We thank the reviewer for the comments and helpful suggestions. In the following we reply to the individual comments and clarify which modifications were made.

The reply is formatted as: reviewer comment, author response, *revised text.*

Review 1

Linda M.J. Kooijmans and coauthors present an evaluation of the implementation of carbonyl sulphide (COS) fluxes in the terrestrial biosphere model SiB4. One can say without exaggeration that the earlier SiB4 implementation of COS is the reference for all current biosphere models that include COS. Kooijmans et al. present a very thorough evaluation with excellent supplementary information that answered almost all questions that arose while reading the manuscript.

My comments are hence minor.

1. I would disagree with the recommendation 4.1. α is not the only reason that COS fluxes are underestimated at some sites. GPP is also underestimated at DK-Sor and AT-Neu. The seasonal shape of GPP is very different in the model at US-IB2 compared to the estimated GPP from observations. So fitting α seems like a fudge factor. I would not recommend this.

> The reviewer makes a valid point, COS flux simulations are currently still linked to GPP simulations through g_s and V_{max} and therefore the accuracy of COS flux simulations go hand in hand with the accuracy of GPP simulations. We have removed the recommendation to adjust alpha. However, we would still like to keep the part of recommendation 4.1 where we emphasize the different temperature response of COS and CO2 uptake.

2. I also regret the wording in recommendation 4.2. The minimum stomatal conductance is called g_0 in the manuscript. It is not explained how it is used in the model. Stomatal conductance most often depends on net assimilation in conductance formulations such as Ball-Berry and its variants. Net assimilation is negative during dawn and dusk. Is stomatal conductance then set to g_0? This is questionable. If I remember well, Ball et al. (1987) said that g_0 is simply the fitted intercept in the empirical formulation during daytime photosynthesis. It is not the nighttime value. If ever I would recommend to look into such formulations as in Barbour and Buckley (PCE 2007).

> We thank the reviewer for pointing out the differences between g_0 and the nighttime conductance (which we now call g_{dark}). We have indicated these differences now in section 3.1.3 and the recommendation in 4.2. Still, we assume that g_0 is representative of g_{dark}, which is a common assumption in land surface models (e.g. Lombardozzi et al., 2017). This is now also explicitly mentioned in the text.

3. I was missing the explanation/discussion that the authors used reanalysis data to drive the model and not local observations. But especially the discussion of the underestimation of GPP and COS fluxes at DK-Sor literally screamed for it. Would it be possible to redo say SiB4_var_Ogee of Figure 2 using local meteo?

The measurement periods of the COS campaigns are not always covered by the FLUXNET, ICOS or AmeriFlux datasets to be used as meteorological driver input. Therefore, we run the SiB4 simulations with MERRA driver data as it provides consistency in data collection, availability and application across sites. We make a note of this in section 2.1.4.

Still, based on the reviewers' comments we also added a comparison of SiB4 runs with MERRA meteorology and observed site meteorology as driver input for two sites that have more than 10 years of observations and where the site meteorology covers the COS measurement period (FI-HYY and DK-SOR) in the supplement (Fig. S7, and shown below). The comparison for DK-SOR shows that the SiB4 run with site meteorology provides a similar drought anomaly in 2016 as the run with MERRA meteorology. The fact that SiB4 is not able to capture the GPP anomaly is thus not due to the driver data used. This is now mentioned in section 3.1.3.

We do also find that GPP is consistently lower with observed site meteorology due to lower PAR in the observations than in MERRA, leading to a larger model-observation bias with the site meteorology. A reason for the larger bias with site meteorology is that all of the tests, development and tuning for SiB4 are done with MERRA driver files, which might be different for site-level input. We made a note of this in the supplement:

Simulated GPP is consistently lower in SiB4_Obs due to lower radiation in the observations than in MERRA2 (Fig. S7c,g), leading to a larger GPP model-observation bias with SiB4_Obs (Fig. S7a,e). A reason for the larger bias when site meteorology is used is that all of the tests, development and tuning for SiB4 are done with MERRA2 driver files, which might be different for site-level meteorological input. The results in the main text are consistently based on SiB4 runs with MERRA2 driver data, and are not biased by using different meteorology than in the SiB4 development.

We also must note that we found that for a few sites we accidentally selected the wrong year Figs. S10 (Tsoil) and S11 (SWC) (i.e., 2015 instead of 2016 for FI-HYY, DK-SOR, ES-LM1 and US-IB2. This made it seem like the MERRA data did not resemble the observed conditions in Figs. S9 and S10 (SWC). This mistake is now corrected.

