Reply to reviewers

Marine Remaud[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, UMR8212, IPSL, Saint-Aubin,France

Correspondence: Marine Remaud (mremaud@lsce.ipsl.fr)

Please find our point-by-point reply below. We would like to thank both reviewers for their careful and constructive evaluation of our paper. Their original text is reproduced in bold type. You will find a track of the changes attached to this document.

1   Reply to Reviewer 1

This paper addresses the global COS budget, and how it can help to constrain GPP. To that end, an analytical inversion
5   **is set up to assimilate CO_2 and COS observations from a limited number of sites globally. COS uptake by the biosphere is coupled to GPP by a Leaf Relative Uptake (LRU) approach. In this way, COS observations inform about GPP, an important term in the CO_2 budget. The impact of adjusted GPP on the CO_2 budget can be compensated by adjusting the CO_2 respiration term, a term that does not affect COS.**

Technically, this is neat piece of work. The results indicate increases in GPP at high Northern latitudes, and reduced
10   **GPP in the tropics, in line with various recent CO_2 inversions.**

Most of the attention is focused on the COS budget. However, the important question is: what is the information that we gain about the CO_2 budget. In that sense, the paper falls a bit short. Is the revised CO_2 budget compatible with "independent" observations (not assimilated)? What follows are a couple of major comments along these lines.

We agree with this summary and thank the reviewer for her/ his appreciation for the large amount of engineering work involved
15   in this study. We want our answers to provide the desired clarifications.

To better understand the role of COS on the Gross Primary Production (GPP), we performed an additional inversion using CO_2 observations (see below). Regarding the compatibility of the revised budget with independent CO_2 observations, the revised budget has been shown to be in better agreement with airborne observations from the HIPPO campaign (see Fig. 10 of the manuscript), in particular in the high latitudes. The simulated concentrations are also closer to observations at most of the 8
20   sites over northern America (see Fig. S7-11). However, we added an evaluation of the CO_2 revised budget against the airborne observations from the ATOMS campaign during the period 2016-2019 (see below).

1.1   Interpretation

One of the main questions is: how can COS help to better constrain GPP/respiration? So, one would expect an inversion without COS, and with COS to compare. Admittedly, a CO_2 inversion that optimizes both GPP and respiration is likely
25   **ill-posed, but with proper priors this should be feasible. In a second step, COS is included. This would be a clean way to**

estimate the potential of COS**. Now, the story is unbalanced, because we cannot "see" the role** COS **observations played.**

Following this recommendation, we performed an inversion using CO_2 observations only aiming at optimizing the GPP and the respiration for each PFT. In the following, this additional inversion is called INV-CO2 compared to the coupled CO_2 - COS inversion called INV-CO2COS. Regarding the suggestion to tune the prior to further reduce ill-posedness, we respectfully
5   disagree: In a Bayesian framework, prior uncertainties are independent from observations. Thus, for INV-CO2, we used the same prior error statistics as for INV-CO2COS. The ratio of CO_2 RMSE of INV-CO2 and INV-CO2COS is between 0.97 and 1.01, with most stations below 1. For the definition of the RMSE, the reader is referred to Equation 7 of the revised manuscript. This suggests no inconsistency between the INV-CO2 results and the COS measurements.

The key figure is Figure 7, which shows the adjustments in GPP + respiration. Clearly, uptake is larger in the NH high-
10   **latitudes (> 50N). Is this supported by independent** CO_2 **observations?**

The larger uptake is supported by independent CO_2 airborne observations from the HIPPO campaign as shown in Figure 10 of the manuscript. The inversion INV-CO2COS clearly reduces the CO_2 residuals between the model and observations in the northern high latitudes. This is further shown in Fig. S11 using independent airborne observations at site ETL in Canada. The inversion corrected the lack of drawdown in the boundary layer in summer.

15   **And the important question is: are the observed changes presented in Figure 8 driven by** COS**?**

To answer this question, we compare on Figure 1a (of this review) the GPP obtained by INV-CO2 with the GPP obtained by INV-CO2COS. We see that the COS observations contribute to increasing the GPP in the boreal regions (PFT 7,8,9,15) and to decreasing the GPP in the tropical regions (PFT2, 3). Figure 1 b) (of this review) further shows that the use of the additional constraint of COS leads to slightly smaller uncertainty on the GPP within each of the PFT compared to an inversion with only
20   CO_2 observations. This is particularly true for the boreal regions. This analysis was added in the Supplement.

The authors show a comparison to SIF only, but I think there are plenty of other metrics, e.g. CO_2 **at independent stations, that could help to assess the realism of the** $COS+CO_2$ **inversion.**

We used the SIF to validate the optimized GPP because (i) this quantity is directly related to the GPP compared to the CO_2 atmospheric concentrations and (ii) the SIF from GOME-2 has a larger spatial cover over the tropics. Moreover, as said
25   earlier, a validation against independent CO_2 airborne observations has already been performed over the Pacific ocean (HIPPO campaign) and at several sites over northern America (Fig. S7-11). To complete the validation, Figure 2 (of this review) compares the simulated CO_2 and COS concentrations against airborne from the ATOM campaign located over the Pacific and Atlantic oceans. Although, some biases remain in the high southern latitudes, the inversion decreases the model-obs mismatch elsewhere and especially in the northern latitudes. As a result, the increase in NBP driven by our inversion is supported by the
30   independent airborne observations from the ATOM campaign.

Now, the authors present a "validation" of COS **by means of MIPAS, HIPPO, NOAA profiles and stations in France and Japan. Specifically the latter comparison is not very useful, although it points to misplaced emissions in the Zumkehr et al. (2018) study.** The latter comparison has actually a different purpose: it evaluates the anthropogenic inventory for Japan.

Also surprising is the lack of interpretation of the posterior covariance matrix (only the error reduction is presented in
35   **Table 6). This is one of the large advantages of the system. This could shed light on how well respiration and GPP can**
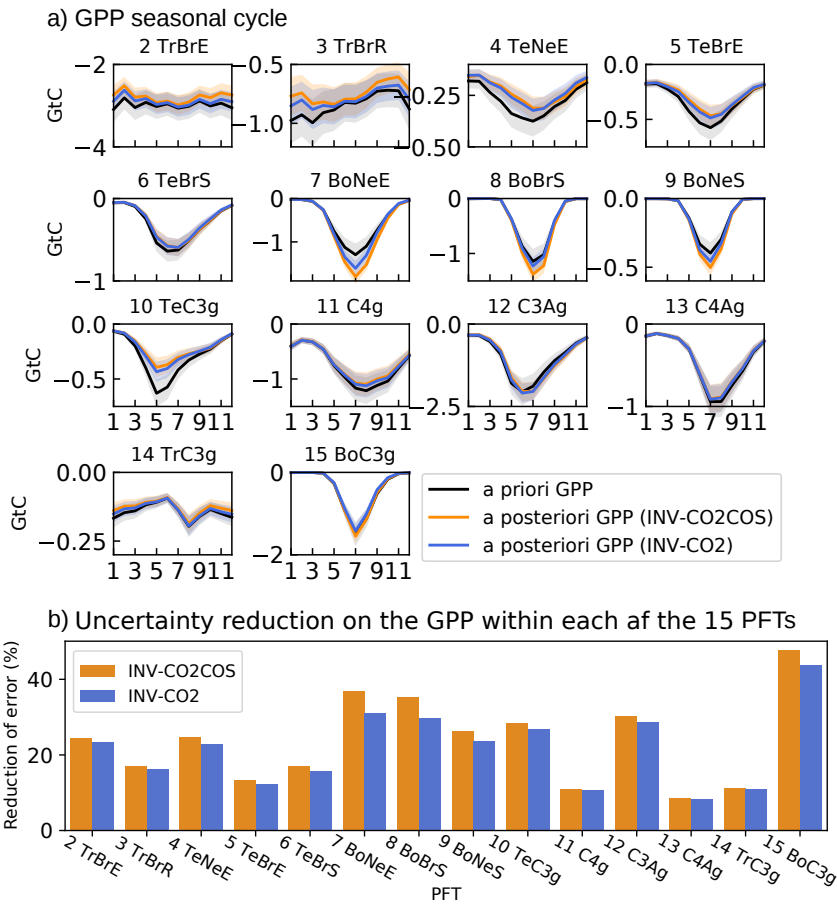
Figure 1. a) Mean seasonal cycle of the total prior (black) and posterior (orange and blue) GPP fluxes and their uncertainties within each of the 15 PFTs during the period 2009-2018. The orange curve is associated with the standard inversion using COS and CO_2 observations. The blue curve is from the inversion using only CO_2 observations. The fluxes have been averaged over 2009-2018. 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 the manuscript. The acronyms Tr, Bo and Te mean Tropical, Boreal and Temperate, respectively. b) Uncertainty reduction for the GPP within each of the 15 PFTs driven by the inversion using CO_2 and COS observations (orange: INV-CO2COS) and CO_2 only observations (blue: INV-CO2) over the years 2009-2019. For the GPP within each PFT, the uncertainty reduction is defined here as 1 minus the root mean square of the ratio between the diagonal terms of the posterior error matrix and those of the prior error matrix.

be separated. To address this shortcomings, Figure 1b (of this review) sheds light on the uncertainty reduction of the annual GPP within each one of the 15 PFTs. The uncertainty reduction is maximal in the boreal regions. Figure 1 has been added in the Supplement.

In the discussion and perspectives session, the "coupled" aspect of the simulations is totally forgotten: it is all about the COS budget, and hardly about CO_2. The issue: "Improving the relationship between COS plant uptake and GPP",
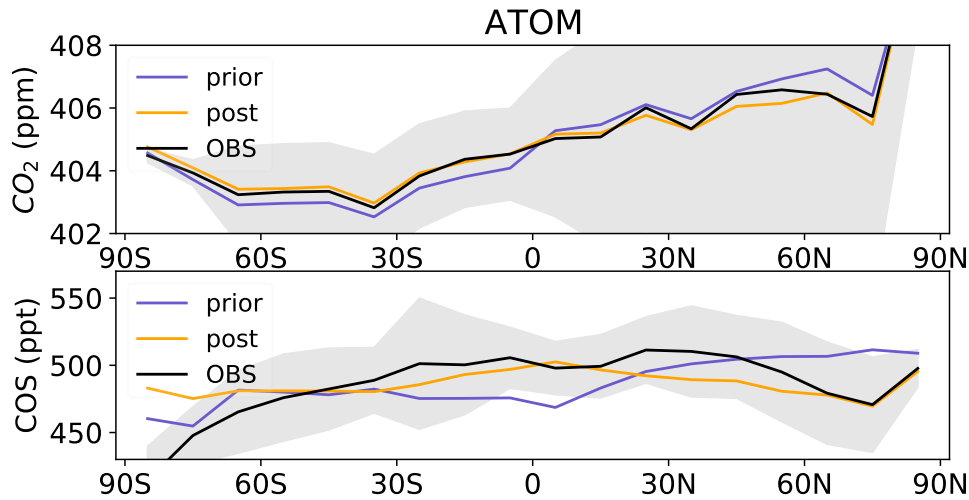
Figure 2. Comparison of the latitudinal variations of the a priori and a posteriori LMDz COS abundance with the ATOM observations. Because of the unbalanced prior, the LMDz COS abundances have been vertically shifted such that the means of the a priori are the same as the mean of the ATOM data (480 ppt). The error bar is calculated as the standard variation of the COS concentration averaged over longitudes and heights.

would be a place to speculate about e.g. the optimization of the LRU per PFT, and how to further address GPP from both COS **and** CO_2.

