad et al. 2017, Jasechko et al. 2019, Jung et al. 2020, Rymer et al. 2011, Badgley et al. 2019, Stocker et al. 2019). The PFTs and their acronyms are defined in Table (TLRU). Note that GPP, respiration, COS vegetation and soil fluxes are null within PFT 1 (base soil).

a supprimé: .

```
\begin{table}[h!]
\resizebox{\textwidth}{!}{%
\begin{tabular}{lccc}
PFT & & Acronym & \\
\hline
```

```

1 - Bare soil & BaS \\
2 - Tropical Broad-leaved Evergreen Forest & TrBrE \\
3 - Tropical Broad-leaved Raingreen Forest & TrBrR \\
4 - Temperate Needleleaf Evergreen Forest & TeNeE \\
5 - Temperate Broad-leaved Evergreen Forest & TeBrE \\
6 - Temperate Broad-leaved Summergreen Forest & TrBrS \\
7 - Boreal Needleleaf Evergreen Forest & BoNeE \\
8 - Boreal Broad-leaved Summergreen Forest & BoBrS \\
9 - Boreal Needleleaf Summergreen Forest & BoNeS \\
10 - Temperate C3 Grass & TeC3g \\
11 - C4 Grass & C4g \\
12 - C3 Agriculture & C3Ag \\
13 - C4 Agriculture & C4Ag \\
14- Tropical C3 grass & TrC3g \\
15- Boreal C3 grass & BoC3g \\
\end{tabular}
\caption{\label{TLRU} List of the PFTs as defined in the ORCHIDEE LSM.}
\end{table}

```

%For the sake of realism, we modify the GPP spatio-temporal distribution in latitude range 50S:50N in a such way that within a PFT, it has the spatial and seasonal variability of the monthly 0.5- ∞ GOME-2 SIF data (Solar Induced Fluorescence) while conserving the same annual total from ORCHIDEE. Given a vegetation fraction $X(p)$ for the PFT p , GPP_{SIF} for the month t within the PFT p are obtained from the transformation:

```

%\begin{equation}
% GPP_{SIF}(lat,lon,t)=SIF(lat,lon,t)\times X(p)
\frac{GPP_{SIF}}{GPP_{tot}}=\frac{SIF_{SIF}}{SIF_{tot}}
%\end{equation}

```

%with GPP_{tot} and SIF_{tot} the annual total over the PFT map. Beforehand, the SIF map was convolved with the vegetation fraction of each PFTs. SIF represents the amount of light reemitted by chlorophyll molecules as a byproduct of photosynthesis. Although the GPP-SIF relationship is unlinear and vegetation dependant in instantaneous leaf level measurements [citep{colombo_variability_2018}](#), spatial and temporal patterns of satellite retrieved SIF have been shown to be highly correlated with GPP at the biome and global levels [citep{yang_solar-induced_2015,li_chlorophyll_2018}](#).

\subsubsection{\mathrm{COS} fluxes}\label{FlxCOS}

The components of the COS budgets that are considered are biomass burning, soil emissions and [uptake](#), anthropogenic emissions, plant uptake, oceanic emissions and the atmospheric oxidation by the OH radical in the troposphere. Photolysis in the stratosphere, estimated to 30 $\mathrm{GgS.yr^{-1}}$ in the LMDz atmospheric transport model (not shown), and volcano emissions, in the range 23-43 $\mathrm{GgS.yr^{-1}}$, have been neglected [citep{whelan_reviews_2018}](#).

```

\begin{table}
\resizebox{\textwidth}{!}{%
\begin{tabular}{|l|c|cccccc|}
\hline
& PROCESSES & \citep{kettle_global_2002} & \citep{montzka_global_2007} & \citep{suntharalingam_global_2008} & \citep{berry_coupled_2013} & \citep{launois_new_2015} & \citep{ma_inverse_2021}&This study \\
\hline

```

a supprimé: sink


```

\hline \hline
SOURCES & Anthropogenic & 180 & 180 & 180 & 180 & 180 & 349 &
398 \\
& Oceanic & 296 & 296 & 230 & 876 \tnote{\dag} & 1027
&277&269 \\
& Biomass Burning & 38 & 106 & 70 & 136 & 70 & 136 &
53 \\
& Anoxic Soils & 26 & 26 & 26 & neglected & 101 & &
neglected & neglected \\
SINKS & Oxic soils & -130 & -130 & -130 & -355 & -510 & &
-236 \\
& Plant uptake & -238 & -1115 & -490 & -738 & ,\Aii1335 & -
1053 \tnote{\textasteriskcentered} & -657 \\
& Atmospheric chemical loss & -121 & -121 & -121 & -101 & -
100 & -141 & -100 \\
\hline \hline
TOTAL & & 52 & -757 & -234 & -1 & -567 & -432 \tnote{\ddag} & -273 \\
\hline

```

```

\end{tabular}
}
\caption{\label{TBud} Overview of the global budget of  $\mathrm{COS}$ .
Units are  $\mathrm{GgS.yr}^{-1}$ . \ \tnote{\dag} \footnotesize{In order
to provide a balanced  $\mathrm{COS}$  budget, the oceanic emissions from
\cite{kettle_global_2002} have been increased by 600  $\mathrm{GgS.yr}^{-1}$ }. \\
\tnote{\ddag} \footnotesize{This unknown source has been optimized using
a global 4-dimensional variational data-assimilation system.} \\
\tnote{\textasteriskcentered} \footnotesize{This number corresponds to
the soil and plant uptake combined.}
}
\end{table}

```

\subsubsection{\bf{Soil}}

Reference air-surface exchanges from oxic soils have been simulated by the steady-state analytical model of \cite{ogee_new_2016} implemented in the ORCHIDEE land surface model with the Zobler soil classification at a 0.5 degree both in longitudes and latitudes, \cite{abadie_2021}. This model is built on the assumptions that the soil atmosphere exchanges are governed by three processes, namely diffusion through the soil column, production and irreversible uptake via hydrolysis. The COS uptake reflects for the most part the activity of CA, ubiquitous in soil microorganism, which efficiently converts COS into $\mathrm{H}_2\mathrm{S}$ and CO_2 , similarly to what happens in plants. The CA activity is represented by the CA enhancement factor or f_{CA} , which is PFT-specific and has been calibrated against measurements performed by \cite{meredith_coupled_2018} on different biomes in the laboratory. The production term simulates the COS abiotic production from soils via the \cite{whelan_carbonyl_2016} model. Its exponential increase with temperature decreases the net soil uptake over the tropics and in mid-latitudes in summer. The soil model has been shown to be in better agreement with measurements than the \cite{berry_coupled_2013} model used in previous top-down studies. As for the contribution of anoxic soils \cite{whelan_salt_2013}, we have not taken them into

a supprimé: 4-

a supprimé: dimensional variational data-assimilation system. over the globe.)

a supprimé: .

a supprimé: means to reflect

a supprimé:

a supprimé: .). A complete validation of the soil fluxes in ORCHIDEE will be the topic of a future publication.

account in the absence of reliable emission maps
`\citep{whelan_reviews_2018}`.

`\subsubparagraph{\bf{Plant uptake}}`

We chose the empirical formulation of the CO_2 uptake by leaves from `\citep{sandoval-soto_global_2005}` given by the linear relationship

$$F_{\mathrm{CO}_2} = GPP \times \frac{v_{\mathrm{CO}_2}}{v_{\mathrm{CO}_2} + v_{\mathrm{CO}_2}}$$

`\begin{equation}`
`\frac{v_{\mathrm{CO}_2}}{v_{\mathrm{CO}_2}}`
`\label{eq1}`
`\end{equation}`

In this equation, F_{CO_2} and GPP are the CO_2 uptake and the CO_2 uptake (both in $\mathrm{ppm}/\mathrm{m}^2/\mathrm{s}$), respectively, C_{CO_2} and C_{CO_2} being the ambient air concentrations of CO_2 and CO_2 . v_{CO_2} and v_{CO_2} are the CO_2 and CO_2 leaf uptake velocities.

The ratio of uptake velocities of CO_2 compared to CO_2 is defined as the LRU: `\begin{equation}`

$$LRU = \frac{v_{\mathrm{CO}_2}}{v_{\mathrm{CO}_2}}$$

`\label{eqLRU}`
`\end{equation}`

We use a zero-order LRU approach (i.e. with no interaction between vegetation and CO_2 mixing ratio), given the complexity of an order approach (i.e. a coupled atmospheric CO_2 concentration - CO_2 flux calculation). To address this shortcoming, we use the time-evolving hemispheric means of the CO_2 and CO_2 atmospheric concentrations, C_{CO_2} and C_{CO_2} as done in `\citep{montzka_global_2007}`. They are computed from monthly means at selected stations or groups of stations weighted by the cosine of latitude of atmospheric boxes encompassing different site groupings in this way:

$$SH_{\mathrm{CO}_2} = [SPO \times 0.408 + CGO \times 0.770 + SMO \times 0.974] / [0.408 + 0.770 + 0.974]$$

`\begin{align}`
`SH_{\mathrm{CO}_2} = [SPO \times 0.408 + CGO \times 0.770 + SMO \times`
`0.974] / [0.408 + 0.770 + 0.974]`
`\end{align}`

`\begin{equation}`

$$NH_{\mathrm{CO}_2} = [(KUM + MLO) \times 0.970/2 + (LEF + NWR + HFM) \times 0.751/3 + (BRW + ALT) \times 0.402/2] / [0.970 + 0.751 + 0.402]$$

`\begin{align}`
`NH_{\mathrm{CO}_2} = [(KUM + MLO) \times 0.970/2 + (LEF + NWR + HFM) \times 0.751/3`
`\ \`
`+ (BRW + ALT) \times 0.402/2] / [0.970 + 0.751 + 0.402]`
`\end{align}`

`\end{equation}`

We have only made a distinction between C4 (LRU=1.21) and C3 plants (LRU=1.68) and disregarded the dependence on light and water vapor deficit that was observed at both leaf `\citep{stimler_relationships_2010}` and ecosystem scales

`\citep{maseyk_sources_2014, commane_seasonal_2015, kooijmans_influences_2019}`. Our LRU set is derived from `\citep{whelan_reviews_2018}` and uses, for C3 plants, the median value of 53 LRU data and, for C4 plants, the median value of 4 LRU data. This simplification is supported by `\citep{hilton_peak_2017, campbell_large_2017}` who showed that the uncertainty on the LRU parameter is of a second order importance compared to the uncertainties on the GPP and the other CO_2 fluxes. Moreover, `\citep{maignan_carbonyl_2020}` showed that using a mechanistic model or its LRU equivalent model (i.e. with a constant LRU per PFT in ORCHIDEE LSM) for the plant uptake leads to similar results when

a supprimé: pm

a supprimé: a one-

a supprimé: }

a supprimé: }

transporting the COS fluxes with LMDz and comparing the COS concentrations at stations of the NOAA network. A physical reason making the LRU simplification acceptable is that the observation sites sample plant sink signals from multiple parts of the day. We have not taken into account the epiphytes which can both emit and absorb COS depending on environmental conditions [\citep{kuhn_environmental_2000,rastogi_large_2018}](#).