In this part of the discussion, we added "A complementary experiment would be to optimize a set of LRU coefficient for each PFT together with the GPP fluxes."

5 ## 1.2 The choice of modelling errors (COS/CO_2)

From the description of the errors, I get the impression that the errors have been chosen constant in time, and only varying depending on the station. This sounds to me an oversimplification, because e.g. summer hemispheric fluxes are much more uncertain (due to the active biosphere and their coarse representation). As a consequence, you would expect modelling errors to be much larger in summer.

10 The reviewer seems to point at the impact of aggregation errors, that increase observation errors when sub-represented-scale variability in the fluxes increases. A seasonal variation in aggregation error is an interesting hypothesis, but the ground truth that we have to evaluate it only comes from sparse eddy covariance measurements that can hardly serve the evaluation of this possible change in variability. In any case, the possibility that this effect dominates observation errors is questionable. We could also suggest that modelling errors could be much larger in summer because the upward transport of air masses by convection is

15 stronger, but similarly, we have no statistics on the seasonal variations of transport errors for COS. We have actually initiated

a model inter-comparison experiment for COS aiming at better characterizing the transport errors for COS. The results will be the subject of a future publication.

1.3 Messy

The manuscript is still very messy, with many small mistakes (C instead of S, comments in French, units wrong (e.g.
5 **S12)). I attach an annotated pdf with some (but surely not all) of the errors corrected.**
We would like to thank the reviewer for pointing out these small mistakes. We have double-checked the manuscript. You will find below our answers to all remarks in the annotated PDF.

1.4 Main text: PDF

Page 1, line 14 - unclear, fluxes, budget, ?? We replaced "components of the COS and CO_2" by "components of the COS
10 and CO_2 budgets".
Page 2, line 3 - . For these Done.
Page 2, line 13 - LSMs We replaced "they" by "LSMs".
Page 2, line 29 - I think you need to refer to Berry (2013) : Done.
Page 2, line 35 We replaced hydrolized by oxidized.
15 **Page 3, line 14 - confusing to add this here** We removed "leading to a high oceanic sources".
Page3, line 17 - Greenland?? Thanks for pointing out this mistake. We replaced "Groenland" by "Greenland".
Page 4, line 1 - Here I would for sure mention that this system assimilates CO_2 **and** COS : We changed the sentence to "Here, we present an update of the Launois et al. (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."
20 **Page 4, line 14 - optimized** We replaced "the fluxes" by the "the optimized fluxes".
Page 4, line 24 - s We replaced "inversion" by "inversions".
Page 5, line 4 - This is too technical We have not removed this part of the sentence that outlines the accuracy of the computation regarding the adjoint test. Indeed, this high accuracy adds value compared to the previous inverse system of Launois et al. (2015) that used a rough approximation of the adjoint called "retro transport".
25 **Page 5, line 7 - We talk here about "the plant"** COS **fluxes only** We changed COS fluxes into COS total fluxes.
Page 5, line 7 - Unclear sentence In the inverse system, LMDz is assimilated to its tangent linear. As the Van Leer (1977) advection scheme adds some non linearity, the atmospheric transport in LMDz has to be linearized to obtain the tangent linear model $\mathbf{M}$. We verified the validity of this approximation through a test for the tangent linear model performed for the COS fluxes x_0 and expected flux increment patterns λ. Specifically, we compared the non linear evolution of $\mathbf{M}(x_0 + \lambda)$ with the
30 linear evolution $M(\lambda)$ around an initial state x_0.
We replaced the sentence "We checked that this linearization using CO_2 was still valid for COS fluxes and expected COS flux increment patterns." by "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 $M(x_0 + \lambda)$ with the linear evolution $\mathbf{M}(\lambda)$ for the COS fluxes x_0 (not shown)."

Page 5, line 14 - that We replaced "which" by "that".

Page 5, line 20 - Do I understand that you smear out 1 observations over an 8-day period? Yes, we changed "In practice, we considered average synthetic observations at each selected measurement site (see Section 2.2.1) for each 8-day period between 2008 and 2019." into "In practice, we considered 8-day-average synthetic observations at each selected measurement site (see Section 2.2.1) between 2008 and 2019."

Page 5 - line 20 : What is the implication of this? The implication is that the atmospheric transport model can not represent the temporal variability within a week. We added this sentence Page 5, line 20 of the revised manuscript.

Page 5, line 27 : "Is this for all simulations 1-1-2008? Or is this 9 months prior to observations? Technically, I do not fully understand this, so please explain better." We have removed the forward counted, which was confusing. For each observation, we run the adjoint model backward in time 9 months from observation time.

Page 5, line 34 : I cannot place this 2 weeks??? 2 weeks correspond to the typical frequency of the COS measurements.

Page 5, line 33 - produced We replaced "given" by "produced".

Page 6, line 5 : I note NOAA is not part of the author team. Better place a proper reference here. The reference is written line 18 : "The data represents an extension of the measurements first published in Montzka et al. (2007)."

Page 6, line 6: I would suggest to put these in a table. No action was taken, we do not see the advantage of putting these in a table.

Page 6 , line 22: Why only two months? Would it not be better to show the full sensitivity? Figure 1 shows the transport sensitivities to the sources integrated over 9 months on average for the period 2016-2019. The spatial distribution is more homogeneous but still shows that the tropics are poorly constraint. We added this figure in the Supplement.

Page 6, line 24: the (unless you want to claim that you performed the measurements). We replaced "our" by "the".

Page 6, line 24 - The tropics. We replaced "Tropics" by "The tropics".

Page 6, line 25, What does this mean? Partition? We removed the part of the sentence "the inversion ... zone.", which was confusing.

Page 7, legend of the Figure 1, I read 8 days before? Thanks for pointing out this mistake, we replaced "weekly mean concentrations" by "monthly mean concentrations". The "weeks" are defined to last 8 days in our study, except the last "week" of the month which lasts between 4 and 8 days. We made a monthly average of all the adjoint outputs within a month (4 per months as there are 4 "weeks" per month).

Page 7, Figure 1, I read two months in the text The sentence has been modified from ".. in the previous month" to ".. in the two previous months."

Page 7, legend of the Figure 1 - but at sea level We replaced "..MLO." by "..MLO but at sea level."

Page 8, line 27, Unclear: who tuned? Did you tune, or did Glatthor tune? We removed the words within parenthesis and added the following sentence: "Satellite retrievals are dry-mole fractions tuned by the data providers to the compound World Meteorological Organization (WMO) mole fraction scale."

Page 7, line 10 - s We replaced "..site.." by "..sites..".
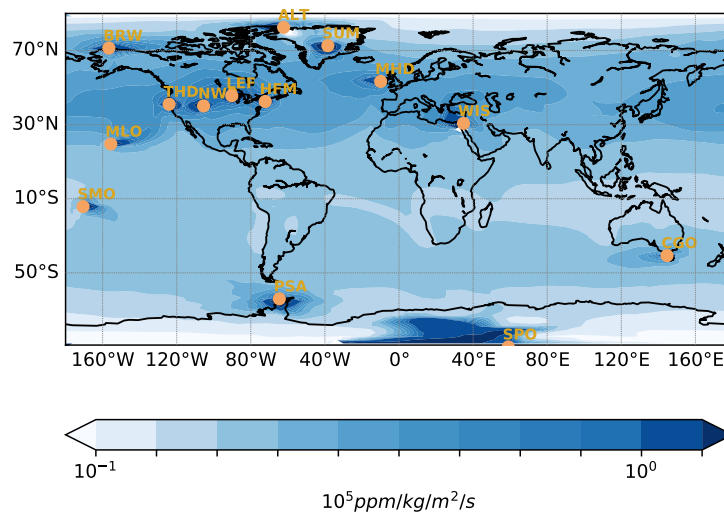
Figure 3. Annual climatology of Jacobians computed by the adjoint of the LMDz model: map of the partial derivatives, in $\mathrm{ppm}/(\mathrm{kg/m^2/s})$, of a monthly mean concentration at all stations from the NOAA network with respect to CO_2 surface fluxes in the nine previous months. The yellow dots denote the location of the surface sites. The site KUM is not depicted as it has the same coordinates than MLO.

Page 7, line 16 - has We replaced ".. COS measurements have been.." by "COS has been..".

Page 8, line 25 - remove (We replace (Belviso et al., 2020) by Belviso et al. (2020).

Page 8, line 25 - "the" We replaced ".. to observations .." by "..to the observations..".

Page 9, line 2 (note again that MIPAS died) No action was requested here and no action was taken.

5 **Page 8, line 29 - mention how many layers ..** We added the number of layers within parenthesis such as ".. MIPAS vertical resolution (60 layers).. ".

Page 9, line 3 : In the MIPAS observations We replaced "In order to dampen the random noise.." by "In order to dampen the random noise in the MIPAS observations..".

Page 9, line 3: Curious that y and yo are used?? We changed y to y^o.

10 **Page 9, line 10 : pseudo (I mean there is the flux limiter issue, right)** We did not add pseudo here as it is called the linear operator. The atmospheric transport has been linearized to become the linear tangent (the linear operator). For the sake of clarity, we replaced "the transport" by "the linearized transport" in this sentence.

Page 10, line 29: above you mentioned 50.. There is no inconsistency. Page 2, line 34, we wrote "OH in the low troposphere while 50 ± 15 GgS.yr^{-1} is photolysed in the stratosphere (Whelan et al., 2018)". This corresponds to the current knowledge

15 but LMDz is in the lower range of the estimates of the statospheric loss.

Page 10, line 43 : al We replaced "observations" by "observational".

Page 11? line 4: reflects? We replaced "means to reflect" by "reflects".

Page 11, Table 2 : I guess this is soil uptake and plant uptake combined. Precision added in the tabular.

Page 11, legend of the Table 2 - global We added the word global "global 4-dimensional variational data-assimilation system".

Page 12, line 9 - ??? We replaced pm by ppm.

Page 12, line 14 - first order We replaced one order by first order.

Page 13, line 2: This number is different from the table. The previous number of 406 $GgS.yr^{-1}$ was for the year 2012 while the number in the Table was for the period 2009-2019. We added the number for the period 2009-2019. **Page 13, line 3: The emissions that are** We replaced Are by 'The emissions that are'.

Page 13, line 11: Formatting Done.

Page 14, line 20 We replaced "small one" by "small control vector".

Page 13, line 13: I would expect two numbers, got three... We replaced "In these, direct, indirect emissions via CS_2 and DMS from the global ocean account for 130 ± 80 GgS.yr^{-1}, 74 ± 120 GgS.yr^{-1} and $65–110$ GgS.yr^{-1}, respectively." by "In these, direct, indirect emissions via CS_2, indirect emissions via DMS from the global ocean account for 130 ± 80 GgS.yr^{-1}, 74 ± 120 GgS.yr^{-1} and $65–110$ GgS.yr^{-1}."

Very confusing. I am under the impression that these are the fields at the start of the simulation (and this is a multiplier to them) True. We removed "monthly".

Page 15, line 14: t We replaced Tropics by tropics.

Page 15, line 15: Is this time-invariant? I do not think that is the best way. In winter many of these errors get much smaller... Here, the reviewer is referred to section 1.2 of this review.

Page 16, line 7: would expect 4 numbers The three numbers correspond the 3 big regions (high-latitudes, mid-latitudes, tropics). For instance over the high latitudes, the PFTs 7,8,9,15 are correlated with a factor of 0.6. Precision added in the text.

Page 17, line 5: The We replaced "Memory effect.." by "The memory effect ..".

Page 17, line 8- Formatting! Done.

Page 17, lines 11 and 12: S? We changed GgC.yr^{-1} by GgS.yr^{-1}.

Page 17, line 30: See also note to cost function, what are y and yo? The variables and the notation were defined in section 2.3. Thanks for pointing out this mistake in the formula that we have corrected.

Page 18, line 3 : from the formula it should be $(N_{obs} + N_{state})$ No action was taken as the written formula is correct. The reader is referred to the publication Chevallier (2007) for a discussion on the χ^2.

Page 18, line 8: including off-diagonal elements Done.

Page 18, line 12 : Do I miss something here? Hxi is the simulated concentration I assumed. We corrected the formula.

Page 18, line 17 : If this is with an unbalanced COS **budget, this is not very surprising!** That is the reason we wrote "as expected" at the beginning of the sentence. No action was taken.

Page 18, line 25: Would it be an idea to present also a metric of the seasonal cycle in the table? (e.g. max-min) The RMSE is a more general metrics, taking into account the errors in both the phase and the amplitude.

Page 18, line 27: Surprising, because the chi-square looks good (around 1). Indeed, these are the stations that receive the

largest error, so this result is what can be expected. To fit better, you need to reduce the errors. The model-obs mismatch can not reach zero further to the aggregation errors (limited number of freedom). In our case, the region to be optimized is too large (the whole northern hemisphere for each PFT) to represent the spatial gradients between LEF and NWR regardless the observation errors. To demonstrate that, we divided the observations errors at three sites by two: LEF, NWR, WIS. Figure 4

5 below shows that the division of observation errors leads to the same observation-model misfit.

Page 18, line 30: Again, fully in line with the error settings As evidenced by the Fig. 4 of this review, smaller observation
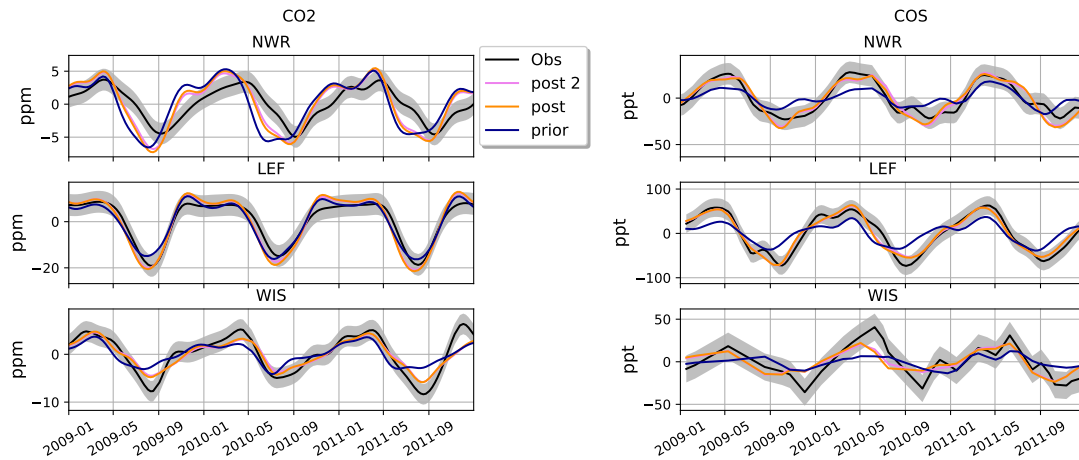
Figure 4. Detrended temporal evolutions of simulated and observed CO_2 and COS concentrations at three selected sites, for the a priori and a posteri fluxes, simulated between 2009 and 2011. Top: Barrow station (BRW, Alaska, USA), middle: Niwot Ridge (NWR, USA) bottom: Park Falls (LEF, USA). The violet curve corresponds to the simulated concentrations for the a posteriori fluxes given by dividing by two the observations errors at LEF, NWR, WIS. 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.

errors do not change the model-observation misfit and thus the results.

Page 18, line 32: ? We replaced "(see Section 9(" by "(see Equation 9)".

Page 19, Table 1: maybe provide units? Done.

10 **Page 19, line 3: how many degrees of freedom are in the state? Since you apply correlations, this is not easy to calculate.** For a description of the method to compute this quantity, the reviewer is referred to Equation 3 of the manuscript. 80% and 20% are the respective contributions of the left and right term to the cost function.

Page 19, line 5 - ??? We replaced "see section 10" by "see Equation 9".

Page 21, line 1 - Well the first reason is that the gap in the prior budget is much smaller in your case We agree with the

15 reviewer. We replaced "First, we assimilated continental surface measurements from the NOAA network through the whole years of 2008-2019 while Kuai et al. (2015a) assimilated a single month of satellite retrievals over the tropical oceans. Second, the Zumkehr et al. (2018) anthropogenic emissions are much higher than the Kettle et al. (2002) one used in these previous studies." by "First, the Zumkehr et al. (2018) anthropogenic emissions are much higher than the Kettle et al. (2002) one used in

these previous studies. Second, we assimilated continental surface measurements from the NOAA network through the whole years of 2008-2019 while Kuai et al. (2015a) assimilated a single month of satellite retrievals over the tropical oceans.

Page 21, line 1: ly We replaced "First" by "Firstly".

Page 21, line 3: ly We replaced "Second" by "Secondly".

Page 21, line 14: i We replaced "oxydation" by "oxidation".

Page 22? Figure 6 We replaced for by over.

Page 22, line 6 - as We replaced than by as.

Page 22, line 11 - calculated as We replaced "given by" by "calculated as".

Page 22, line 12 - calculated from We replaced "one given by" by "flux calculated from".

Page 23, line 3 : material We replaced "Supplementary" by "supplementary material".

Page 23, line 5 - . Done.

Page 23, line 8 - larger We replaced "higher" by "larger".

Page 23, line 15 - the seasonal cycle of the GOME-2 SIF product. We replaced "the one of the SIF from the GOME 2 product." by "the seasonal cycle of the GOME-2 SIF product.".

Page 23, line 17: why minimum and not (also) maximum? I understand that a minimum in SIF corresponds with maximum GPP? For the sake of clarity, we added the sentence: At the ecosystem scale, SIF is anti-correlated with the GPP: a maximum in SIF corresponds with a minimum in GPP. We also replaced "minimum" by "maximum". In the supplementary material, on the Fig. S6, the SIF values have been multiplied by (-1).

Page 23, line temporal dynamics in GPP. We replaced "GPP temporal dynamic" by "temporal dynamics in GPP".

Page 23, line 18 Units and range (0 –> -1) correct? Yes.

Page 23, line 22 - opposite We replaced "in opposition of phase" by "in opposite phase".

Page 23, lines 22 and 23 We removed the two "the".

Page 23, line 23 : Conclude, there is not a clear improvement We added the conclusion: "To conclude, the atmospheric inversion does not lead to a clear improvement in the representation of the GPP seasonal cycle."

Page 23, line 26: Unclear why this is limited to COS? This is not limited to COS. In Figure 10, the CO_2 latitudinal distribution simulated by LMDz were also compared against the airborne observations HIPPO. We also added a comparison with the airborne observations ATOM.

Page 24, legend of the Figure 8 - over We replaced "between" by "over".

Page 24, legend of the Figure 8 - over We removed the "the".

Page 24, line 6 - small over We replaced "weak in" by "small over."

Page 24, line 10 - larger We replaced "weaker" by "larger."

Page 25, line 9 - s We added a s to profile.

Page 24, line 1: Again not surprising since the prior budget is not balanced.. The decrease in RMSE is not obvious as the bias in the prior concentrations has been removed before computing the RMSE. This means that the seasonality of the posterior concentrations is in better agreement with the MIPAS satellite retrievals. We added in the legend of the Figure 9 (manuscript)

"The bias in the prior concentrations has been removed before computing the RMSE."

Page 25, Figure 9: Seems to be a result of the unbalanced budget. RMSE: how calculated ? Unit? In the legend, we added the sentence : The RMSE (see Equation 8) is shown above each panel. We further define the RMSE and its unit in Section 2.7.

Page 25, Figure 9 :??? Seems that SON still has the effects of unbalanced budget??? Since the negative bias in the prior concentrations has been removed, the phase of the seasonal cycle can not be analyzed. However, the latitudinal gradient can be compared against the observations. This pictures only shows that overall the COS values are overestimated in the high latitudes and underestimated over the tropics.

Page 26, line 3: I read in the caption of these figures that a bias was removed? So, it is hard to judge The bias at 3.5 km was removed in the prior vertical profile. This implies that no bias was removed in the posterior vertical profile.

Page 26, Figure 10 - Mention here: becuase ot the unbalanced prior! Done.

Page 26, Figure 10: But over latitudes, longitudes and heights? Some are in the stratosphere..? Unclear There are only three values that are located close to the tropopause (85°N, 8500 m). Removing these values does not change the Figure (see Figure 5 below).
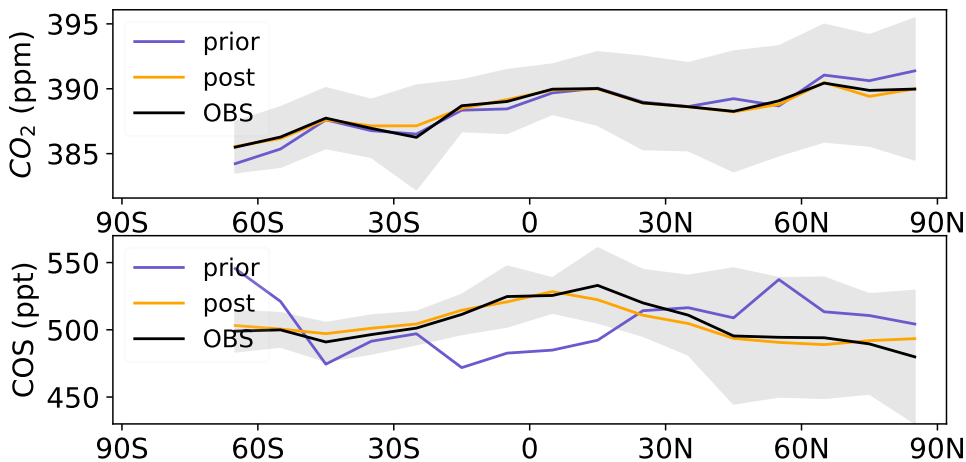
Figure 5. Comparison of the latitudinal variations of the a priori and a posteriori LMDz COS abundance with the HIPPO observations. Because of the unbalanced prior, the LMDz COS abundances have been vertically shifted such that the means of the a priori are the same as the mean of the HIPPO data (521 ppt). The error bar is calculated as the standard variation of the COS concentration averaged over longitudes and heights. The values located in the stratosphere have been removed.

Page 26, line 11- Technically, this should be mixing ratio. We replaced 'concentrations' by 'mixing ratio'.

Page 27, line 2: I think you should present the comparison as time-series, and not as 'average' map. The resolution is relatively coarse, and mis-alignment of transport can easily happen. Within the inversion time window (2009-2019), there

are less than 10 observations available for each Japanese site, which would limit the interpretation of the time series.

Page 27, line 5: I would argue here, that, without assimilation of the observations, the system has many degree of freedom. In that sense, the question is also how the prior emissions perform. The validation is not very helpful. Indeed, in winter, the comparison between the observed and simulated atmospheric concentrations comes down to evaluating the prior

5   anthropogenic inventory and in particular, the placements of the anthropogenic sources. This analysis shows that the inventory needs to be improved over China and Japan in order to assimilate these three stations. Given the observation footprints, in summer, the comparison between the observed and simulated atmospheric concentrations is useful to evaluate the posterior oceanic sources.

Page 27, line 12: not an issues, since the lifetime against oxidation by OH is very long (>10 years).. We removed this part

10   of the sentence.