The LRUs associated with each PFT are presented in Table [\ref{TLRU}](#).

In order to quantify the impact of the LRU on the inverse results, we use three different sets of PFT-dependent LRU presented in Table [\ref{TLRU}](#). The first set is derived from whelan and uses, for C3 plants, the median value of 53 LRU data and for C4 plants, the median value of 4 LRU data. The second is derived from [\cite{seibt_kinetic_2010}](#), using the LRU values displayed in their Table 3. The last one is from Maignan et al. who computed a LRU equivalent for each PFT based on a monthly climatology of modelled COS and GPP fluxes.

```

% \begin{table}[h!]
% \resizebox{\textwidth}{!}{%
% \begin{tabular}{lccc}
% PFT & LRU & \cite{whelan_reviews_2018} & \\
% \hline
% 1 - Bare soil & & 0.00 & \\
% 2 - Tropical Broad-leaved Evergreen Forest & & 1.68 & \\
% 3 - Tropical Broad-leaved Raingreen Forest & & 1.68 & \\
% 4 - Temperate Needleleaf Evergreen Forest & & 1.68 & \\
% 5 - Temperate Broad-leaved Evergreen Forest & & 1.68 & \\
% 6 - Temperate Broad-leaved Summergreen Forest & & 1.68 & \\
% 7 - Boreal Needleleaf Evergreen Forest & & 1.68 & \\
% 8 - Boreal Broad-leaved Summergreen Forest & & 1.68 & \\
% 9 - Boreal Needleleaf Summergreen Forest & & 1.68 & \\
% 10 - Temperate C3 Grass & & 1.68 & \\
% 11 - C4 Grass & & 1.21 & \\
% 12 - C3 Agriculture & & 1.68 & \\
% 13 - C4 Agriculture & & 1.21 & \\
% 14- Tropical C3 grass & & 1.68 & \\
% 15- Boreal C3 grass & & 1.68 & \\
% \end{tabular}
% \caption{\label{TLRU} Table of LRU values per PFT, derived from
% \cite{whelan_reviews_2018}. }
% \end{table}

```

\subsubsection{Anthropogenic fluxes}

For anthropogenic fluxes, we use the inventory of [\cite{zumkehr_global_2018}](#) for the period 1980-2012 that corresponds to a global source of 398 GgS.yr^{-1} (range of 223 - 586 GgS.yr^{-1}) for the period 2009-2019. Emissions after 2012 are taken from the year 2012. The inventory accounts for direct COS emissions and indirect emissions through the oxidation of CS_2 into the atmosphere. The considered emissions are, in order of importance: emissions from rayon (staple and yarn) industry, residential coal, pigments, aluminium melting, agricultural chemicals and tires. Compared to [\cite{kettle_global_2002}](#), the majority of the sources has shifted over time from the US to China which encompasses now 45% of the total emissions.

a supprimé: 406

a supprimé: in year 2012

a supprimé: Are

\subsubparagraph{\bf{Ocean}}

COS is directly emitted by the ocean in places where the sea water is saturated in COS . Emissions typically happen in summer in high latitudes. COS is also indirectly emitted through the oxidation of DMS and CS_2 in the atmosphere, which are both produced in seawater. We use the indirect and direct COS emissions from \cite{lennartz_direct_2017,lennartz_monthly_2020} whose total emissions account for $285 \pm 10 \text{ GgS.yr}^{-1}$. In these, direct, indirect emissions via CS_2 , indirect emissions via DMS from the global ocean account for $130 \pm 80 \text{ GgS.yr}^{-1}$, $74 \pm 120 \text{ GgS.yr}^{-1}$ and $65 \pm 110 \text{ GgS.yr}^{-1}$. These emissions have been all computed using box models calibrated with ship-borne measurements made in different parts of the globe \cite{lennartz_direct_2017,lennartz_monthly_2020}. The DMS emissions are taken from the \cite{lana_updated_2011} climatology. The latitudinal distributions of each of the three terms of the oceanic emissions are depicted on Fig. \ref{FF2}.

```
\begin{figure}[H]
\centering
\includegraphics[scale=0.6]{figs/Prior_ocean.pdf}
\caption{Zonal mean distribution of the prior oceanic fluxes as a function of latitude averaged over the year 2010. The direct  $\text{COS}$  emissions are shown in blue whereas the indirect  $\text{COS}$  emissions through  $\text{DMS}$  ( $\text{CS}_2$ ) are depicted in brown (green).}
\label{FF2}
\end{figure}
```

We have not considered DMS and CS_2 as separate tracers as done in \cite{ma_inverse_2021}. CS_2 has a lifetime estimated between 4 days \cite{khan_global_2017} and 12 days \cite{khalil_global_1984} and DMS has a lifetime of 1.2 days. For the sake of simplicity, the oxidation of CS_2 and DMS by OH has been assumed to happen instantly in the atmosphere.

\subsubparagraph{\bf{Biomass burning}}

We use the inventory emissions from \cite{stinecipher_biomass_2019} with a global estimate of $60 \pm 37 \text{ GgS.yr}^{-1}$ for the period 1997-2016. These authors used CO as a reference species to compute the COS biomass burning emissions. To do that, they combined emission factors of COS to CO from the literature and applied them to the CO emissions. These CO emissions were computed beforehand from the GFED Global Fire Emissions Database (GFED version 4, <https://www.globalfiredata.org/>). The resulting biomass emissions are classified into four categories: savanna and grassland, boreal forests, temperate forests, tropical deforestation and degradation, peatland fires, and agricultural waste burning. The savanna was shown to be the largest contributor to the global biomass burning emissions and therefore to the overall uncertainty. These new estimates are lower than the previous estimate of open burning emissions. The latter were also positively biased by a strong emission factor derived from measurements over the peatlands. Moreover, their weak inter-annual variability was shown to better reproduce the annual trend in atmospheric concentration

a supprimé: CS_2 , indirect emissions via DMS from the global ocean account for $130 \pm 80 \text{ GgS.yr}^{-1}$, $74 \pm 120 \text{ GgS.yr}^{-1}$ and $65 \pm 110 \text{ GgS.yr}^{-1}$. These emissions have been all computed using box models calibrated with ship-borne measurements made in different parts of the globe

a supprimé:). Direct COS emissions are shown in green.)

a supprimé: estimation as these

at the Jungfraujoch station, the long-term trend being primarily driven by changes in anthropogenic activity \citep{zumkehr_gridded_2017}. It should be also noted that, compared to the \cite{KettleGlobal2002} inventory, the inventory emissions from \cite{StinecipherBiomass2019} do not include biomass burning sources from agriculture residues and biofuels. The latter were estimated to be about 3 times as large as open burning emissions \cite{CampbellAtmospheric2015}.

\subsubparagraph{\bf(OH sink)}

Since the highest reaction rate is close to the surface \citep{KettleGlobal2002}, we represent the OH sink by a surface flux. As done in \cite{LaunoisNew2015}, we take the spatial patterns of monthly maps of the OH radical concentrations and we distribute both horizontally and temporally a total annual tropospheric sink of $100 \mathrm{GgS.yr}^{-1}$, suggested by previous estimates \cite{KettleGlobal2002, BerryCoupled2013-1}. We use monthly maps of OH radical concentrations from an update of \cite{HauglustaineInteractive2004}.

\subsection{Inversion configuration}

\subsubsection{Control vector}

Our inversion window covers 12 years. The spatiotemporal resolution of the control vector $\{\mathbf{x}\}$ over this period represents a compromise between the assumed resolution of the errors of the reference fluxes, the expected resolution of the flux increments that can be inferred by the sparse site distribution (see Figure \ref{Obs}), and considerations on computing time.

Typically, a large control vector (i.e. many controlled regions and types of emission) may represent the complexity of reality better than a small control vector (i.e. few regions and emission processes), but also increases the inversion calculation load without always improving inversion skill, given the scarce and uneven observation network. The variables in the control vector are therefore all multipliers of the above-described gridded reference fluxes, as described as follows, rather than grid-point fluxes themselves. The choice of multipliers rather than increments implies that the initial sub-control-scale patterns are kept. The prior control vector $\{\mathbf{x}^{\{b\}}\}$ is simply a vector of ones.

a supprimé: one

We control COS oceanic fluxes in three latitudinal bands : the tropics, the northern latitudes and the southern latitudes. This separation allows the inverse system to modify the latitudinal distribution of the reference emissions, which remains subject to large uncertainties, while preserving the prior longitudinal patterns. This amounts to saying that the coastal sites located in the northern hemisphere constrain the total oceanic emissions over the whole northern hemisphere above 30°N . The ocean emissions are only modified within each of the three latitudinal bands by a single specific factor. Because the role of indirect COS emissions through DMS is still a matter of debate (Von Hobe, M., personal communication), we take all ocean emissions as a whole. On the continents, for respiration, GPP and soil fluxes, we distinguish the two hemispheres for eight of the 15 PFTs which are present in both (4, 5, 6, 10, 11, 12, 13, see Table \ref{TLRU}) to take into account the different seasonality. For the anthropogenic COS emissions, we control a single annual emission coefficient

and rely on the reference distribution of sources between Europe, Asia and America: the lack of observations in the Asia-Pacific region does not allow us to separately optimize Asian emissions. All parameters are optimized on a monthly scale with the exception of anthropogenic emissions which are assumed to be constant throughout the year.

For the CO_2 , we neglect the uncertainty on the oceanic, fire and anthropogenic CO_2 emissions compared to that of the sum of the respiration and GPP. The parameters of the control vector are described in Table \ref{TPar}.

```

\begin{table}[h!]
\begin{tabular}{l|cccccc}
\hline
Parameters & species & number of PFTs & number of regions & frequency & units & number of parameters \\
\hline
Oceanic fluxes &  $\mathrm{COS}$  & 0 & 3 & \text{\textcircled{d}} & monthly & GgS & 36 \\
\hline
Soil fluxes &  $\mathrm{COS}$  & 15 & 2 & \text{\textcircled{d}} & monthly & GgS & 4140 \\
GPP &  $\mathrm{CO}_2$  -  $\mathrm{COS}$  & 15 & 2 & \text{\textcircled{d}} & monthly & GtC & 4140 \\
Respiration &  $\mathrm{CO}_2$  & 15 & 2 & \text{\textcircled{d}} & monthly & GtC & 4140 \\
\hline
Anthropogenic emissions &  $\mathrm{COS}$  & 0 & 1 & annual & GgS & 1 \\
\hline
Biomass burning emissions &  $\mathrm{COS}$  & 0 & 1 & monthly & GgS & 12 \\
\hline
Background concentration &  $\mathrm{CO}_2$  & & & 1 & 1 & & & a supprimé: monthly \\
ppm & 1 & & & & & & & \\
\hline
Background concentration &  $\mathrm{COS}$  & & & 1 & 1 & & & a supprimé: monthly \\
& ppt & 1 & & & & & & \\
\hline
\end{tabular}
\caption{\label{TPar} Controlled variables for one year. The size of the control vector is equal to 149630 for the inversion period 2008-2019. \text{\textcircled{d}} \footnotesize{The ocean flux is divided into 3 regions : 30\text{\textcircled{d}} N:90\text{\textcircled{d}} N, 90\text{\textcircled{d}} S:30\text{\textcircled{d}} S, 30\text{\textcircled{d}} S:30\text{\textcircled{d}} N.} \text{\textcircled{d}} GPP, respiration and soil fluxes of PFTs 4, 5, 6, 10, 11, 12, 13 are divided into 2 hemispheres : 0\text{\textcircled{d}} N:90\text{\textcircled{d}} N, 0\text{\textcircled{d}} S:90\text{\textcircled{d}} S.} \\
\end{caption}
\end{table}

```

```

\subsubsection{Prior and observation error covariance matrices}
\begin{figure}[h!]
\centering
\includegraphics[scale=0.4]{figs/Error_CO2.pdf}
\includegraphics[scale=0.4]{figs/Error_COS.pdf}

\caption{Assigned error standard deviations for each station and for a)  $\mathrm{CO}_2$  and b)  $\mathrm{COS}$ . Stations are ordered from the South Pole (on the left) to the North Pole (on the right).}

```

a supprimé: monthly

a supprimé: monthly

a supprimé: 0

```
\label{Error1}
\end{figure}
```

Observation errors are defined with respect to the observation operator \mathbf{H} and are actually dominated by the errors of \mathbf{H} . As explained in Section 2.3, \mathbf{H} is made of a disaggregation operator, a transport model and a sampling operator. For the transport model error statistics, we follow the detail of the approach described by [Chevallier et al. \(2010\)](#) who used the statistics of the difference between the raw time series and the corresponding smooth curve as a proxy. This approach yields one error standard deviation per station. The procedure to derive the smooth curve is explained in Section 2.6. We doubled the resulting standard deviation at each station in order to account for the error induced by the disaggregation operator. The error is likely larger at stations NWR, LEF, HFM and WIS partly because of the larger influence of nearby fluxes and we have applied an additional twofold factor there. For instance, LEF is located in the Midwestern States, a region contributing half of the summer carbon uptake in North America [\(Sweeney et al. 2015\)](#). Similarly, the standard deviation is also multiplied by two at station SMO further to the challenging representation of sub-grid-scale transport by deep convective clouds in the tropics. The resulting observation error standard deviation at each stations is shown in Figure [Error1](#).

Another justification to a larger uncertainty is that these four continental stations are influenced by the diurnal rectifier effect [\(Denning et al. 1995\)](#); such phenomenon is not represented in our observation operator since we have considered monthly fluxes. So we compared the concentrations at all surface station given by the transport of biospheric COS fluxes with hourly time variations and those given by the transport of monthly fluxes. Results shown in the Appendix that the impact of the diurnal cycle doesn't exceed the measurement accuracy (6 ppt) except at continental sites HFM et LEF. Then, we add this difference at these two stations to the observation uncertainties.

```
\begin{table}[h!]
\begin{tabular}{c|cccc}
{} & \{Oceanic fluxes\} & \{Soil fluxes\} & \{GPP and \} & \{Biomass
burning\} & \{Anthropogenic\} \\
{} & {} & {} & {} & {} & \{Respiration\} & \{emissions (COS)\} & \{
emissions\} \\
\hline
{} & \{Error standard deviation\} & \{2.0\} & \{0.3\} & \{0.1-0.2\} & \{0.9\}
& & \\
{} & \{Lag-1 autocorrelation coefficient\} & \{0.5 (60)\} & \{0.9 (100)\} &
\{0.9 (90)\} & \{0.5 (60)\} & \{0.5 (500)\} \\
{} & \{temporal correlation in days\} & {} & {} & {} & {} & {} \\
{} & {} & {} & {} & {} & {} & {} \\
{} & \{Correlations between PFTs\} & \{-\} & \{0.5-0.6\} & \{0.5-0.6\} & \{-\} & \{-\}
\end{tabular}
```

Description of the prior error covariance matrix. Since the control vector is made of low-resolution multipliers to reference maps, the standard deviations are fractions of the reference values. The lag-1 autocorrelation coefficients are the correlations assigned between

a supprimé: Tropics

```

two consecutive time steps for each controlled variable, the time step
being defined in Table \ref{TPar}. \\
% \tnote{\dag} \footnotesize{Given its large weight in terms of
GtC, we assigned an error of 0.1. for PFT 2, Tropical Broad-leaved
Evergreen Forests. }
\\
\label{Tab1}
\end{table}

```

Our prior error covariance matrix B (that applies to \mathbf{x} , a vector of ones, cf. Section 2.5.1) is described in Table \ref{Tab1}.
 The error apply to scaling factor (no units) that have a prior value equal to 1.
 The uncertainty of each parameter (or standard deviation σ) is a percentage of the value of the parameter.
 Although the large number of parameters offers very diverse possibilities for the definition of the error covariance matrix, we present only one scenario that is optimal in terms of fit to observations among those that we find compatible with our knowledge of the errors of the reference maps.

For instance:

```

\begin{itemize}
\item GPP and respiration.

```

The monthly-mean GPP from ORCHIDEE within each of the PFTs agrees with site-level GPP estimates from eddy covariance measurements in the range of 20 % (not shown). For PFT 2 (Tropical Broad-leaved Evergreen Forests), we reduce the 1-sigma uncertainty to 10%, a more realistic value given the large gross fluxes over the tropics. We introduce some non diagonal terms in the prior error covariance matrix to represent likely error correlations between PFTs given that they share for most processes the same equations in the ORCHIDEE model. Thus, the errors in the PFTs mainly located over the high latitudes (PFTs 7, 8, 9, 15), the mid-latitudes (PFTs 4, 5, 6, 10), the tropics (PFTs 2, 3, 11, 14, see Table \ref{TLRU} for a description of the PFTs) are set to be correlated with a factor 0.6 (high latitudes), 0.5 (mid-latitudes) and 0.6 (the tropics), respectively. Thus, over the high latitudes, the PFTs 7,8,9,15 are correlated with a factor of 0.6. We further introduce temporal correlations for GPP and respiration. At the first order, we expect that the errors associated to the monthly GPP simulated by ORCHIDEE are positively correlated because: i) errors in the structure of the ORCHIDEE model likely lead to positively correlated flux errors, ii) parametric errors will also provide similar correlations. However, errors in the meteorological forcing may de-correlate the gross flux errors, which could justify for an exponential decay as a function of time. The memory effect linked for example to soil moisture (and thus precipitation) may also induce error correlation \citep{stocker_drought_2019}. For the annual global GPP, this set-up leads to a 1-sigma uncertainty of 5 GtC.yr^{-1} for a reference value here of 125 GtC.yr^{-1} : this uncertainty may look small compared with the range of GPP estimates found in the literature (see Table S1) but is in agreement with the most recent estimation of $125 \pm 5.2 \text{ GtC.yr}^{-1}$ from \citep{stocker_drought_2019}. The same set-up has been chosen for plant respiration. There are error correlations between GPP and respiration but these are neglected in this study.

a supprimé: fom

a supprimé: ,

a supprimé: ,

a supprimé: Memory

a supprimé: \$

a supprimé: }\$

\item Oceanic emissions. Our resulting 1-sigma uncertainty of 350 GgS.yr^{-1} for the globe and the year, given a reference value of 271 GgS.yr^{-1} (see Fig. \ref{Tab1}), is consistent with

a supprimé: GtC

a supprimé: GtC

\cite{lennartz_direct_2017,lennartz_influence_2019,lennartz_monthly_2020} who estimated the ocean emissions between 120 - 600 GgS.yr^{-1} .
\item Anthropogenic emissions. Our correlation length of 500 days damps interannual variations, consistent with \cite{zumkehr_global_2018} who found that they do not vary by more than 5 % from one year to the next. The resulting 1-sigma uncertainty of 197 GgS.yr^{-1} for the globe and the year, given a reference value of 370 GgS.yr^{-1} (see Fig. \ref{Tab1}), is consistent with the estimation of 223-586 GgS.yr^{-1} given by \cite{zumkehr_global_2018}.

a supprimé: GtC

a supprimé: GtC

\item Soil fluxes. Our choice of a standard deviation of 30 % is rather arbitrary given the lack of measurements to evaluate the reference soil flux within each PFT. We also assign a large autocorrelation length (100 days) to damp month-to-month variations, consistent with local measurements made at Harvard and Gif-Sur-Yvette \cite{belviso_top-down_2020,commane_seasonal_2015}.

\end{itemize}

\subsection{Post-processing of the CO_2 and COS simulations and measurements}

a supprimé: \$

a supprimé: }\$

The seasonal cycle is derived from the surface data using the CCGVU curve fitting procedure developed by \cite{thoning_atmospheric_1989} (<http://www.esrl.noaa.gov/gmd/ccgg/mb1/crvfit/crvfit.html>). The procedure estimates a smooth function by fitting the time series to a first order polynomial equation for the growth rate combined with a two-harmonic function for the annual cycle, and a low-pass filter with 80 and 667 days as short-term and long-term cutoff values, respectively.