Page 27, line 12 - small We replaced "low" by "small".

Page 27, line 15 - observations We replaced "the one observed" by "the observations".

Page 27, line 17 - in this season ? We added in this season at the end of the sentence.

Page 27, line 24 - weakened We replaced weaken by weakened.

15   **Page 27, line 34: Should?? Is this planned work, or a recommendation?** Since this is a recommendation, we added the word "should".

Page 28, line 3: Vague recommendation? Correct the spatial distribution of the emissions? Using a model? What a resolution? In fact, using IASI with a low sensitivity to the surface, this will be very difficult. It would be better to establish more measurements, e.g. sample industrial facilities to validate the inventories. We agree with the reviewer

20   and we reformulated the sentence: "Further work should include a more thorough evaluation of the spatial distribution of the European anthropogenic sources using COS retrievals from Fourier transform infrared spectrometry (Wang et al., 2016; Krysztofiak et al., 2015) along with a high-resolution (e.g., 0.5 °in latitude and longitude) chemical transport model. Additional in-situ measurements in the vicinity of industrial facilities are also needed to validate the anthropogenic inventory."

Page 29, line 7: maybe include also a statement about the light-dependence of LRU, which is not taken into account.

25   **Also, stomatal aperture at night might also be important for the COS flux (and less for the CO_2 flux, since uptake is light dependent).** We did not add such a statement since Maignan et al. (2020) showed, using a mechanistic model of the COS plant uptake, that the light-dependence of the LRU does not affect the simulated concentrations at the NOAA stations.

Page 29, line 17: Unclear sentence We reformulated into : "In particular, the DMS emissions simulated by the NEMO PISCES ocean model Belviso et al. (2012) are higher over most of oligotrophic subtropical zones compared to the DMS emissions of

30   Lana et al. (2011). This means that the spatial distribution of the DMS oceanic emissions is highly uncertain."

Page 29, line 22 :again, is this ongoing work, or a recommendation? We added " Future work should include..

Page 29, lin 30 : ?? Sentence does not run In winter, Japan is downwind of the main anthropogenic sources located over Eastern China. We replaced "However, Ma et al. (2021) showed that such simplifications could modify the average COS surface concentrations up to 80 ppt outflow of the anthropogenic sources in winter." by "However, Ma et al. (2021) showed

35   that such simplifications could modify the average COS surface concentrations up to 80 ppt over Eastern China and Japan".

Page 29, line 26 - S We added an S to Teragram.

Page 29, line 12 We replaced "months lag" by "months-lag".

Page 30, line 19 - COS I guess? We replaced "CO_2" by COS.

Page 30, line 24 - negative sign We added the negative sign.

Page 30, line 30: Would this be really needed? The 100 GgS/yr chemistry sink is not very sensitive to the OH distribution, and also e very slow process. So, the exact details will not make a huge difference. The same holds for the stratospheric sink. Once you have the transport pathway to the stratosphere correct, a simple COS **photolysis calculation would suffice....** Ma et al. (2021) showed that including the chemistry of CS_2 and DMS in their transport model could influence the posterior surface COS concentrations up to 40 ppt close in the vicinity of the anthropogenic sources. Thus, we can guess that the distribution of OH could be important if the oxidation of CS_2 and DMS are considered in addition to the oxidation of COS.

Page 30, line 15: One of the recommendations should also be to increase the amount of measurements, specifically over the tropics...You mainy address inventories in the recommendations, but observations are key! The need for additional observations over the tropics has been earlier addressed at the beginning of the discussion.

Page 31, line 4 - Indeed? We replaced "probably" by "indeed".

Pge 31, line 6 - In contrast We replaced contrariwise by "in contrast".

Page 31, line 9: Japan is not tropics We suppressed the last sentence.

1.5 Supplementary

Figure S7 : So there might be over/underestimation? In the legend, we replaced "In order to highlight the difference in profile shape" by "Because of the unbalanced prior budget".

2 Reply to Elliott Campbell

This study advances the COS inversion approach to a new level by simultaneously incorporating COS and CO2. The investigators provide a thorough presentation of the many aspects of their inversion framework and their use of validation data shows the strength of their approach. I have some minor comments below that may improve the communication of this work.

We thank the reviewer for his appreciation of our work and his comments. We followed his recommendations and made some corrections to the manuscript as explained below.

Anthropogenic: It is nice to see our anthropogenic inventory being applied in this study because much has changed in the understanding of anthropogenic emissions since the earlier Kettle inventory was published. While our approach is based on the best available bottom-up information, there are high uncertainties in both the emission factors and the proxy data used for spatial scaling. These uncertainties mean that the anthropogenic inventory is more suitable for analysis over very large regions such as analysis with atmospheric COS measurements from background sites (e.g. MLO). This inventory is less suitable for analysis with atmospheric COS measurements that have significant influence from emissions in smaller regions (e.g. Paris/GIF). A few clarifying points could be added to the mansucript to help the reader understand this distinction.

We removed the sentence "More generally, the disagreement between simulated and observed COS concentrations tested at these sites indicated that there is a large uncertainty in this inventory." and added the following sentences "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)." Concerning GIF, we found no evidence in the measurement record of an emission hot spot nearby: the GIF time series exhibits seasonal variations and a long-term decreasing trend as elsewhere in the northern hemisphere.

Biomass Burning: I am also happy to see our open burning inventory applied in this study. An important note is that the open burning inventory does not include biofuels or agriculture waste. Biofuels and agriculture waste were estimated in Campbell et al. (GRL, 2015), with biofuels being about 3 times as large as open burning. Consider using all three components (biofuel, open, ag residue) in your study or add a note that open burning was included while biofuels and agriculture waste were neglected.

Thank you for pointing out this shortcoming. We added a note at the end of the section 2.4.2 - Biomass burning : "It should be also noted that, compared to the Kettle et al. (2002) inventory, the inventory emissions from Stinecipher et al. (2019) do not include biomass burning sources from agriculture residues and biofuels. The latter were estimated to be about three times as large as open burning emissions (Campbell et al., 2015)." We added the sentence in "including potentially important missing sources" of the discussion: "For instance, the biomass burning sources from biofuels are not included in the Stinecipher et al. (2019) inventory although they were previously estimated to be three times as large as the sources from open burning (Campbell et al., 2015)."

Plant Sink: Some quantification of the error due to the use of a zero-order plant sink instead of an online, first-order approach would be helpful. One way to quantify this error might be to use your posterior COS concentrations and use the lowest model layer in a temporally-explicit way to adjust the plant sink and then run that adjusted plant sink through your transport model and compare to your posterior COS at the sample sites. Perhaps make another version of Fig 5 for the supplement with an additional line that is the run discussed above. And perhaps also add a tropical forest site. Although there are no observation sites in the tropics, it may still be helpful to see how different a zero order and first order approach might be in this region.

As the scaling factor between the GPP and the vegetation fluxes of COS is averaged over each hemisphere, our analytical inverse system can not use an order 1 approach for the vegetation fluxes of COS. However, we have assessed here the impact of using a 0 order approach compared to an order 1 approach. Figure 6 compares the optimized vegetation fluxes of COS (0 order) with the optimized vegetation fluxes of COS which has been rescaled with the varying optimized concentrations (order 1). We see that the inclusion of the order 1 approach decreased the total vegetation fluxes of COS up to 100 GgS annually. The decrease is located in the tropical regions, especially within the PFT2, and to a lesser extent in boreal latitudes. We added the two Figures below in the supplementary material.

Figure 7 of the review further shows that the order 1 approach has a significant influence on the atmospheric COS concentrations at sites AMA, BRW, HFM. For instance at AMA, the seasonal cycle has decreased of 50 ppt.

If using such an approach in the future it might also be important to further estimate the difference between the lowest model layer in the atmospheric transport model and the leaf boundary layer. Your approach to adjusting the zero order plant sink by the time evolving hemispheric means looks like a great way to account for seasonal changes in the mean hemispheric concentrations. You may also want to discuss the large geographic and vertical changes or try and quantify these in your work. The drawdown from the free troposphere to the canopy can be quite large. For example our work in the redwoods, we report measurements from the free troposphere, boundary layer, and canopy that each have a significant drop (Campbell et al., JGR-B, 2017).

Thank you for this interesting point. We added the two sentences in the discussion: "This will involve representing the sharp drop of COS between the canopy and the boundary layer, which can reach 70 ppt in redwood forests (Campbell et al., 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 (Naudts et al., 2015) but more research is needed before they can be used for our application."

Consider adding a sentence explaining how you estimated the constants in equations 6, 7, and 8.

Page 12, equations 6 and 7 from the revised manuscript: We replaced the sentence "They are computed from monthly means at selected stations in this way:" by "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:".

In disregarding the LRU and referencing previous studies that have shown that this is an acceptable simplification for
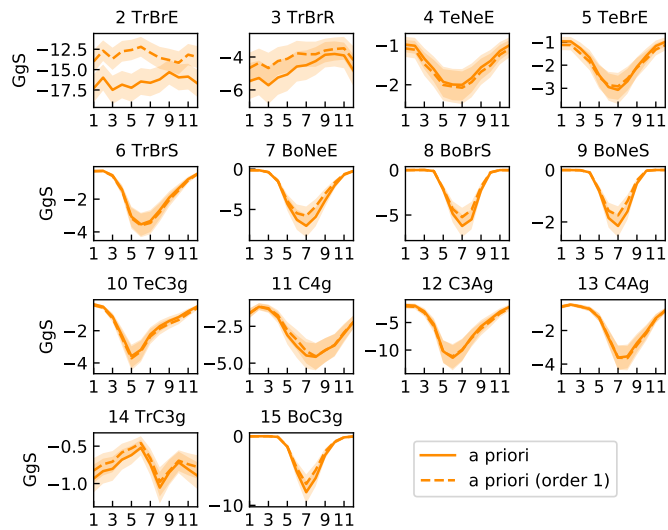
Figure 6. Mean seasonal cycle of the total posterior (orange) and the posterior GPP fluxes using the order 1 approach (orange, dashed line) and their uncertainties within each of the 15 PFTs during the period 2009-2018. The orange curve is associated with the standard inversion using COS and CO_2 observations. The fluxes have been averaged over 2009-2018. 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 acronyms Tr, Bo and Te mean Tropical, Boreal and Temperate, respectively.

regional and global atmospheric simulations you might also explain the physical reason for why this is acceptable (atmospheric mixing causes the plant sink signal from multiple parts of the day).

We added the following sentence Page 13 in the first paragraph: "A physical reason making the LRU simplification acceptable is that the observation sites sample plant sink signals from multiple parts of the day."

5 **Also when noting the LRU relationship to light at ecosystem scale (P12,L21) you might add a reference for where this was first observed at ecosystem scale in Maseyk et al. (2014)**

Thank you for the Maseyk et al. (2014) reference that we added on page 12, line 26 of the revised manuscript.