\subsection{Metrics}

a supprimé: Consistency metrics

The simulated atmospheric concentrations (for CO_2 or COS here) are evaluated against measurements using the Root Mean Square Error, RMSE, defined as:

a supprimé: %

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^N (C^{\text{Obs}}(n) - C^{\text{Mod}}(n))^2}{N}}$$

a supprimé: % \

a supprimé: %

a supprimé: →

where N is the number of considered observations, $C^{\text{Obs}}(n)$ is the n^{th} observed concentrations and $C^{\text{Mod}}(n)$ is the n^{th} modelled concentration. The unit of RMSE is in ppm (ppt) for CO_2 (COS).

a supprimé: % \

a supprimé: %

a supprimé: %modelled concentration. χ^2

The global χ^2 is equal to twice the cost function $J(\mathbf{x})$ at its minimum (see Equation \ref{CostF} for the general definition of the cost function):

\begin{align}

```

\chi^2 = (\bf{y^o} - \textbf{H} \bf{x})^T \textbf{R}^{-1} (\bf{y^o} -
\textbf{H} \bf{x}) + (\bf{x} - \bf{x^b})^T \textbf{B}^{-1} (\bf{x} -
\bf{x^b})
\label{chi2}
\end{align}

```

The variables are defined in the section 2.3. This metric allows us to check the consistency of the error covariance matrices. The χ^2 follows the so-called chi-square law, with the the number of degrees of freedom equal to the number of observations (N_{obs}) (as in our case the observation error covariance matrix is diagonal). The ratio χ^2/N_{obs} (normalized χ^2), should therefore be close to 1. This means that the residuals between observed and simulated concentrations should be aligned with the assigned measurement errors, and the residuals should be distributed as a Gaussian around the observed values. A value larger (respectively smaller) than 1 may indicate that the assigned uncertainties (of the measurements and/or from the a priori fluxes) are too small (respectively too large). However, tuning the prior and observation covariance matrices with the sole normalized χ^2 may actually be misleading since the matrices involve many variables (including off-diagonal elements) that may play compensating roles in the χ^2 \citep{chevallier_impact_2007}.

a supprimé: outside the

The χ^2 per station, χ_i^2 , represents the contribution of each site to the first term of the global χ^2 . For a station i , the metric is defined as:

```

\begin{align}
\chi_i^2 = (\bf{y_i^o} - \textbf{H} \bf{x_i})^T \textbf{R}_i^{-1}
(\bf{y_i^o} - \textbf{H} \bf{x_i})
\label{chi22}
\end{align}

```

with $\bf{x_i}^T$ and $\bf{y_i^o}$ being the simulated and observed concentrations at station i . This value, divided by N_{obs} (normalized χ_i^2), should ideally be close to 1. A value larger (respectively smaller) than 1 may indicate that the assigned uncertainties of the measurements at this station are too small (respectively too large).

a supprimé: \$\

a supprimé: y

```

\section{Inverse results}
\subsection{Comparison to the assimilated surface measurements}
\begin{table}[h]
\centering
\includegraphics[scale=0.7]{figs/RMSE.pdf}
\caption{ Column "RE" presents the fractional reduction of the model vs.
assimilated measurement RMSE
( $1 - \frac{\text{RMSE}_{\text{post}}}{\text{RMSE}_{\text{prior}}}$ ). Column " $\text{RMSE}_{\text{prior}}^{\text{seas}}$ "
presents the RMSE of the a priori detrended time series compared to the
assimilated measurement time series. Column " $\text{RMSE}_{\text{post}}^{\text{seas}}$ "
presents the RMSE of the a posteriori detrended time series. Column
" $\text{RE}^{\text{seas}}$ " presents the reduction of uncertainties using the RMSE
metrics applied to the detrended time series ( $1 -
\frac{\text{RMSE}_{\text{post}}^{\text{seas}}}{\text{RMSE}_{\text{prior}}^{\text{seas}}}$ ). Column " $\chi^2$ "
presents the reduced chi-squared statistics (without unit) for each
station. The detrended curves have been filtered to remove the synoptic
variability (see Sect. 2.6). The RMSE is in ppm (ppt) for  $\text{CO}_2$ 
( $\text{CO}$ ). All statistics are for the period 2009–2019.}
\label{Error1}
\end{table}

```

Table \ref{Error1} shows the error reduction achieved by the inversion in terms of RMSE between the simulated and the observed concentrations. As expected, the inverse system has reduced the observation-model mismatch by about 85 \% at most stations. Of interest in Table \ref{Error1} is also the error reduction for the detrended smooth curves in which only seasonal variations are retained. It is indeed important to accurately represent the large COS and CO_2 surface depletion in spring as it mainly reflects the amplitude of the GPP over the continents. The seasonal error reduction is usually smaller than the overall error reduction: the COS inversion mainly corrects the negative tendency in COS mixing ratio arisen from the unbalanced prior budget. For instance at MLO between 2008 and 2011, the tendency of the CO_2 (COS) concentrations a priori is $3.9 \mathrm{ppm.yr}^{-1}$ ($-57 \mathrm{ppt.yr}^{-1}$) against $2.0 \mathrm{ppm.yr}^{-1}$ ($1.4 \mathrm{ppt.yr}^{-1}$) in the observations. Yet, the inversion has reduced the seasonal misfits to observations at most sites except at LEF and MLO for CO_2 and MLO, THD, WIS for COS . At the northernmost sites (ALT, BRW, SUM, MHD), the error reduction exceeds 50\% for both compounds. Despite some improvements, the inversion still struggles to represent the seasonal cycle of the COS measurements at sites WIS, HFM, THD for which the RMSE remains greater than 15 ppt. THD is a coastal station which suffers from the influence of fluxes nearby \citep{riley_influence_2005}. For this reason, modelling the variability of its CO_2 and COS mixing ratio has been shown to be particularly challenging \citep{ma_inverse_2021}. The inverse system also struggles to match CO_2 measurements at sites WIS, NWR, LEF with a seasonal RMSE greater than 1.5 ppm.

a supprimé: raw

The consistency of the estimate with the measurement errors and the a priori flux errors assumed is analyzed first with the global normalized chi-squared statistic (see Equation \ref{chi2}). This metric should ideally be close to 1. In our case, the normalized χ^2 equals to 1.04, a value consistent with a fair configuration. The relative contribution of the measurement term (first term of Equation \ref{chi2}) to the total χ^2 (Equation \ref{chi2} or cost function at its minimum) is much larger than that of the flux term (80\% versus 20\% on average), suggesting that the a priori constraint is rather loose.

a supprimé: Section

a supprimé: CostF

In addition to the global consistency between data errors and a priori flux errors, the validity of the relative weights (inverse of the squared data error) assumed for the individual measurement residuals (i.e., at each station) is assessed (see Equation \ref{chi2}). To this end, Table \ref{Error1} shows the χ^2 per station. The value is less than 1 for seven stations out of 15 for both compounds, meaning that the residuals are within the range of the assigned observation uncertainty. Among the stations with χ^2 values greater than 1, HFM stands out and likely we assigned too small uncertainties to this station.

a supprimé: Section

a supprimé: have

a supprimé: low

In order to better visualize the improvement on the seasonal cycle, we compare in Figure \ref{Error2} the simulated a priori and a posteriori concentrations against observations at three sites: BRW, NWR and LEF. These time series have been detrended beforehand to retain the seasonal cycle. At BRW, the inversion has corrected the too low seasonal amplitude and the phase lag in the a priori concentrations within the range of observation uncertainties. At LEF, the a priori concentrations were already in good agreement with the observations and the inversion has not

improved the simulated concentrations much. However, at NWR, the inversion struggles to correct the advanced phase, especially in the CO_2 simulations, consistent with a χ^2 greater than 1. One likely explanation is that our biome-scaling approach with one coefficient per PFT is too coarse to correct the spatial distribution of the prior fluxes, especially between relatively close sites such as NWR and LEF. The latter are more prone to be influenced by local fluxes than ocean stations such as MHD for example.

```
\begin{figure}[h!]  
\centering  
\includegraphics[scale=0.5]{figs/FitStatCOS.pdf}  
\includegraphics[scale=0.5]{figs/FitStatCO2.pdf}  
  
\caption{Detrended temporal evolutions of simulated and observed  
 $\text{CO}_2$  and  $\text{COS}$  concentrations at three selected  
sites, for the a priori and a posteriori fluxes, simulated between 2009  
and 2011. Top: Barrow station (BRW, Alaska, USA), middle: Niwot Ridge  
(NWR, USA) bottom: Park Falls (LEF, USA). The curves have been  
detrended beforehand and filtered to remove the synoptic variability (see  
Sect. 2.6). The grey bar represents the 1-sigma error bar of the  
observations.}  
\label{Error2}  
\end{figure}
```

a supprimé:) . }

```
\subsection{Optimized fluxes}
```

Table \ref{Opt1} summarizes our top-down assessment of the COS and the CO_2 budgets. The inversion doubled the COS oceanic emissions to 530 GgS.yr^{-1} . Given the missing source in the reference fluxes, the ocean dominance in the measurement footprints, and the efficient reduction of the global error by 90%, the increase of oceanic emissions is an expected behaviour of the Bayesian inverse system. In contrast, the inversion marginally decreased the total soil and vegetation absorption likely due to the seasonal constraints. Following a decrease of 7 GtC.yr^{-1} of the GPP to match the COS constraint, the respiration has decreased by 10 GtC.yr^{-1} in order to keep a land carbon sink in agreement with the global atmospheric CO_2 budget. Thus, on a global scale, the inversion seems to have corrected the overestimated prior atmospheric trend by a larger decrease in respiration than in GPP. All residuals between the total prior and the posterior fluxes are within the assumed 1-sigma range of the prior uncertainty, except for respiration, where the increment is twice as large as the standard deviation. The residuals are even much smaller than the prior standard deviation for the anthropogenic and the biomass burning emissions, suggesting that we could have narrowed the initial errors for those components.

The total oceanic COS emission remains lower than previous top-down studies using different configurations and observations, which instead estimated an oceanic source between 700 and 1000 GgS.yr^{-1} (\citep{berry_coupled_2013,kuai_estimate_2015,launois_new_2015}). Several reasons could explain these differences. Firstly, the \cite{zumkehr global 2018} anthropogenic emissions are much higher than the \cite{kettle global 2002} one used in these previous studies.

a supprimé: First

Secondly, we assimilated continental surface measurements from the NOAA network through the whole years of 2008-2019 while \cite{kuai_estimate_2015} assimilated a single month of satellite retrievals over the tropical oceans. Finally, the prior biospheric and oceanic fluxes used, especially over the tropical domain, a region that is poorly constrained in the inversion, could explain the differences with the previous COS budgets. \cite{launois_new_2015} noticed a dependence between the magnitude of the optimized ocean source and the prior vegetation uptake. The larger biospheric sink used in \cite{launois_new_2015, berry_coupled_2013} requires a larger oceanic source over the tropics to close the COS budget. This is particularly true for \cite{berry_coupled_2013} who used a fixed large biospheric sink of 1100 GgS.yr^{-1} .

a supprimé: Second, the \cite{zumkehr_global_2018} anthropogenic emissions are much higher than the \cite{kettle_global_2002} one used in these previous studies.

a supprimé: which

```
\begin{table}[h!]
\includegraphics[scale=0.6]{figs/Budget.pdf}

%\begin{tabular}{l|cc}
%\hline
% Component of the & Prior (GgS/y) & Post (GgS/y) \\
% COS budget & & \\

%\hline
% Anthropogenic emissions & 398.3 \pm 199.1 & 360.5 \pm 21.9 \\
% Oceanic fluxes & -263.3 \pm 1077.2 & 530.1 \pm 186.9 \\

% Biomass burning emissions & 43.3 \pm 32 & 72.4 \pm 27.5 \\
% Soil fluxes & -236.1 \pm 70.8 & -226 \pm 69.5 \\
% Vegetation sink & -656.8. \pm 131.4 & -646.4 \pm 114.6 \\
% OH sink & -100 \pm 0 & -100 \pm 0 \\
%\hline
%\end{tabular}

%\begin{tabular}{l|cc}
%\hline
% Component of the & Prior (GtC/y) & Post (GtC/y) \\
% CO2 budget & & \\

%\hline
% GPP & -126.7 \pm 25.3 & -124.9 \pm 22.2 \\
% Respiration & 124.6 \pm 24.9 & 120.1 \pm 22.3 \\
% Fire (unoptimized) & 1.9 & 1.9 \\
% Fossil fuel (unoptimized) & 9.6 & 9.6 \\
% Oceanic fluxes (unoptimized) & -1.9 & -1.9 \\

%\hline
%\end{tabular}

\caption{Prior and posterior total fluxes and their associated 1-sigma uncertainty as part of the  $\text{COS}$  (left) and the  $\text{CO}_2$  (right) budgets. The mean magnitude of the different types of fluxes is given for the period 2009-2019. The vegetation sink is computed from the vegetation uptake (Table on the right) using the LRU relationship described in Equation \ref{eq1}. The components of the  $\text{CO}_2$  and  $\text{COS}$  budgets, as written here, have been obtained by adding all the related optimized parameters (see Table \ref{TPar} for a
```

description of the parameters). The flux convention is positive upwards (from the surface to the atmosphere). For a given component, the associated uncertainty is the root-mean square of the sum of all the posterior error covariance terms related to the component divided by the number of years (11 here.)

```
\label{Opt1}
\end{table}
```

Figure \ref{Opt2} represents the zonal mean of the prior and posterior oceanic and continental COS fluxes as a function of latitude. The inversion increased ocean fluxes over the tropics while decreasing them in the high latitudes. This behaviour was already noticed by \cite{berry_coupled_2013} who used a different inverse system and the \cite{kettle_global_2002} oceanic flux as a prior. Over the tropics, COS and CS_2 measurements in sea waters do not support this increase as already mentioned in the introduction \cite{lennartz_direct_2017,lennartz_monthly_2020}. However, COS emissions through DMS oxidation in pristine marine environment, could play a role in sustaining this tropical source. Over the northern and southern oceans, high emissions in our reference oceanic flux from \cite{lennartz_direct_2017} mainly arise from the direct oceanic emissions (see Fig. \ref{FF2}). The latter could be overestimated: the COS concentrations simulated by the ocean box model are higher than most of the measurements made in sea waters sampled over different parts of the globe \cite{lennartz_direct_2017}. This remark supports the inversion decrease of the oceanic emissions over the mid and high latitudes. The decrease beyond 50° towards the poles also reflects a seasonal cycle in COS sea water concentrations of a much lower amplitude than the one in atmospheric COS in the marine boundary layer \cite{lennartz_marine_2020}. This strong marine seasonal cycle is not attenuated enough by mixing processes within the boundary layer and the inversion weakened the oceanic release to match the seasonal cycle in atmospheric COS concentrations at BRW and ALT. In particular, the emissions in the northern high latitudes have been suppressed in summer to correct the late peak in the time series at BRW on Figure \ref{Error2}. While oceanic emissions decrease in the high latitudes, the terrestrial sink tends to increase. The change in terrestrial sink is mainly attributed to vegetation (see Fig. [S4](#)). The change in soil fluxes goes in the same direction as the change in COS vegetation uptake.

a supprimé: oxydation

a supprimé: S3

a supprimé: than

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.6]{figs/Gradient_flux.pdf}
\caption{Latitudinal distribution of the prior (dashed line) and
posterior fluxes (full line) for the continental (brown) and oceanic
components (blue) of the  $\mathrm{COS}$  budget. The fluxes have been
averaged over the years 2009–2019. }
\label{Opt2}
\end{figure}
```

a supprimé: for

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.6]{figs/Gradient_CO2.pdf}
\caption{Latitudinal distribution of the prior (dashed line) and
posterior net  $\mathrm{CO}_2$  fluxes from the terrestrial vegetation
```

(full line). Vegetation fluxes are the sum of GPP and respiration fluxes. The fluxes have been averaged for the years 2009-2019. }

```
\label{Opt2bis}  
\end{figure}
```

Regarding the impact on the CO_2 budget, Figure \ref{Opt2bis} shows the latitudinal distribution of the net CO_2 vegetation fluxes defined as the sum of respiration and GPP before and after inversion. The inversion has increased almost threefold the net vegetation absorption above 50°N . This response is a common feature of the current inverse systems which, by assimilating CO_2 measurements only into an atmospheric transport model, infer a higher net vegetation sink in the high latitudes than land-surface models. Indeed, in Fig. 8 of \cite{friedlingstein_global_2020}, the net land sink (above 30°) calculated as the average of 17 process models is between 0.5 GtC/y and 1.5 GtC/y whereas the flux calculated from 6 inverse systems is between 1 and 2.5 GtC/y averaged over the last ten years. More specifically, Figure \ref{Opt3} illustrates how the inversion changes the seasonal cycle of GPP and respiration within each of the 15 PFTs of the ORCHIDEE model. The changes in the global total per PFT are shown separately in the supplementary material (see Fig. S4). In the tropics within PFTs 2 and 3 (Tropical Broad-leaved Evergreen and Raingreen Forests, see Table \ref{TLRU}), the inversion decreased GPP by about $4 \mathrm{GtC.yr}^{-1}$ whereas respiration lost $1 \mathrm{GtC.yr}^{-1}$, leading to a small source of CO_2 . In the mid-latitudes (PFTs 4, 5 and 10, Table \ref{TLRU}), the inversion weakened GPP and respiration by $5 \mathrm{GtC.yr}^{-1}$ and $2 \mathrm{GtC.yr}^{-1}$, respectively. The second salient change is an increase in CO_2 absorption within the high latitudes covered by PFTs 7, 8, 9 and 10 (see Table \ref{TLRU}). Indeed, GPP increased by almost $2 \mathrm{GtC.yr}^{-1}$ while respiration only decreased by $0.2 \mathrm{GtC.yr}^{-1}$ in total. The increased GPP over the boreal latitudes explains the larger seasonal cycle of the a posteriori CO_2 and CO_2 concentrations at sites BRW and ALT. The comparison of GPP and respiration from ORCHIDEE against eddy covariance measurements at several sites around the globe pointed at an underestimation of these components, consistent with our inversion results (not shown). A complete validation of this ORCHIDEE version will be the topic of a future publication.

```
\subsection{Comparison with independent observations}
```

```
\subsubsection{Evaluating the seasonal cycle with SIF data}
```

In order to assess the realism of the a posteriori GPP, its seasonal cycle is compared with seasonal cycle of the GOME-2 SIF product. Although the ecosystem-dependent bias in the SIF products makes a direct comparison with GPP impossible, SIF has been recognized as a good indicator of the temporal dynamics in GPP. At the ecosystem scale, SIF is anti-correlated with the GPP: a maximum in SIF corresponds with a minimum in GPP. Figure \ref{Opt3} superimposed the maximum of the SIF on the GPP seasonal cycle. The normalized SIF seasonal cycle is further shown on Fig. S6. Ideally, the maximum coincides with the minimum of the GPP seasonal cycle. Overall, the inversion has not altered the timing of the

a supprimé: given by

a supprimé: one given by

a supprimé: Supplementary

a supprimé: S3

a supprimé: }\$

a supprimé: higher

a supprimé: the one

a supprimé: SIF from the

a supprimé: dependant

a supprimé: GPP

a supprimé: dynamic.

a supprimé: minimum

a supprimé: S5

a supprimé: minimum

CO_2 seasonal depletion. The optimized seasonal cycle disagrees with the SIF satellite retrievals within PFT 2 (Tropical Broadleaved Evergreen), PFT 3 (Tropical Boreal Raingreen Forest) and PFT 14 (Tropical C3 grass), questioning the realism of a weaker CO_2 and CO_2 absorption over the tropics. Within PFT 2, the inversion tends to produce a seasonal signal in opposite phase with SIF. In the mid-latitudes, the seasonal phase of GPP is slightly degraded within PFT 4 (Temperate Needle-leaved Evergreen Forest) while it is improved within PFT 12 (C3 Agricultural Land). In the high latitudes, the phase of the seasonal cycle, which was in quite good agreement with the SIF in the GPP a priori, has not been altered by the inversion. To conclude, the atmospheric inversion does not lead to a clear improvement in the representation of the GPP seasonal cycle.

a supprimé: is degraded

a supprimé: the

a supprimé: the

a supprimé: opposition of

a supprimé: the

a supprimé: the

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.64]{figs/Cycle_GPP2.pdf}
\includegraphics[scale=0.64]{figs/Cycle_Resp.pdf}