Ocean fluxes from direct and indirect sources are optimized using the same control parameter but these direct and indirect ocean sources come from very different processes that have distinct geographic and seasonal variation. It might

10 **be important to try an alternative simulation in which the direct and indirect sources are optimized using distinct sets of control variables. Perhaps this could be discussed as a possible direction for future work.**

The physical processes underlying indirect oceanic emissions through DMS are highly uncertain and optimizing them would be interesting. However, there are still large uncertainties regarding the spatial distribution of the DMS emissions (Lana et al., 2011; Belviso et al., 2012) and the chemistry of DMS into the atmosphere (Von Hobe, personal communications). Because of
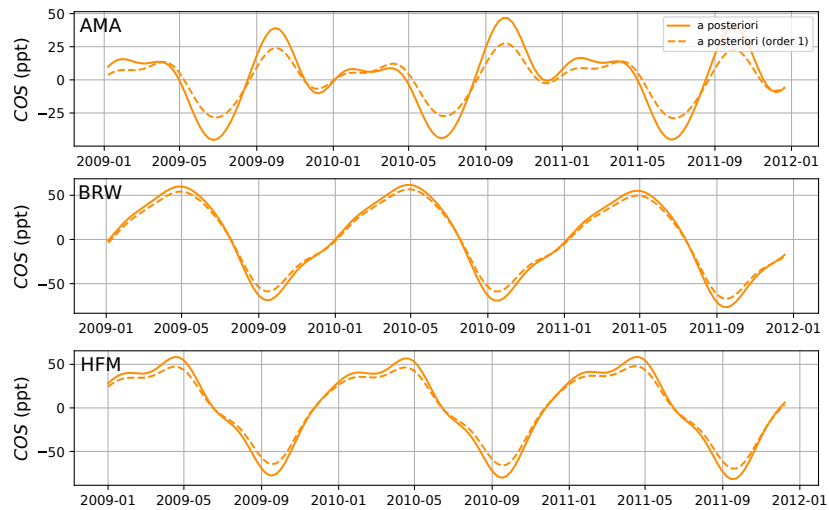
Figure 7. Detrained temporal series of the optimized concentrations using the 0 order approach (full line, orange) and the optimized concentrations using the order 1 approach (dashed line, orange) at sites AMA (top), BRW(middle) and HFM (bottom). The site AMA is located in the Amazon basin in Brazil. The mean differences between the two curves are 40, 30, 50 ppt at BRW, HFM, AMA over the years 2008-2011.

these reasons, the inversion cannot really separate the direct from the indirect oceanic sources, even with an extended control vector. We have added the following sentence in the second paragraph of section 2.5.1. "The ocean emissions are 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, we take all ocean emissions as a whole." In addition, in the discussion, we added the following

5   sentence "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.".

P5L19 "In practice, we considered only…" Consider adding some text to not if this leaves you with enough observations at each site. Since your control variable is amplitude does this mean you need enough observed samples per month to constrain the amplitude at each site? For example, would you want to make sure a site had at least one measurement

10 **per month in a given year to use that year's data in your inversion?**

We added the Figure below in the supplementary material and added these two sentences at Page 6 at the end of the first paragraph: "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)."
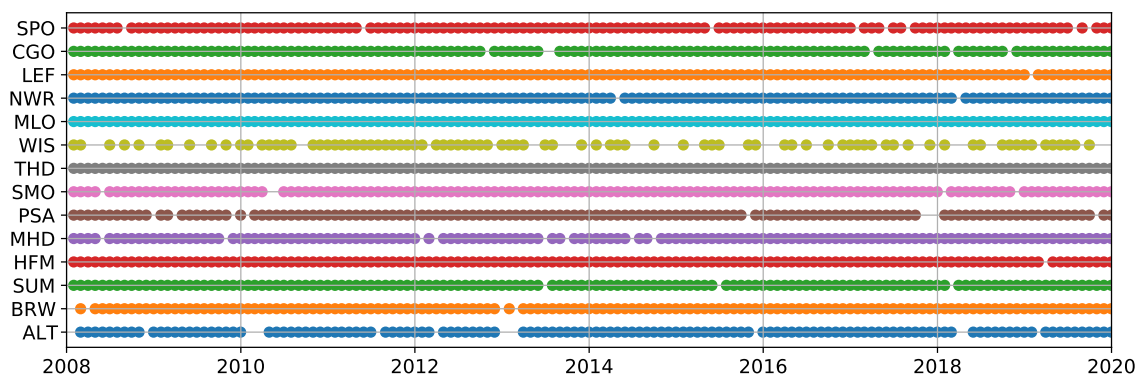
Figure 8. Sampling dates of the COS measurements for the stations of the NOAA network

P23L20 "The seasonal cycle is degraded..." you might want to expand this sentence into a few sentences to explain more clearly to the reader what you are talking about.

We have replaced the sentence "The seasonal cycle is degraded 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 COS
5 absorption over the tropics." by "The optimized seasonal cycle disagrees with the SIF satellite retrievals ... ".

Figure 1 is a very helpful illustration of the regions of influence. However some additional text could be helpful here. In particular, this figure (as well as some of the text) will leave readers thinking that the inversion has very little to offer in terms of information about the tropics. However as we see from your validation data in the tropics (HIPPO and MIPAS), the inversion does an excellent job of improving model skill in the tropics in general. While there might not
10 **be much to offer in terms of the tropical forests, it seems that SMO and possibly other sites are sufficient for improving the marine boundary layer in the tropics. Providing this context along with Figure 1 could be helpful.**

Further to a request from Reviewer 1, we have added the same figure in the supplementary material but with the observation footprints integrated for 9 months backward. It seems that the tropics are constrained in a homogeneous way by remote air masses coming from the assimilated stations. We have added the following sentence in the main text: "However, the tropical
15 areas are slightly constrained by well mixed air masses coming from remote stations (see Fig. S2)."

Overall, this is an exciting paper that pushes the COS inversion approach into new territory. In particular, their study points to the need for additional tropical observations and improved bottom-up information on anthropogenic sources. Given the dearth of measurements on these two important points, it seems that there is significant opportunity to further advance the method and generate new discoveries in carbon cycle science.
20 We thank the reviewer again for his positive appreciation.

References

Belviso, S., Masotti, I., Tagliabue, A., Bopp, L., Brockmann, P., Fichot, C., Caniaux, G., Prieur, L., Ras, J., Uitz, J., Loisel, H., Dessailly, D., Alvain, S., Kasamatsu, N., and Fukuchi, M.: DMS dynamics in the most oligotrophic subtropical zones of the global ocean, Biogeochemistry, 110, 215–241, https://doi.org/10.1007/s10533-011-9648-1, https://doi.org/10.1007/s10533-011-9648-1, 2012.

5 Belviso, S., Lebegue, B., Ramonet, M., Kazan, V., Pison, I., Berchet, A., Delmotte, M., Yver-Kwok, C., Montagne, D., and Ciais, P.: A top-down approach of sources and non-photosynthetic sinks of carbonyl sulfide from atmospheric measurements over multiple years in the Paris region (France), PLOS ONE, 15, e0228 419, https://doi.org/10.1371/journal.pone.0228419, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228419, 2020.

Campbell, J. E., Whelan, M. E., Seibt, U., Smith, S. J., Berry, J. A., and Hilton, T. W.: Atmospheric carbonyl sulfide

10 sources from anthropogenic activity: Implications for carbon cycle constraints, Geophysical Research Letters, 42, 3004–3010, https://doi.org/10.1002/2015GL063445, https://onlinelibrary.wiley.com/doi/abs/10.1002/2015GL063445, 2015.

Chevallier, F.: Impact of correlated observation errors on inverted CO2 surface fluxes from OCO measurements, Geophysical Research Letters, 34, https://doi.org/https://doi.org/10.1029/2007GL030463, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030463, 2007.

15 Kettle, A. J., Kuhn, U., Hobe, M. v., Kesselmeier, J., and Andreae, M. O.: Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks, Journal of Geophysical Research: Atmospheres, 107, ACH 25–1– ACH 25–16, https://doi.org/10.1029/2002JD002187, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002187, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2002JD002187, 2002.

Krysztofiak, G., Té, Y. V., Catoire, V., Berthet, G., Toon, G. C., Jégou, F., Jeseck, P., and Robert, C.: Carbonyl Sulphide (OCS) Variabil-

20 ity with Latitude in the Atmosphere, Atmosphere-Ocean, 53, 89–101, https://doi.org/10.1080/07055900.2013.876609, https://doi.org/10.1080/07055900.2013.876609, 2015.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochemical Cycles, 25, https://doi.org/https://doi.org/10.1029/2010GB003850, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/

25 2010GB003850, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010GB003850, 2011.

Launois, T., Belviso, S., Bopp, L., Fichot, C. G., and Peylin, P.: A new model for the global biogeochemical cycle of carbonyl sulfide – Part 1: Assessment of direct marine emissions with an oceanic general circulation and biogeochemistry model, Atmospheric Chemistry and Physics, 15, 2295–2312, https://doi.org/https://doi.org/10.5194/acp-15-2295-2015, https://www.atmos-chem-phys.net/15/2295/2015/, publisher: Copernicus GmbH, 2015.

30 Ma, J., Kooijmans, L. M. J., Cho, A., Montzka, S. A., Glatthor, N., Worden, J. R., Kuai, L., Atlas, E. L., and Krol, M. C.: Inverse modelling of carbonyl sulfide: implementation, evaluation and implications for the global budget, Atmospheric Chemistry and Physics, 21, 3507–3529, https://doi.org/10.5194/acp-21-3507-2021, https://acp.copernicus.org/articles/21/3507/2021/, publisher: Copernicus GmbH, 2021.

Maignan, F., Abadie, C., Remaud, M., Kooiijmans, L. M. J., Kohonen, K.-M., Commane, R., Wehr, R., Campbell, J. E., Belviso, S., Montzka, S. A., Raoult, N., Seibt, U., Shiga, Y. P., Vuichard, N., Whelan, M. E., and Peylin, P.: Carbonyl Sulfide: Comparing a Mechanistic Repre-

35 sentation of the Vegetation Uptake in a Land Surface Model and the Leaf Relative Uptake Approach, Biogeosciences Discussions, pp. 1–41, https://doi.org/https://doi.org/10.5194/bg-2020-381, https://bg.copernicus.org/preprints/bg-2020-381/, publisher: Copernicus GmbH, 2020.

Montzka, S. A., Calvert, P., Hall, B. D., Elkins, J. W., Conway, T. J., Tans, P. P., and Sweeney, C.: On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO2, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD007665, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007665, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD007665, 2007.

5 Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., Campioli, M., Ghattas, J., De Groote, T., Haverd, V., Kattge, J., MacBean, N., Maignan, F., Merilä, P., Penuelas, J., Peylin, P., Pinty, B., Pretzsch, H., Schulze, E. D., Solyga, D., Vuichard, N., Yan, Y., and Luyssaert, S.: A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes, 8, 2035–2065, https://doi.org/10.5194/gmd-8-2035-2015, https://gmd.copernicus.org/articles/8/2035/2015/, 2015.

10 Stinecipher, J. R., Cameron-Smith, P. J., Blake, N. J., Kuai, L., Lejeune, B., Mahieu, E., Simpson, I. J., and Campbell, J. E.: Biomass Burning Unlikely to Account for Missing Source of Carbonyl Sulfide, Geophysical Research Letters, 46, 14 912–14 920, https://doi.org/10.1029/2019GL085567, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085567, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL085567, 2019.

Van Leer, B.: Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection, Journal of
15 Computational Physics, 23, 276–299, https://doi.org/10.1016/0021-9991(77)90095-X, http://www.sciencedirect.com/science/article/pii/002199917790095X, 1977.

Wang, Y., Deutscher, N. M., Palm, M., Warneke, T., Notholt, J., Baker, I., Berry, J., Suntharalingam, P., Jones, N., Mahieu, E., Lejeune, B., Hannigan, J., Conway, S., Mendonca, J., Strong, K., Campbell, J. E., Wolf, A., and Kremser, S.: Towards understanding the variability in biospheric CO_2 fluxes: using FTIR spectrometry and a chemical transport model to investigate the sources and sinks of carbonyl
20 sulfide and its link to CO_2, Atmospheric Chemistry and Physics, 16, 2123–2138, https://doi.org/https://doi.org/10.5194/acp-16-2123-2016, https://acp.copernicus.org/articles/16/2123/2016/, publisher: Copernicus GmbH, 2016.