%\begin{table}[h!]
\resizebox{\textwidth}{!}{%
\begin{tabular}{l|cccccccccccc}
& TrBrE & TrBrR & TeNeE & TeBrE & TeBrS & BoNeE & BoBrS & BoNeS &
TeC3g & C4g & TrC3Ag & C4Ag& TrC3g & BoC3g\\
\hline
Prior & -0.3& 0.5& 1.0& 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.9 & 0.9 & 1.0 & 0.9
& 0.8 & 1.0\\
Posterior& -0.8& 0.1& 0.9 & \textbf{1.0} & \textbf{1.0} & \textbf{1.0} & \textbf{1.0} & \textbf{1.0} & \textbf{0.9} & \textbf{0.9} & \textbf{1.0} & \textbf{0.9}
& \textbf{1.0} & 0.6 & \textbf{1.0}\\
\end{tabular}}
%\end{table}
```

\caption{Mean seasonal cycle of the total prior (black) and posterior (orange) GPP (a) and respiration (b) fluxes and their uncertainties within each of the 15 PFTs during the period 2009-2018. The maximum of the mean seasonal cycle of the SIF from GOME-2 has been superimposed on the GPP seasonal cycle in green. The fluxes have been averaged over 2009-2018. Below are the correlation coefficient between the monthly SIF and the GPP averaged during the period 2009-2018. The values in bold indicate the PFTs with a GPP improved or unchanged by the inversion. PFT 1, bare soil, is not shown as respiration and GPP are null. Only the values integrated over the Northern Hemisphere are shown for PFTs 4, 5, 6, 10, 11, 12 and 13. The identifiers of the PFTs are described in Table \ref{TLRU}. The acronyms Tr, Bo and Te mean Tropical, Boreal and Temperate, respectively. }

a supprimé: minimum

a supprimé: between

a supprimé: the

\subsubsection{Comparison with independent atmospheric observations}

As a second step, we assess the a posteriori concentrations using several datasets : the MIPAS satellite retrievals, the HIPPO airborne measurements and the surface measurements over Japan and France (see section 2.2). In particular, the MIPAS retrievals of CO_2

a mis en forme : Français

atmospheric concentrations at 250 hPa in the tropics give insight into the magnitude of the main biospheric sink located over Brazil during the wet season, when convective air masses reach the upper troposphere \citep{glatthor_global_2017}. First, Figure \ref{Opt4} shows the a posteriori and a priori COS seasonal concentrations at 250 hPa, convolved with the MIPAS averaging kernels and averaged over the period 2009-2012. We see that the inversion reduced the RMSE by more than one third throughout the whole year. The inversion removed the positive bias above 50° N in DJF and under 50° N in MAM (as a result of lower oceanic emissions in the high latitudes) and the negative bias over the tropical oceans (as a result of higher tropical oceanic emissions). Such an increase is consistent with \citet{glatthor_tropical_2015}, who also needed to multiply the vegetation sink and the oceanic sources from \citet{kettle_global_2002} by 4 to better match the MIPAS retrievals. However, there are some remaining deficiencies. In particular, the COS depletion observed between Brazil and Africa is well reproduced but its amplitude is slightly underestimated. The simulated COS concentrations are also too small over the Pacific Ocean. The reasons could be an underestimation of the tropical emissions or a too homogeneous distribution of these emissions through the longitudes. We have to remember that we have optimized a single factor for the oceanic emissions over the whole tropical band and thus the spatial gradients within the tropical band have not been optimized. This could explain the lack of variability over the ocean. Over the mid-latitudes, the smaller concentrations in spring point at a too weak terrestrial sink or too strong oceanic emissions. The lack of stratospheric COS loss could also be responsible for these underestimated concentrations since they are close to the tropopause near 60° .

a supprimé: weak in

a mis en forme : Français

a supprimé: weaker

a mis en forme : Français

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.4]{figs/MIPAS_obs.pdf}
\includegraphics[scale=0.4]{figs/MIPAS_prior.pdf}
\includegraphics[scale=0.4]{figs/MIPAS_post.pdf}
```

\caption{Climatological seasonal COS distributions at 250 hPa measured by (left column) MIPAS and simulated using the prior scenario (middle column) and (right column) the optimized scenario. The data sets cover the years 2008,–2012, and the displayed seasons are (top row) December to February, (second row) March to May, (third row) June to August, and (fourth row) September to November. White areas are data gaps, and dark blue COS amounts above the Amazonian region bottom left) are below 450 pptv. The negative bias in the prior concentrations, which results from the unbalanced COS prior budget, has been removed on panel c. The RMSE (see Equation \ref{ERMSE}) is shown above each panel. The bias in the prior concentrations has been removed before computing the RMSE.

a mis en forme : Français

a supprimé: .[FC: a justifier, pas clair]

a mis en forme : Français

```
\label{Opt4}
\end{figure}
```

We further assess the latitudinal distribution of the COS sources and sinks given by the inversion with the help of the HIPPO airborne measurements. For this purpose, Figure \ref{Opt5} compares the inter-hemispheric gradient in the a posteriori and a priori COS and CO_2 concentrations against the HIPPO airborne measurements. We have verified beforehand that the transport model performs well at sites LEF and THD (see Fig. [S7](#)) whose continental

a supprimé: S6

a mis en forme : Français

and coastal locations respectively emphasize transport errors. The representation of vertical mixing is indeed crucial for continental sites \citep{geels_comparing_2007} such as LEF whereas coastal sites such as THD are difficult to represent in coarse resolution models \citep{riley_influence_2005}. Given the good agreement between modelled and observed vertical profiles at these two sites (see Fig. [S7](#)), transport errors are assumed here to be of secondary importance compared to the uncertainty in the fluxes and differences between the concentrations a priori and a posteriori are ascribed to differences in the surface fluxes. Figure \ref{Opt5} shows that the a posteriori better matches the observed latitudinal distribution. Especially, the shared positive bias in the northern latitudes between COS and CO_2 has been corrected as a result of higher GPP. The improvement is also noticeable in the COS and CO_2 vertical profiles over Northern America (see Supplementary material). In contrast to the \cite{ma_inverse_2021} top-down study, there is no significant negative bias in the COS vertical profiles here (see Fig. [S7-11](#)).

- a supprimé: profile
- a mis en forme : Français
- a supprimé: S6
- a mis en forme : Français

- a supprimé: S6-10
- a mis en forme : Français

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.5]{figs/HIPPO.pdf}
\caption{Comparison of the latitudinal variations of the a priori and a posteriori LMDz  $\mathrm{COS}$  abundance with the HIPPO observations. Because of the unbalanced prior, the LMDz  $\mathrm{COS}$  abundances have been vertically shifted such that the means of the a priori are the same as the mean of the HIPPO data (521 ppt). The error bar is calculated as the standard variation of the  $\mathrm{COS}$  concentration averaged over longitudes and heights.}
\label{Opt5}
\end{figure}
```

- a supprimé: The
- a mis en forme : Français

- a supprimé: averaged
- a mis en forme : Français
- a mis en forme : Français

The optimized COS fluxes are now assessed at three surface sites in Japan : Miyakojima (MIY; 24°N , 125°E), Yokohama (YOK; 35°N , 139°E), and Otaru (OTA; 43°N , 141°E). In winter, these sites sample air masses coming from the eastern northern edge of China (see \cite{hattori_constraining_2020} and the LMDz footprints on Fig. S11). The confrontation of the posterior concentrations against measurements serves at evaluating the spatial distribution of the \cite{zumkehr_global_2018} anthropogenic inventory over the eastern part of China. In Figure \ref{Opt6} a and b, we show a comparison between the a posteriori and observed COS concentrations at each of the three sites for both winter and summer 2019. The averaged COS surface concentrations during February-March 2019 and July-August 2019 are also shown in Figure \ref{Opt6}. At the northernmost site OTA, the overestimation of the COS mixing ratios of 40 ppt points at too strong anthropogenic sources in Northern China in the modified \cite{zumkehr_global_2018} inventory. The site located in middle Japan, YOK, has a simulated concentration of almost 100 ppt higher than the observations. This implies an error in the inventory which indicates a source above the site (see Figure S9). As for the southern site MIY, the model underestimates the COS concentration by 100 ppt, pointing at an underestimation of the anthropogenic sources over the eastern edge of China or Korea.

- a supprimé: concentrations
- a mis en forme : Français

- a supprimé: one observed
- a mis en forme : Français

In summer, sites YOK and OTA sample air masses coming both from continental Japan and from the Pacific Ocean at the East of Japan. The southernmost site MIY seems to be mostly affected by oceanic sources originating from the east (see the LMDz footprints on Figure S12). The sites OTA and YOK overestimate the COS concentrations by 60 and 150 ppt and reflect the influence of the misplaced anthropogenic source in center Japan (Figure S13). At MIY, the comparison with observations suggests that the oceanic source is too strong because the atmospheric concentrations are overestimated by 40 ppt in southern Asia and in northern Japan. However, the oceanic source may not be overestimated in Southern Asia because we have assumed that CS_2 is emitted as COS . `\citet{ma_inverse_2021}` showed that implementing the CS_2 oxidation process into the atmosphere leads to a decrease in surface COS mixing ratio of 40 ppt in the vicinity of Japan. Also, there is an oceanic hot spot located in the footprint of the site (see Figure S13) which might not be reliable.

The spatial pattern of the `\citet{zumkehr_global_2018}` inventory seems to show too strong sources over Japan and too weak sources in the eastern edge of China. The inversion system could therefore have compensated the lack of anthropogenic source in the eastern part of China by increasing the oceanic source. However, it is difficult to extrapolate conclusions drawn from a specific region to a larger scale. There is also no clear indication that the oceanic sources are overestimated eastward of Japan.

Finally, we perform a similar assessment of the optimized COS fluxes in winter at station GIF in France. The footprint of the station covers central France and countries at the eastern edge such as Belgium and the eastern part of Switzerland (see Figure S14). The confrontation of the posterior concentrations against measurements serves at evaluating the `\citet{zumkehr_global_2018}` anthropogenic inventory and, in particular, its spatial distribution over central France since the terrestrial sink is assumed to be much smaller. Station MHD provides very small constraints over France and Eastern Europe as its footprint is mainly oceanic. The comparison between the posterior concentrations and atmospheric measurements on Figure `\ref{Opt6}`c indicates that the anthropogenic sources within the footprint of the station are also overestimated: the a posteriori concentrations are more than 130 ppt higher than the one observed. This confirms the study of `\citet{belviso_top-down_2020}` which reported a misplaced hot-spot on Paris (see Fig. S14). In reality, the concentrations at GIF are 10 ppt lower than the one at the background MHD, reflecting a dominant influence of the biospheric sink, in this season.

```
\begin{figure}[h!]
\centering
\includegraphics[scale=0.55]{figs/Japan-WIN.pdf}
\includegraphics[scale=0.55]{figs/Japan-SUM.pdf}
\includegraphics[scale=0.55]{figs/GIF-Win.pdf}
\caption{Mean  $\text{COS}$  concentration sampled at the first level of the LMDz model in a) winter 2019 (February-March), b) in summer 2019 (July-August) and c) in winter (Dec-Feb) during the period 2016-2019. The values within the yellow frames correspond to the mean  $\text{COS}$  observed and modelled concentrations, and their standard deviation at four surface sites: Miyakojima (24°N,
```

a supprimé: S8

a mis en forme : Français

a supprimé: S10

a mis en forme : Français

a supprimé: concentrations

a mis en forme : Français

a supprimé:

a mis en forme : Français

a supprimé: S9

a mis en forme : Français

a supprimé: S13

a mis en forme : Français

a supprimé: and that the oxidation rate is low during this period of the year.

a mis en forme : Français

a supprimé: low

a mis en forme : Français

a supprimé: .

a mis en forme : Français

```

125\degree27E), Yokohama (35\degree51N, 139\degree48E), and Otaru
(43\degree14N, 141\degree16E) and GIF (48\degree42, Å6N - 2\degree08, Å6E).
Station MHD has been assimilated and is shown here as a reference.)
\label{Opt6}
\end{figure}

```

```
\section{Discussion and perspectives}
```

To conclude, there is an need of continuous in-situ carbonyl sulfide observations. The lack of continuous in-situ observations, especially over the tropics, limits our capacity to infer the COS surface fluxes by inverse modelling and therefore to optimize GPP. There is some hope that new satellite products could address this issue but at this stage, current COS retrievals have also their limitations such as, for instance, cloud interference or the lack of sensitivity to the surface fluxes \citep{glatthor_global_2017,kuai_estimate_2015-1,vincent_fast_2017}. Letting aside this obvious lack of observations to be assimilated, we are now discussing the way forward to improve our knowledge of the COS budget.

```
\begin{itemize}
\item \textbf{Improving the anthropogenic inventory} The inverse system
has weakened the global anthropogenic source by almost 20 \%. It is
unclear whether this decrease results from an overestimation of the
global emissions or from misplaced hot-spots within the footprints of the
assimilated stations. For instance, the overestimated concentrations in
the model at a site located in middle Japan point to a misplaced hot-spot
near the station. If these measurements were assimilated, the inverse
system would tend to produce an unrealistic negative flux increment over
the area to match the observed concentrations. A similar inconsistency
has been reported between measurements at the Gif-sur-Yvette background
site and the hot-spot to the north, over Paris, stated in the
\cite{zumkehr_global_2018} inventory. Thus, the reported hot spot
locations and magnitudes must be improved to be able to benefit from
these new observations at Gif-Sur-Yvette and in Japan. Further work
should includes a more thorough evaluation of European anthropogenic
sources using  $\text{COS}$  retrievals from Fourier transform infrared
spectrometry \citep{wang_towards_2016,kryzstofiak_carbonyl_2015} along
with a high-resolution (e.g., 0.5 \degree) chemical transport model in
order to correct the spatial distribution of these emissions. Samples in
industrial facilities could also be made to validate anthropogenic
inventories. Currently, due to large uncertainties in the emission
factors and the use of a proxy for spatial disaggregation, the
anthropogenic inventory is more appropriate for interpreting atmospheric
COS measurements from background sites like MLO than atmospheric COS
measurements which have a significant influence from nearby emissions
(e.g. Japan/YOK).
```

```
\item \textbf{Improving the relationship between  $\text{COS}$  plant
uptake and GPP.} For the LRU values, we have only made a distinction
between C4 and C3 plants. A complementary experiment would be to optimize
a set of LRU coefficient for each PFT together with the GPP fluxes. We
plan to include the PFT dependence of the LRU by using the LRU dataset of
\cite{maignan_carbonyl_2020} derived from a mechanistic vegetation
model, and for which conductances will be further tuned with eddy-
covariance flux measurements. LRU absolute values are indeed critical.
For instance, if the LRU were larger at high latitude, the inversion
```

a mis en forme : Français

a supprimé: weaken

a mis en forme : Français

a supprimé: at

a mis en forme : Français

a supprimé: in the vicinity of

a mis en forme : Français

a supprimé: \cite{belviso_top-down_2020}.

a mis en forme : Français

a supprimé: More generally, the disagreement between simulated and observed COS concentrations tested at these sites indicated that there is a large uncertainty in this inventory. Further work first

a mis en forme : Français

a supprimé: IASI

\cite{serio_simultaneous_2020} and from

a mis en forme : Français

a supprimé: chemical transport model in order to correct their spatial distribution...

a mis en forme : Français

a supprimé: }

a mis en forme : Français

a mis en forme : Français

would not need to increase the GPP as much. However, LRUs have been estimated to be lower in the boreal ecosystems (around 1 and 1.8 for \cite{maignan_carbonyl_2020} and \cite{seibt_kinetic_2010} respectively) than in the tropical and temperate ecosystems (around 1.3 and 3 for \cite{maignan_carbonyl_2020} and \cite{seibt_kinetic_2010} respectively). So, using another existing LRU dataset will likely lead to a comparable GPP sink in the high latitudes. Another simplification of our study is that the current LRU approach does not take into account the feedback between COS vegetation sink and atmospheric concentrations. The atmospheric concentrations vary on seasonal and interannual timescales but have been indeed considered constant per latitudinal band. Such a feedback might have significant impact on the total vegetation sink (see Fig. S18), in particular over the Amazon. Thus, refining our inverse system would require including the feedback between the atmospheric concentrations and the COS vegetation sink (first order approach). This will involve representing the sharp drop of COS between the canopy and the boundary layer, which can reach 70 ppt in redwood forests \cite{campbell_atmospheric_2015}. However, current global models do not represent the turbulence within the canopy and the link with the atmospheric boundary layer, which does not allow to correctly simulate the vertical gradient of concentrations between the lowest layer of an atmospheric model and the canopy. Some promising developments were made with the ORCHIDEE LSM \cite{naudts_vertically_2015} but more research is needed before they can be used for our application.

\item \textbf{Increasing the realism of the soil fluxes} The GPP estimate strongly relies on the realism of the soil fluxes. The soil fluxes need to be more constrained and their errors better defined. In particular, more attention should be paid to the seasonality of soil fluxes compared to the one of the vegetation fluxes in the field measurements. For instance, this would help to know whether the two months-lag between the soil and vegetation fluxes in the high latitudes is realistic.

\item \textbf{Improving the prior COS oceanic fluxes with the help of an ocean model.} Prior oceanic emissions are probably

overestimated in the high and mid latitudes as shown by \cite{lennartz_direct_2017,lennartz_monthly_2020} and as suggested by the inverse system. A possible reason could be the lack of horizontal transport and downward mixing within the water column in their ocean box model. Another estimation of the oceanic emissions based on an ocean general circulation and biogeochemistry model which fully represents the transport of water masses would help to better define the range of uncertainties of these emissions. In particular, the DMS emissions simulated by the NEMO PISCES ocean model \cite{belviso_dms_2012} are higher over most of oligotrophic subtropical zones compared to the DMS emissions of \cite{lana_updated_2011}. This means that the spatial distribution of the DMS oceanic emissions is highly uncertain. In the past, direct oceanic emissions have been simulated in the Nucleus for European Modelling of the Ocean (NEMO) Pelagic Interaction Scheme for Carbon and Ecosystem Studies (PISCES) ocean model by

\cite{launois_new_2015-1} but the hydrolysis rate has later been found erroneous. Future work should include the correction of the rate in the NEMO-PISCES ocean model and also the implementation the oceanic physical processes responsible for the CS_2 emissions. Moreover, the single factor used to convert DMS into COS is very uncertain and may not apply to any atmospheric conditions (Von Hobe, 2020, personal communication). Since there is so much DMS emitted by the ocean (ca. 28 Teragrams per year), a small change in the conversion factor (e.g. from the current 0.7\% to e.g. 2.5\%) could make a large

a supprimé: for example the discussion in \cite{ma_inverse_2021}),

a mis en forme : Français

a mis en forme : Français

a mis en forme : Français

a supprimé: }

a mis en forme : Français

a supprimé: for DMS , it has been shown that...

a mis en forme : Français

a supprimé: use of

a mis en forme : Français

a supprimé: that accounts for nutrient dynamics, has resulted in...

a mis en forme : Français

a supprimé: emissions

a mis en forme : Français

a supprimé: } \cite{belviso_dms_2012}.

a mis en forme : Français

a mis en forme : Français

a supprimé: includes

a mis en forme : Français

a supprimé: Teragrams

a mis en forme : Français

difference. When the relative contribution of indirect COS sources to total ocean emissions is better known, an extension of this work could be to optimize each oceanic process separately.

a mis en forme : Français

\item \textbf{Implementing a complete chemistry of COS into the LMDz atmospheric transport model} For an economy of computation time, we have assumed that the DMS and CS_2 oxidation into COS happens instantly in the atmosphere. However, \cite{ma_inverse_2021} showed that such simplifications could modify the average COS surface concentrations up to 80 ppt over Eastern China and Japan in winter. These chemical reactions need be implemented in the LMDz atmospheric model in order to properly evaluate the \cite{zumkehr_global_2018} inventory with the help of COS atmospheric measurements. The lifetime of the DMS , CS_2 and, to a lesser extent, COS into the atmosphere depends on the realism of the OH fields. Therefore, the impact of their uncertainty on the inverse results needs also to be quantified. Chemical transport models disagree on the spatial distribution of the OH fields and using other OH fields could significantly alter the COS budget as it was demonstrated for the methane budget

a supprimé: outflow of the anthropogenic sources...

a mis en forme : Français

\cite{zhao_influences_2020,zhao_role_2020}. In addition, we plan to introduce the stratospheric chemistry of COS into the LMDz atmospheric transport model. The implementation of a complete chemistry while keeping a multi-year inversion window requires using a variational approach: the chemical reactions are indeed more difficult to implement in an analytical inverse system using pre-computed Jacobian matrices.

\item \textbf{Including potentially important missing sources} For instance, the biomass burning sources from biofuels are not included in the \cite{stinecipher_biomass_2019} inventory although they were previously estimated to be three times as large as the sources from open burning \cite{campbell_atmospheric_2015}. We have also neglected the contribution of volcanic emissions on the COS budget. As the locations and time of the volcanic eruptions are well known, emissions based on existing measurements

a supprimé: we have

a mis en forme : Français

\cite{belviso_estimate_1986,chiadini_carbonyl_1991,symonds_origin_1992,sawyer_investigation_2008,notsu_chemical_2010} could be mapped and given as an input to the atmospheric model. Likewise, DMS emissions from vegetation, tropical forests, soil and wetlands \cite{yi_exchange_2008,kanda_emission_1992,minami_emission_1993} have not been included although their contribution to the total DMS release have been estimated in the past between 2 and 15 \% \cite{watts_mass_2000,gondwe_contribution_2003}. Moreover, we have neglected the COS emissions from the anoxic soils that might be a part of the missing tropical source, in particular within the waterlogged soils of the rice paddies \cite{yi_exchange_2008}.

\end{itemize}

\section{Conclusion}

We have developed an analytical system that optimizes GPP, plant respiration CO_2 flux and COS soil fluxes within the 15 PFTs defined in the ORCHIDEE terrestrial model, enabling to take into account the ecosystem-dependence of the fluxes. The LRU approach was used to link the GPP to the COS plant uptake. With this system, we have performed a joint assimilation of CO_2 and COS atmospheric measurements into the LMDz atmospheric transport model for the period 2008–2019. From a technical point of view, the inverse system is able to find the components of the CO_2

and CO_2 budgets that give a good fit with assimilated measurements. Inverse results point at a large oceanic CO_2 source between 450 and 600 $\mathrm{GgS.yr}^{-1}$, most of it located in the tropics. The inversion leads to a GPP increase of a few GtC in the high latitudes and a decrease in the same order of magnitude in lower latitudes (tropics and mid-latitudes) compared to the initial prior estimates from the ORCHIDEE LSM. For CO_2 , this means a vegetation sink of around $-620 \mathrm{GgS.yr}^{-1}$, which is in the lower range of recent estimates based on top-down approaches ([launois_new_2015](#)): $-663-772 \mathrm{GgS.yr}^{-1}$, [ma_inverse_2021](#)): $-557--1053 \mathrm{GgS.yr}^{-1}$). The soil sink and the anthropogenic sources have both decreased and amount to -210 and $335 \mathrm{GgS.yr}^{-1}$, respectively. Biomass burning emissions have been slightly revised upward to $65 \mathrm{GgS.yr}^{-1}$. Compared to GPP, plant respiration has almost not been affected in the high latitudes whereas its total value has decreased by only one quarter of the change in GPP in lower latitudes. The resulting CO_2 biospheric fluxes, defined here as the sum of the respiration and GPP, has lost $2 \mathrm{GtC.yr}^{-1}$ above 30°N compared to the prior fluxes simulated by the ORCHIDEE LSM. This behaviour is shared by current inverse systems which infer the net CO_2 fluxes from atmospheric CO_2 measurements ([friedlingstein_global_2020](#)).

Several aspects of the inferred CO_2 fluxes, such as the inter hemispheric gradient, the tropical spatial distribution, the anthropogenic emissions over Japan, China and France, were evaluated with independent atmospheric measurements over different parts of the globe. In the tropics, independent observations of the upper-troposphere CO_2 partial column and the SIF weaken our confidence in the change in tropical GPP; the inverse system actually lacks measurements in this area to ensure a robust partitioning between the oceanic and the continental components of the CO_2 budget. Indeed, the footprint map of the assimilated measurements indicates that the tropical areas, in particular the continents, are **poorly constrained** by the inverse system. The inverse system partly relies on the terrestrial reference fluxes and adjust the tropical source to match the surface measurements over the tropics. If the tropical oceanic release is **indeed** underestimated in the reference fluxes, its magnitude remains highly uncertain. **In contrast**, in the high-latitudes, independent measurements suggest that the inversion has rightly corrected an underestimation of the GPP in the ORCHIDEE land surface model. Concerning the CO_2 anthropogenic sources, Japanese measurements suggest that these are underestimated in Eastern China, **highlighting the need for an improved anthropogenic inventory**.

`\begin{acknowledgements}`

This study was funded by the CO2 Human Emissions (CHE) project which received funding from the European Union, the Horizon 2020 research and innovation program under grant agreement no. 776186. The authors kindly thank the scientists who provided the measurements used in this study. In particular, the MIPAS averaging kernels were provided by Michael Kiefer. The surface measurements from the NOAA network have been performed by scientists affiliated with NOAA (S. Montzka, C. Siso, B. Miller, F. Moore). The airborne measurements from HIPPO have been made with the help of E. Atlas at Univ of Miami (for the HIPPO campaign). D. Yakir facilitates the collection of flask samples at WIS. S. Hattori facilitates the collection of flask samples over Japan. The authors also

a supprimé: CO₂

a mis en forme : Français

a supprimé: -

a mis en forme : Français

a mis en forme : Français

a supprimé: unconstrained

a mis en forme : Français

a supprimé: probably

a mis en forme : Français

a supprimé: Contrariwise

a mis en forme : Français

a supprimé: . Thus,

a mis en forme : Français

a supprimé: large amount of oceanic source could be an upper limit to compensate

a supprimé: too low

a mis en forme : Français

a mis en forme : Français

a supprimé: sources in the tropics.

a mis en forme : Français

thank the LMDz developers for maintaining the dynamism and creativity of this climate model.

`%\appendixee`

`%\section{} %% Appendix A`

`%\subsection{} %% Appendix A1, A2, etc.`

`▲\noappendix %% use this to mark the end of the appendix section`

a mis en forme : Français

%% Regarding figures and tables in appendices, the following two options are possible depending on your general handling of figures and tables in the manuscript environment:

%% Option 1: If you sorted all figures and tables into the sections of the text, please also sort the appendix figures and appendix tables into the respective appendix sections.

%% They will be correctly named automatically.

%% Option 2: If you put all figures after the reference list, please insert appendix tables and figures after the normal tables and figures. %% To rename them correctly to A1, A2, etc., please add the following commands in front of them:

`\appendixfigures %% needs to be added in front of appendix figures`

`\appendixtables %% needs to be added in front of appendix tables`

%% Please add `\clearpage` between each table and/or figure. Further guidelines on figures and tables can be found below.

`\codedataavailability{The COS data were processed at KIT-IMK with the IMK/IAA retrieval processor from http://www.imk-asf.kit.edu/english/308.php. The LMDz model is available from http://web.lmd.jussieu.fr/trac under the CeCILL v2 Free Software License. The COS measurements at station GIF can be downloaded at https://mycore.core-cloud.net/index.php/s/XwULEXsgwFMlTus/download } %% use this section when having data sets and software code available`

`\authorcontribution{MR, FC and PP conceived the research with contributions from SB. MR built the analytical inverse system and conducted the analysis. AP helped in the analysis of the simulation outputs. CA and FM conducted the ORCHIDEE simulations. SL provided the oceanic fluxes of COS and CS_2 . SB provided the observations at GIF site. AB developed the facility used to compute the observation footprints. CB and FM were consulted on their expertise of the SIF data. All coauthors reviewed the paper. } %% optional section`

a supprimé: `%%\authorcontribution{TEXT} %% optional section`

`▲\competinginterests{The authors declare no competing interest} %% this section is mandatory even if you declare that no competing interests are present`

a mis en forme : Français

`%\disclaimer{TEXT} %% optional section`