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

```
%% 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
l

%% 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

```latex
% 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 Dioxyde and Carbonyl Sulphide
measurements into the LMDz transport model: implication for the global
$\mathrm{COS}$ 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]{Fr√©d√©ric}{Chevallier}
\Author[1]{Fabienne}{Maignan}
\Author[1]{Sauveur}{Belviso}
\Author[1]{Antoine}{Berchet}
\Author[1]{Alexandra}{Parouffe}
\Author[1]{Camille}{Abadie}
\Author[1]{C√©dric}{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]{C√©dric}{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}

%\affil[]{ADDRESS}

%% 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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ deposition velocities in plants. A possible scenario for the period 2008-2019 leads to a global biospheric sink of 800 $\mathrm{GgS.yr^{-1}}$, 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.

\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 \citep{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
\citep{anav_spatiotemporal_2015}.

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 \citep{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
\citep{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 \citep{anav_spatiotemporal_2015}.
Satellite retrievals of, e.g., solar-induced fluorescence (SIF) or

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 $\mathrm{CO_2}$ atmospheric budget.

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 $\mathrm{CO_2}$ is re-emitted into the atmosphere through respiration, COS is nearly irreversibly hydrolized 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 $\mathrm{CO_2}$ 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}.

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 $\mathrm{GgS.yr^{-1}}$, \citet{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 ($\mathrm{CS_2}$) are two other candidates that may

a supprimé: the
a mis en forme : Français
a mis en forme : Français
a supprimé: hydrolized
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, \citet{lennartz_monthly_2020} developed a mechanistic model to simulate COS emissions via $\mathrm{CS_2}$ and estimated a global source of 70 $\mathrm{GgS.yr^{-1}}$, 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 NOx at 298 K \citep{barnes_ftir_1996}. The recent identification of novel DMS oxidation products \citep{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 \citet{kettle_global_2002} used by most top-down studies has been demonstrated to be incomplete \citep{blake_carbonyl_2004,du_important_2016}. The anthropogenic inventory has been revised upward from 200 $\mathrm{GgS.yr^{-1}}$ to 400 $\mathrm{GgS.yr^{-1}}$, with the largest source shifting from North America to Asia \citep{zumkehr_global_2018}. Yet, firn air sampled in Antartica and Greenland suggests that anthropogenic emissions are still underestimated and are closer to 600 $\mathrm{GgS.yr^{-1}}$ \citep{aydin_anthropogenic_2020}.

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 $\mathrm{GgS.yr^{-1}}$ \citep{suntharalingam_global_2008,berry_coupled_2013,kuai_estimate_2015,launois_new_2015}. This finding is consistent with the high concentrations of COS observed over tropical waters \citep{montzka_global_2007,glatthor_tropical_2015,kuai_estimate_2015}, but remains preliminary due to the scarcity of observations \citep{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 \citet{berry_coupled_2013} calibrated a single scaling factor for the oceanic source per latitudinal band. \citet{launois_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, \citet{kuai_estimate_2015} divided the tropics into several regions and optimized one scaling coefficient of the oceanic source per region. Recently, \citet{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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ 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$^o\times$1.9^o (longitude times latitude) with 39 layers in the vertical, based on the general circulation model of the Laboratoire de M\'et\'eorologie 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 $\textbf{M}{\bf x}$ and $\textbf{M}^T{\bf y}^*$, with $\textbf{M}$ the Jacobian matrix of LMDz, ${\bf x}$ a vector of input variables of LMDz (i.e. tracer surface fluxes and initial tracer values), and ${\bf y}^*$ a vector of size the number of output variables (i.e. the four-dimentional 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 \citet{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 $\textbf{M}{\bf x}$ and $\textbf{M}^T{\bf y}^*$, with $\textbf{M}$ the Jacobian matrix of LMDz, ${\bf x}$ a vector of input variables of LMDz (i.e. tracer surface fluxes and initial tracer values), and ${\bf y}^*$ 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 $\mathrm{CO_2}$ surface fluxes from the Copernicus Atmosphere Monitoring Service (https://atmosphere.copernicus.eu/). "We checked that this linearization using $\mathrm{CO_2}$ 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 $M(x_0+\lambda)$ with the linear evolution $\textbf{M}(\lambda)$ 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 \citep{launois_new_2015,peylin_new_2016} which used a rough approximation of the adjoint $\textbf{M}^T{\bf y}^*$, called "retro-transport", in which the direction of time was simply reversed in LMDz without strict inversion of the order of calculations \citep{hourdin_eulerian_2006}. In addition, we use a much more recent version of LMDz here (LMDz6A, \citet{remaud_impact_2018}, vs. LMDz3, \citet{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 $\textbf{M}$ that correspond to the places where, and the times when, we have observations to assimilate. By definition, each value of $\textbf{M}$ 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 $\textbf{M}^T{\bf y}^*$ for each observation to assimilate, with the elements of ${\bf y}^*$ set to zero or one. We use the Community Inversion Framework (CIF, \citet{berchet_community_2020}) to manage these computations.

a supprimé: patterns

a supprimé: ,

a supprimé: which

%In our case, $\bf{M}^T y^*$ 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 $\bf{y}$ at a given station and
time to a source $\bf{S}$ through the relation
\citep{hourdin_eulerian_2006-1}:
%with $y{*}=(Y-H(X))R^{-1}$.
%\begin{equation}
%\bf{y} = H_{0}^T C|t_0+ H_s^{T}S \approx H_s^{T} S
%\label{EqAdjoint}
%\end{equation}
%with $\bf{H_0}^{T}$ and $\bf{H_s}^{T}$ being the sensitivity of the
model concentration to the initial conditions $\bf{C|t_0}$ and to the
sources S, respectively. In our case, the source 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 providen that the mass of a passive traceur 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 \citet{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 ${\bf y}^*$), 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: %\textbf{MR: Le paragraphe
suivant est long et trV®s technique. Au
cas ou le journal choisi n'est pas gmd ,
peut √™tre 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 $\mu\mathrm{mol.mol^{-1}}$, or parts per million, ppm). We have verified that the $\mathrm{CO_2}$ and COS concentrations obtained by the resulting Jacobian matrix ($\textbf{M}{\bf x}$) match well the one produced by the full LMDz transport model over the period (See Fig. S3 of the Supplementary materiel).

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 $\bf{H}$ has been computed using the retro-transport of the LMDz atmospheric model instead of the adjoint transport \citep{peylin_daily_2005,launois_new_2015,peylin_new_2016}. Even though the adjoint transport is mathematically equivalent to the retro traceur transport \citep{hourdin_eulerian_2006}, the two quantity are conceptually different and obtained through two different ways. The adjoint transport is computed using systematic math technic 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 \citep{remaud_impact_2018} for the computation of the adjoint, the version used in
\citet{peylin_daily_2005,launois_new_2015} was LMDz3
\citep{hourdin_lmdz4_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: COS and $\mathrm{CO_2}$ surface sites}
We used the NOAA/ESRL measurements of both $\mathrm{CO_2}$ 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,Äîinlet 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\degree N, 35.06\degree E, 151 asl), Palmer Station, Antarctica, United States (PSA, 64.77\degree S, 64.05\degree W, 10.0 asl), South Pole, Antarctica, United

a supprimé: given

a supprimé: S2

States (SPO, 89.98\degree S, 24.8\degree 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}.

The Jacobian Matrix $\textbf{M}$ 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\degree 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.
\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 $\mathrm{ppm/(kg/m^2/s)}$, of a monthly mean concentration at all stations from the NOAA network with respect to $\mathrm{CO_2}$ 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.}
\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.

a supprimé: citep

a supprimé: Tropics

a supprimé: our

a supprimé: inversion will not be able to partition the COS and $\mathrm{CO_2}$ components in this zone. The…

a supprimé: north

a supprimé: south

a supprimé: weekly

a supprimé: month

\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.

%\textbf{MR: Est ce que je dois introduire les observations a√©roport√©es
de la NOAA m√™me si elles sont montr√©es uniquement dans le Supplement et
qu'elles sont mentionn√©es une fois dans le texte (en renvoyant au
Suppl√©ment)?}

The first observation program is the HIAPER Pole-to-Pole Observations
(HIPPO, \citet{wofsy_hiaper_2011}). HIPPO consisted of five aircraft
transects of many trace gas measurements, including for COS
and $\mathrm{CO_2}$, 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 site in Japan from \citet{hattori_constraining_2020} have been
used as well: Miyakojima (24¬∞80,Ä≤N, 125¬∞27,Ä≤E), Yokohama (35¬∞51,Ä≤N,
139¬∞48,Ä≤E), and Otaru (43¬∞14,Ä≤N, 141¬∞16,Ä≤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 \citep{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 \citep{belviso_top-
down_2020}. According to the recent COS global gridded
anthropogenic emission inventory of \citep{zumkehr_global_2018}, the
Paris region is an important source of COS (79\overline{1} 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 \citet{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
\citep{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
\citep{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 \citet{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{frankenberg_new_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: $https://avdc.gsfc.nasa.gov/pub/data/satellite/MetOp/GOME\_$F/v28/. We used the monthly level 3 product gridded at a 0.5\degree 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, \citet{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-independant.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\degree 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:
%\begin{equation}
%  X_s   =X_a +A[X_m - X_p].
%\end{equation}
%where $Xs$, $Xa$, and $Xm$ are smoothed, a priori, and model vertical
profile, respectively, and $A_k$ is the averaging kernel matrix.

The a priori profile for the $\mathrm{COS}$ retrievals is a zero profile
\citep{glatthor_tropical_2015}, hence it had not to be taken into
account. As done in \citet{glatthor_tropical_2015}, we focus here on the
spatial distribution of the $\mathrm{COS}$ mixing ratio at the 250 hPa
pressure level (still after convolution of the model with the averaging
kernels) for the period 2008‚Äì2012. In order to dampen the random noise
in the MIPAS observations, we aggregate the retrievals into in
5\degree$\times$15\degree latitude-longitude bins.

\subsection{Inverse framework}

Our inverse system seeks to estimate the amplitude of $n$ sources or
sinks of $\mathrm{CO_2}$ and $\mathrm{COS}$ gathered in a vector $\bf{x}$
by reducing the mismatch between the observed concentrations gathered in
a vector $\bf{y^o}$ and those simulated with the atmospheric transport
model $\textbf{M}$ 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 $\textbf{M}$ is part of the linear observation
operator $\textbf{H}$ that relates $\bf{x}$ and the model-equivalent
$\mathrm{CO_2}$ and $\mathrm{COS}$ measurements $\bf{y}$ at the sites
shown in Fig. \ref{Obs}:
\begin{equation}
    \textbf{y} = \textbf{H} \textbf{x}
\end{equation}

In order to regularize the inverse problem corresponding to Eq. (1), we
use a Bayesian framework involving an a priori control vector, $\bf{x^b}$
with its associated uncertainty statistics, summarized in covariance
matrix $\textbf{B}$. 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
\citet{tarantola_1989}) for the the posterior control vector $\bf{x}^a$
and the uncertainty covariance matrix $\textbf{P}^a$ :
\begin{equation}
\begin{aligned}
\bf{x} & = \bf{x^b} + \textbf{B} \textbf{H}^T (\textbf{H} \textbf{B}
\textbf{H}^{T} + \textbf{R})^{-1}(\bf{y^o} - \textbf{H} x^{b})\\
\textbf{P}^a & = \textbf{B} -  \textbf{B} \textbf{H}^T (\textbf{H}
\textbf{B} \textbf{H}^{T} + \textbf{R})^{-1} \textbf{HB}
\end{aligned}
\end{equation}

with $\textbf{R}$ the error covariance matrix of the observations,
encompassing measurement errors and $\textbf{H}$ 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(\bf{x}) =\frac{1}{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{CostF}
\end{align}

\subsection{Gridded reference fluxes}

In the following, we call "reference fluxes" the maps of $\mathrm{CO_2}$
and COS fluxes that are used in the observation operator, the
control vector $\bf{x}$ being a low-resolution multiplier to these (see
Section 2.5.1).
For use at resolution 3.75\degree$\times$1.90\degree, the maps of the
following components of the $\mathrm{CO_2}$ and COS fluxes
have been interpolated from their native resolution. All projections
conserved mass.