```
%% REFERENCES
```

```
%% The reference list is compiled as follows:
```

```
\bibliographystyle{copernicus}
```

```
\bibliography{Mabiblio.bib}
```

```
▲\begin{thebibliography}{}
```

a mis en forme : Français

```
%%\end{thebibliography}
```

```
%% Since the Copernicus LaTeX package includes the BibTeX style file  
copernicus.bst,
```

```
%% authors experienced with BibTeX only have to include the following two  
lines:
```

```
%%
```

```
%% \bibliographystyle{copernicus}
```

```
%% \bibliography{example.bib}
```

```
▲
```

a mis en forme : Français

```
%% URLs and DOIs can be entered in your BibTeX file as:
```

```
%%
```

```
%% URL = {http://www.xyz.org/~jones/idx_g.htm}
```

```
%% DOI = {10.5194/xyz}
```

```
%% LITERATURE CITATIONS
```

```
%%
```

```
%% command & example result
```

```
%% \citet{jones90}| & Jones et al. (1990)
```

```
%% \citep{jones90}| & (Jones et al., 1990)
```

```
%% \citep{jones90,jones93}| & (Jones et al., 1990, 1993)
```

```
%% \citep[p.~32]{jones90}| & (Jones et al., 1990, p.~32)
```

```
%% \citep[e.g.,][jones90]| & (e.g., Jones et al., 1990)
```

```
%% \citep[e.g.,][p.~32]{jones90}| & (e.g., Jones et al., 1990, p.~32)
```

```
%% \citeauthor{jones90}| & Jones et al.
```

```
▲\citeyear{jones90}| & 1990
```

a mis en forme : Français

```
%% FIGURES
```

```
%% When figures and tables are placed at the end of the MS (article in
```

```
one-column style), please add \clearpage
```

```
%% between bibliography and first table and/or figure as well as between
```

```
each table and/or figure.
```

```
%% ONE-COLUMN FIGURES
```

```
%%f
```

```
%%\begin{figure}[t]
```

```
%%\includegraphics[width=8.3cm]{FILE NAME}
```

```
%%\caption{TEXT}
```

```
%%\end{figure}
```

```
%%
```

```
%%% TWO-COLUMN FIGURES
```

```
%%
```

```
%%f
%\begin{figure*}[t]
%\includegraphics[width=12cm]{FILE NAME}
%\caption{TEXT}
%\end{figure*}
%
```

TABLES

```
%% The different columns must be separated with a & command and should
%% end with \\ to identify the column brake.
```

ONE-COLUMN TABLE

```
%%t
%\begin{table}[t]
%\caption{TEXT}
%\begin{tabular}{column = lcr}
%\topline
%
%\middleline
%
%\bottomhline
%\end{tabular}
%\belowtable{} % Table Footnotes
%\end{table}
```

TWO-COLUMN TABLE

```
%%t
%\begin{table*}[t]
%\caption{TEXT}
%\begin{tabular}{column = lcr}
%\topline
%
%\middleline
%
%\bottomhline
%\end{tabular}
%\belowtable{} % Table Footnotes
%\end{table*}
```

LANDSCAPE TABLE

```
%%t
%\begin{sidewaystable*}[t]
%\caption{TEXT}
%\begin{tabular}{column = lcr}
%\topline
%
%\middleline
%
%\bottomhline
%\end{tabular}
%\belowtable{} % Table Footnotes
%\end{sidewaystable*}
```

a mis en forme : Français

```

%%% MATHEMATICAL EXPRESSIONS
%
%%% All papers typeset by Copernicus Publications follow the math
typesetting regulations
%%% given by the IUPAC Green Book (IUPAC: Quantities, Units and Symbols
in Physical Chemistry,
%%% 2nd Edn., Blackwell Science, available at:
http://old.iupac.org/publications/books/gbook/green_book_2ed.pdf, 1993).
%%%
%%% Physical quantities/variables are typeset in italic font (t for time,
T for Temperature)
%%% Indices which are not defined are typeset in italic font (x, y, z, a,
b, c)
%%% Items/objects which are defined are typeset in roman font (Car A, Car
B)
%%% Descriptions/specifications which are defined by itself are typeset
in roman font (abs, rel, ref, tot, net, ice)
%%% Abbreviations from 2 letters are typeset in roman font (RH, LAI)
%%% Vectors are identified in bold italic font using \vec{x}
%%% Matrices are identified in bold roman font
%%% Multiplication signs are typeset using the LaTeX commands \times (for
vector products, grids, and exponential notations) or \cdot
%%% The character * should not be applied as multiplication sign
%
%
%%% EQUATIONS
%
%%% Single-row equation
%
%\begin{equation}
%
%\end{equation}
%
%%% Multiline equation
%
%\begin{align}
%& 3 + 5 = 8\\
%& 3 + 5 = 8\\
%& 3 + 5 = 8
%\end{align}
%
%
%%% MATRICES
%
%\begin{matrix}
%& x & y & z \\
%& x & y & z \\
%& x & y & z \\
%\end{matrix}
%
%
%%% ALGORITHM
%
%\begin{algorithm}
%\caption{...}
%\label{a1}
%\begin{algorithmic}
%...

```

```
%\end{algorithmic}
%\end{algorithm}
%
%
%%% CHEMICAL FORMULAS AND REACTIONS
%
%%% For formulas embedded in the text, please use \chem{}
%
%%% The reaction environment creates labels including the letter R, i.e.
(R1), (R2), etc.
%
%\begin{reaction}
%%% \rightarrow should be used for normal (one-way) chemical reactions
%%% \rightleftharpoons should be used for equilibria
%%% \leftrightharpoons should be used for resonance structures
%\end{reaction}
%
%
%%% PHYSICAL UNITS
%%%
%%% Please use \unit{} and apply the exponential notation

\end{document}
```

a mis en forme : Français