\subsubsection{$\mathrm{CO_2}$ 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) \citep{jones_gridded_2021}. Biomass burning fluxes vary
inter-annually and are described by the GFED 4.1s database
(https://www.globalfiredata.org/data.html). Monthly air‚Äêsea
$\mathrm{CO_2}$ exchange is prescribed from the Copernicus Marine
Environment Service database \citep{denvil-sommer_lsce-ffnn-v1_2019}.
The GPP and respiration fluxes have been simulated at a resolution of
0.5\degree~ both in longitude and latitude by the ORCHIDEE land surface
model \citep{krinner_dynamic_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 \citep{poulter_plant_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 $\mathrm{CO_2}$ 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 $\mathrm{GtC.y^{-1}} $ during
2008-2019. This value is within the range of the GPP estimates (106-137
$\mathrm{GtC.yr^{-1}}$) based on photosynthesis proxies (see Table S1)
\citep{beer_temporal_2009,beer_terrestrial_2010,welp_interannual_2011,ale
mohammad_water_2017,jasechko_global_2019,jung_scaling_2020,ryu_integratio
n_2011,badgley_terrestrial_2019,stocker_drought_2019}. The PFTs and their
acronyms are defined in Table \ref{TLRU}. Note that GPP, respiration,
COS vegetation and soil fluxes are null within PFT 1 (base
soil).

\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_{tot}}{SIF_{tot}}
%\end{equation}
%with $GPP_{tot}$ and $SIF_{t}t$ 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 cholorphyle
mollecules 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 $\mathrm{COS}$ budgets that are considered are
biomass burning, soil emissions and uptake, anthropogenic emissions,
plant uptake, oceanic emissions and the atmospheric oxidation by the
$\mathrm{OH}$ 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||ccccccc|}
\hline
   & PROCESSES & \citet{kettle_global_2002} & \citet{montzka_global_2007}
& \citet{suntharalingam_global_2008}& \citet{berry_coupled_2013} &
\citet{launois_new_2015} & \citet{ma_inverse_2021}&This study \\
```

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   & ‚Äì1335 & -
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
\citet{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}

\subsubparagraph{\bf{Soil}}

Reference air-surface exchanges from oxic soils have been simulated by
the steady-state analytical model of \citet{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 \citep{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
$\mathrm{COS}$ uptake reflects for the most part the activity of CA,
ubiquitous in soil microorganism, which efficiently converts
$\mathrm{COS}$ into $\mathrm{H_2S}$ and $\mathrm{CO_2}$, 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
\citet{meredith_coupled_2018} on different biomes in the laboratory. The
production term simulates the $\mathrm{COS}$ abiotic production from
soils via the \citet{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 \citet{berry_coupled_2013}
model used in previous top-down studies. As for the contribution of
anoxic soils \citep{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 COS uptake by leaves
from \citet{sandoval-soto_global_2005} given by the linear relationship
\begin{equation} F_{COS} = GPP \times \frac{[COS]}{[CO_2]} \times
\frac{v_{COS}}{v_{CO_2}}
\label{eq1}
\end{equation}

In this equation, F_{COS} and GPP are the COS uptake and
the $\mathrm{CO_2}$ uptake (both in $\mathrm{ppm/m^2/s}$), respectively,
$[COS]$ and $[CO_2]$ being the ambient air concentrations of
COS and $\mathrm{CO_2}$. v_{COS} and v_{CO_2} are the
COS and $\mathrm{CO_2}$ leaf uptake velocities.
The ratio of uptake velocities of COS compared to
$\mathrm{CO_2}$ is defined as the LRU: \begin{equation}
LRU= \frac{v_{COS}}{v_{CO_2}}
\label{eqLRU}
\end{equation}
We use a zero-order LRU approach (i.e. with no interaction between
vegetation and COS mixing ratio), given the complexity of an
order approach (i.e. a coupled atmospheric COS concentration -
COS flux calculation). To address this shortcoming, we use the
time-evolving hemispheric means of the COS and $\mathrm{CO_2}$
atmospheric concentrations, NH_{mean} and NS_{mean} as done in
\citet{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:
\begin{align}
SH_{mean}= [SPO \times 0.408 + CGO \times 0.770 + SMO \times
0.974]/[0.408 + 0.770 + 0.974]
\end{align}
\begin{equation}
\begin{align*}
NH_{mean}= [(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_201
9}. Our LRU set is derived from \citet{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
\citet{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 COS fluxes.
Morevoer, \citet{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

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 epyphites 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 \citet{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 \citet{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
\citet{whelan_reviews_2018}. }
%\end{table}

\subsubparagraph{\bf{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 $\mathrm{GgS.yr^{-1}}$ (range of 223 - 586
$\mathrm{GgS.yr^{-1}}$) 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
$\mathrm{CS_2}$ 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 \citet{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.

```
\subsubparagraph{\bf{Ocean}}

$\mathrm{COS}$ is directly emitted by the ocean in places where the sea
water is saturated in $\mathrm{COS}$. Emissions typically happen in
summer in high latitudes. $\mathrm{COS}$ is also indirectly emitted
through the oxidation of $\mathrm{DMS}$ and $\mathrm{CS_2}$ in the
atmosphere, which are both produced in seawater. We use the indirect and
direct $\mathrm{COS}$ emissions from
\citet{lennartz_direct_2017,lennartz_monthly_2020} whose total emissions
account for 285 $\mathrm{GgS.yr^{,àí}}$. In these, direct, indirect
emissions via $\mathrm{CS_2}$, indirect emissions via $\mathrm{DMS}$ from
the global ocean account for $130 \pm 80$ $\mathrm{GgS.yr^{-1}}$, $74\pm
120$  $\mathrm{GgS.yr^{-1}}$ and 65,Äì110 $\mathrm{GgS.yr^{-1}}$. These
emissions have been all computed using box models calibrated with ship-
borne measurements made in different parts of the globe
\citep{lennartz_direct_2017,lennartz_monthly_2020}. The $\mathrm{DMS}$
emissions are taken from the \citet{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
$\mathrm{COS}$ emissions are shown in blue whereas the indirect
$\mathrm{COS}$ emissions through $\mathrm{DMS}$ ($\mathrm{CS_2}$) are
depicted in brown (green).}
\label{FF2}
\end{figure}

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

\subsubparagraph{\bf{Biomass burning}}

We use the inventory emissions from \citet{stinecipher_biomass_2019} with
a global estimate of $60\pm37$ $\mathrm{GgS.yr^{-1}}$ for the period
1997-2016. These authors used CO as a reference species to compute the
$\mathrm{COS}$ biomass burning emissions. To do that, they combined
emission factors of $\mathrm{COS}$ 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: $

a supprimé: }$.

a supprimé: }$ and

a supprimé: }}$, respectively.

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 \citet{kettle_global_2002}
inventory, the inventory emissions from \citet{stinecipher_biomass_2019}
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 \citep{campbell_atmospheric_2015}.

\subsubparagraph{\bf{OH sink}}

Since the highest reaction rate is close to the surface
\citep{kettle_global_2002}, we represent the OH sink by a
surface flux. As done in \citet{launois_new_2015}, 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 \citep{kettle_global_2002, berry_coupled_2013-1}. We use
monthly maps of OH radical concentrations from an update of
\citet{hauglustaine_interactive_2004}.

\subsection{Inversion configuration}

\subsubsection{Control vector}

Our inversion window covers 12 years. The spatiotemporal resolution of
the control vector ${\bf x}$ 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 ${\bf x^{b}}$ is simply a vector of ones.

 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\degree 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

a supprimé: one

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 $\mathrm{CO_2}$, we neglect the uncertainty on the oceanic, fire
and anthropogenic $\mathrm{CO_2}$ 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\tnote{\dag} & monthly&
GgS & 36 \\
 Soil  fluxes          & $\mathrm{COS}$ & 15     & 2\tnote{\ddag}  &
monthly      & GgS  & 4140\\
GPP          & $\mathrm{CO_2}$ - $\mathrm{COS}$ & 15    & 2\tnote{\ddag} &
monthly &       GtC   & 4140\\
Respiration &  $\mathrm{CO_2}$        &              15   & 2\tnote{\ddag}
& monthly &       GtC   & 4140
\\
Anthropogenic emissions   & $\mathrm{COS}$ &  0   & 1 & annual & GgS & 1
\\
Biomass burning  emissions &  $\mathrm{COS}$ &   0   & 1 & monthly & GgS
& 12  \\
Background concentration &  $\mathrm{CO_2}$        &             1& 1  & &
ppm  & 1 \\
Background concentration &  $\mathrm{COS}$        &             1   & 1  &
&       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.\\
\tnote{\dag} \footnotesize{The ocean flux is divided into 3 regions :
30\degree N:90\degree N, 90\degree S:30\degree S, 30\degree S:30\degree
N.}\\
 \tnote{\ddag} GPP, respiration and soil fluxes of PFTs 4, 5, 6, 10, 11,
12, 13 are divided into 2 hemispheres : 0\degree N:90\degree N, 0\degree
S:90\degree S.
 \\
}}
\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: O

\label{Error1}
\end{figure}

Observation errors are defined with respect to the observation operator
${\textbf{H}$ and are actually dominated by the errors of ${\textbf{H}$.
As explained in Section 2.3, ${\textbf{H}$ 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
\citet{chevallier_co2_2010} who used the statistics of the difference
between the raw times 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 \citep{sweeney_seasonal_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 \ref{Error1}.
%Another justification to a larger uncertainty is that these four
continental stations are influenced by the diurnal rectifier effect
\citep{denning_latitudinal_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|ccccc}
    {} & {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}
&0.5\\
    {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}
        \caption{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 $\textbf{B}$ (that applies to ${\bf
x^b}$, 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 $\sigma$) 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
$\mathrm{GtC.yr^{-1}}$ for a reference value here of 125
$\mathrm{GtC.yr^{-1}}$: 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 $\mathrm{125\pm 5.2
GtC.yr^{-1}}$ from \citet{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: }$

```latex
    \item Oceanic emissions. Our resulting 1-sigma uncertainty of 350
$\mathrm{GgS.yr^{-1}}$ for the globe and the year, given a reference
value of 271 $\mathrm{GgS.yr^{-1}}$ (see Fig. \ref{Tab1}), is consistent
with
\citet{lennartz_direct_2017,lennartz_influence_2019,lennartz_monthly_2020
} who estimated the ocean emissions between 120 - 600 $\mathrm{GgS.yr^{-
1}}$ .
    \item Anthropogenic emissions. Our correlation length of 500 days
damps interannual variations, consistent with
\citet{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
$\mathrm{GgS.yr^{-1}}$ for the globe and the year, given a reference
value of 370 $\mathrm{GgS.yr^{-1}}$ (see Fig. \ref{Tab1}), is consistent
with the estimation of 223-586 $\mathrm{GgS.yr^{-1}}$ given by
\citet{zumkehr_global_2018}.
    \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   \citep{belviso_top-
down_2020,commane_seasonal_2015}.

\end{itemize}

\subsection{Post-processing of the $\mathrm{CO_{2}}$ and $\mathrm{COS}$
simulations and measurements}

The seasonal cycle is derived from the surface data using the CCGVU curve
fitting procedure developed by \citet{thoning_atmospheric_1989}
(http://www.esrl.noaa.gov/gmd/ccgg/mbl/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, nd a low-pass filter with 80 and 667 days
as short-term and long-term cutoff values, respectively.

\subsection{Metrics}

The simulated atmospheric concentrations (for $\mathrm{CO_2}$ or
$\mathrm{COS}$ here) are evaluated against measurements using the Root
Mean Square Error, RMSE, defined as:
\begin{equation}
RMSE=\sqrt{\frac{\sum_{n=1}^N(C^{Obs} (n)-C^{Mod} (n))^2 }{N}}
\label{ERMSE}
\end{equation}
where $N$ is the number of considered observations, $C^{Obs} (n)$ is the
n$^{th}$ observed concentrations and $C^{Mod} (n)$ is the n$^{th}$
modelled concentration. The unit of RMSE is in ppm (ppt) for
$\mathrm{CO_2}$ ($\mathrm{COS}$).

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

a supprimé: GtC
a supprimé: GtC
a supprimé: GtC
a supprimé: GtC
a supprimé: $
a supprimé: }$
a supprimé: Consistency metrics
a supprimé: %
a supprimé: %\
a supprimé: %
a supprimé: →
a supprimé: %\
a supprimé: %
a supprimé: %modelled concentration. ¶
\label{chi2}

```
\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}.

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).

```
\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
($1-\frac{RMSE_{post}}{RMSE_{prior}}$). Column "$RMSE_{prior}^{seas}$"
presents the RMSE of the a priori detrended time series compared to the
assimilated measurement time series. Column "$RMSE_{post}^{seas}$"
presents the RMSE of the a posteriori detrended time series. Column
"$RE^{seas}$" presents the reduction of uncertainties using the RMSE
metrics applied to the detrended time series ($1-
\frac{RMSE_{post}^{seas}}{RMSE_{prior}^{seas}}$). 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 $\mathrm{CO_2}$
($\mathrm{COS}$). All statistics are for the period  2009-2019.}
\label{Error1}
\end{table}
```

a supprimé: outside the

a supprimé: $\

a supprimé: y

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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ (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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ and COS mixing ratio has been shown to be particularly challenging \citep{ma_inverse_2021}. The inverse system also struggles to match $\mathrm{CO_2}$ measurements at sites WIS, NWR, LEF with a seasonal RMSE greater than 1.5 ppm.

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.

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{chi22}). 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.

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

a supprimé: raw

a supprimé: HFM,

a supprimé: Section

a supprimé: CostF

a supprimé: Section

a supprimé: have

a supprimé: low

improved the simulated concentrations much. However, at NWR, the
inversion struggles to correct the advanced phase, especially in the
$\mathrm{CO_2}$ 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
$\mathrm{CO_2}$ and COS concentrations at three selected
sites, for the a priori and a posteri 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}

\subsection{Optimized fluxes}

Table \ref{Opt1} summarizes our top-down assessment of the COS
and the $\mathrm{CO_2}$ budgets. The inversion doubled the COS
oceanic emissions to 530 $\mathrm{GgS.yr^{-1}}$. 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 $\mathrm{GtC.yr^{-1}}$ of the GPP to match the
COS constraint, the respiration has decreased by 10
$\mathrm{GtC.yr^{-1}}$ in order to keep a land carbon sink in agreement
with the global atmospheric $\mathrm{CO_2}$ 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
$\mathrm{GgS.yr^{-1}}$
\citep{berry_coupled_2013,kuai_estimate_2015,launois_new_2015}. Several
reasons could explain these differences. Firstly, the
\citet{zumkehr_global_2018} anthropogenic emissions are much higher than
the \citet{kettle_global_2002} one used in these previous studies.

Secondly, we assimilated continental surface measurements from the NOAA network through the whole years of 2008-2019 while \citet{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. \citet{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 \citet{launois_new_2015,berry_coupled_2013} requires a larger oceanic source over the tropics to close the COS budget. This is particularly true for \citet{berry_coupled_2013} who used a fixed large biospheric sink of 1100 $\mathrm{GgS.yr^{-1}}$.

\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 COS (left) and the $\mathrm{CO_2}$ (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 $\mathrm{CO_2}$ and COS budgets, as written here, have been obtained by adding all the related optimized parameters (see Table \ref{TPar} for a

---

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

a supprimé: which

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
\citet{berry_coupled_2013} who used a different inverse system and the
\citet{kettle_global_2002} oceanic flux as a prior. Over the tropics,
COS and $\mathrm{CS_2}$ measurements in sea waters do not
support this increase as already mentioned in the introduction
\citep{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 \citet{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 \citep{lennartz_direct_2017}.
This remark supports the inversion decrease of the oceanic emissions over
the mid and high latitudes. The decrease beyond 50\degree 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 \citep{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.

\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 COS budget. The fluxes have been
averaged over the years 2009-2019. }
\label{Opt2}
\end{figure}

\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 $\mathrm{CO_2}$ budget, Figure \ref{Opt2bis}
shows the latitudinal distribution of the net $\mathrm{CO_2}$ 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\degree N. This response is a common
feature of the current inverse systems which, by assimilating
$\mathrm{CO_2}$ 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 \citet{friedlingstein_global_2020},
the net land sink (above 30\degree) 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 $\mathrm{GtC.y^{-1}}$ whereas respiration lost 1
$\mathrm{GtC.yr^{-1}}$, leading to a small source of $\mathrm{CO_2}$. 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 $\mathrm{CO_2}$ absorption within the high latitudes covered
by PFTs 7, 8, 9 and 10 (see Table \ref{TLRU}). Indeed, GPP increased by
almost 2 $\mathrm{GtC.y^{-1}}$ 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
COS and $\mathrm{CO_2}$ 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 an 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

COS 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 $\mathrm{CO_2}$ and
COS 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.

\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|cccccccccccccc}

    & 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{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 an
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. }
\label{Opt3}
\end{figure}

\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 COS

a supprimé: is degraded

a supprimé: the

a supprimé: the

a supprimé: opposition of

a supprimé: the

a supprimé: the

a supprimé: minimum

a supprimé: between

a supprimé: the

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\degree N in DJF and under 50\degree 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\degree.

\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,Äì2012, 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.}
\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 $\mathrm{CO_2}$ 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: weak in

a mis en forme : Français

a supprimé: weaker

a mis en forme : Français

a mis en forme : Français

a supprimé: .[FC: a justifer, pas clair]}

a mis en forme : Français

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 apriori and aposteriori 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
$\mathrm{CO_2}$ has been corrected as a result of higher GPP. The
improvement is also noticeable in the COS and $\mathrm{CO_2}$
vertical profiles over Northern America (see Supplementary material). In
contrast to the \citet{ma_inverse_2021} top-down study, there is no
significant negative bias in the COS vertical profiles here
(see Fig. S7-11).

\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 COS abundance with the HIPPO observations.
Because of the unbalanced prior, the LMDz 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 COS concentration averaged over longitudes
and heights.}
\label{Opt5}
\end{figure}

The optimized COS fluxes are now assessed at three surface
sites in Japan : Miyakojima (MIY; 24\degree80N, 125\degree27E), Yokohama
(YOK; 35\degree 51N, 139\degree48E), and Otaru (OTA; 43\degree14N,
141\degree16E). In winter, these sites sample air masses coming from the
eastern northern edge of China (see \citet{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 \citet{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 \citet{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.

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
$\mathrm{CS_2}$ is emitted as COS. \citet{ma_inverse_2021}
showed that implementing the $\mathrm{CS_2}$ 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 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 COS
observed and modelled COS concentrations, and their standard
deviation at four surface sites: Miyakojima (24\degree80N,

125\degree27E), Yokohama (35\degree51N, 139\degree48E), and Otaru
(43\degree14N, 141\degree16E) and GIF (48\degree42,ÄôN - 2\degree08,ÄôE).
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
\citet{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 COS retrievals from Fourier transform infrared
spectrometry \citep{wang_towards_2016,krysztofiak_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 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
\citet{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

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 \citet{maignan_carbonyl_2020} and \citet{seibt_kinetic_2010} respectively) than in the tropical and temperate ecosystems (around 1.3 and 3 for \citet{maignan_carbonyl_2020} and \citet{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 \citep{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 \citep{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 \citet{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 \citet{belviso_dms_2012} are higher over most of oligotrophic subtropical zones compared to the DMS emissions of \citet{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 \citet{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 $\mathrm{CS_2}$ 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, personnal 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 \citet{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: } \citep{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.

\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 $\mathrm{CS_2}$ oxidation into COS happens instantly in the atmosphere. However, \citet{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 \citet{zumkehr_global_2018} inventory with the help of COS atmospheric measurements. The lifetime of the DMS, $\mathrm{CS_2}$ 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 \citep{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 \citet{stinecipher_biomass_2019} inventory although they were previously estimated to be three times as large as the sources from open burning \citet{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 \citep{belviso_estimate_1986,chiodini_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 \citep{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 \% \citep{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 \citep{yi_exchange_2008}.
\end{itemize}

\section{Conclusion}
We have developed an analytical system that optimizes GPP, plant respiration $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$ 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 $\mathrm{CO_2}$

and COS budgets that give a good fit with assimilated measurements. Inverse results point at a large oceanic COS 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 COS, 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 (\citet{launois_new_2015}: -663-772 $\mathrm{GgS.yr^{-1}}$, \citet{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 $\mathrm{CO_2}$ biospheric fluxes, defined here as the sum of the respiration and GPP, has lost 2 $\mathrm{GtC.yr^{-1}}$ above 30\degree N compared to the prior fluxes simulated by the ORCHIDEE LSM. This behaviour is shared by current inverse systems which infer the net $\mathrm{CO_2}$ fluxes from atmospheric $\mathrm{CO_2}$ measurements \citep{friedlingstein_global_2020}.

Several aspects of the inferred COS 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 COS 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 COS 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 COS 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‚Äôs 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

```
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

%% 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/XwULEXsgwFM1Tus/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 $\mathrm{CS_2}$. 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

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

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

a mis en forme : Français

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

a mis en forme : Français

```
%% REFERENCES

%% The reference list is compiled as follows:
\bibliographystyle{copernicus}
\bibliography{Mabiblio.bib}
%\begin{thebibliography}{}

%\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}
%%
%% 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

%% 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
%
```

a mis en forme : Français

a mis en forme : Français

a mis en forme : Français

```
%%f
%\begin{figure*}[t]
%\includegraphics[width=12cm]{FILE NAME}
%\caption{TEXT}
%\end{figure*}
%
%
%%% TABLES
%%%
%%% The different columns must be seperated 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}
%\tophline
%
%\middlehline
%
%\bottomhline
%\end{tabular}
%\belowtable{} % Table Footnotes
%\end{table}
%
%%% TWO-COLUMN TABLE
%
%%t
%\begin{table*}[t]
%\caption{TEXT}
%\begin{tabular}{column = lcr}
%\tophline
%
%\middlehline
%
%\bottomhline
%\end{tabular}
%\belowtable{} % Table Footnotes
%\end{table*}
%
%%% LANDSCAPE TABLE
%
%%t
%\begin{sidewaystable*}[t]
%\caption{TEXT}
%\begin{tabular}{column = lcr}
%\tophline
%
%\middlehline
%
%\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 mutliplication 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
%%% \leftrightarrow 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