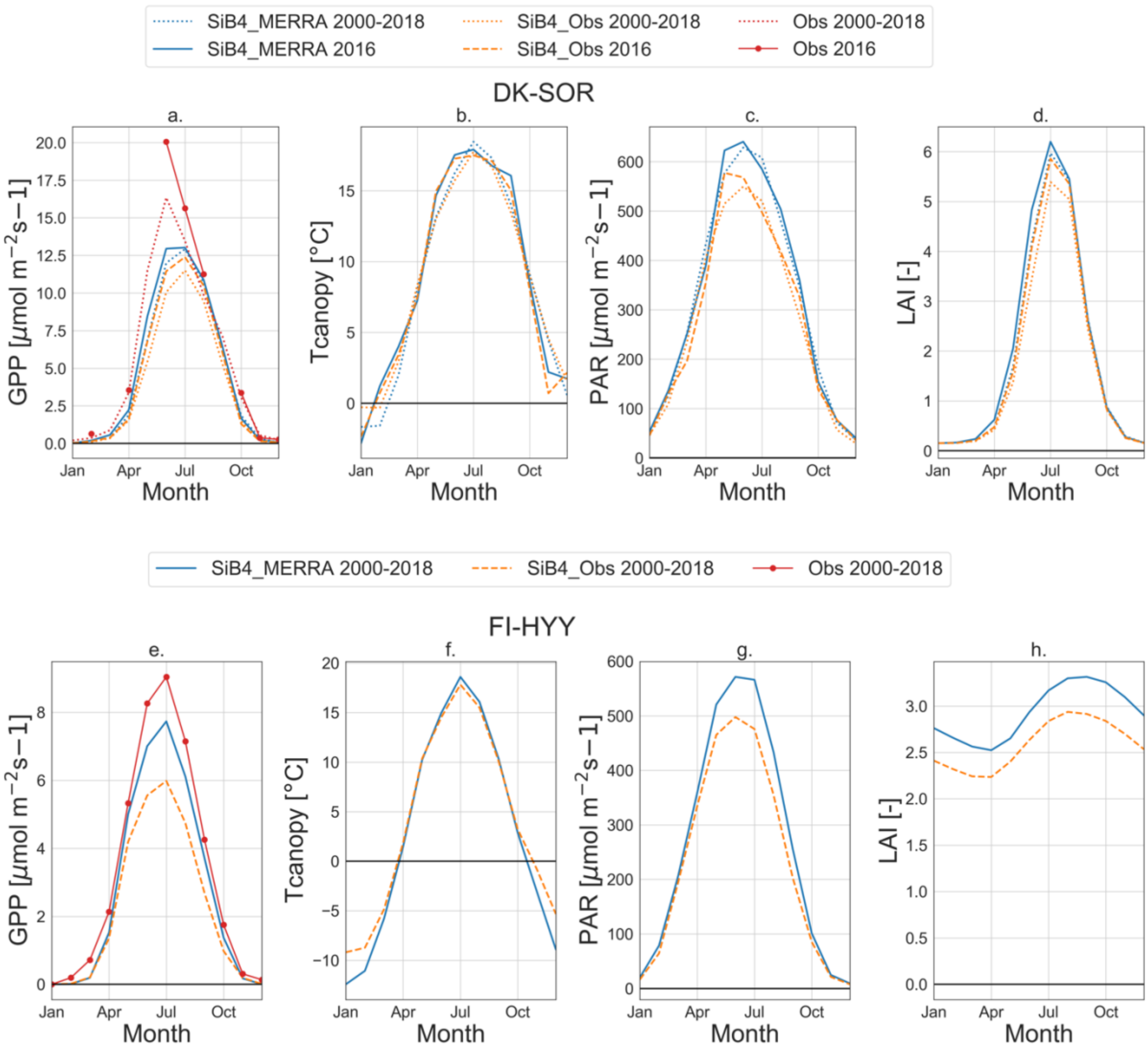

Figure S7: Seasonal cycles of GPP, canopy temperature (Tcanopy), photosynthetically active radiation (PAR) and leaf area index (LAI) at DK-SOR (top) and FI-HYY (bottom) as simulated by SiB4 with MERRA2 driver data (blue) and as simulated by SiB4 with observed site meteorology (orange) and compared with GPP observations from ICOS.

4. I would have loved to see the comparison of SiB4 output with the inverted fields of Ma et al. (2021).

We have added the figure below to the supplementary material and discuss it briefly in section 3.3:

The biosphere flux resulting from inverse modelling by Ma et al. (2021) indicates COS emissions in the Amazon (Fig. S17). While biosphere emissions over the Amazon are unrealistic (Glatthor et al., 2015), it reflects the large missing source in the Tropics (land and ocean) that we are not able to attribute to the biosphere; see a comparison of our SiB4 biosphere flux with the inverted biosphere flux by Ma et al. (2021) in Fig. S17. A potential reason for unrealistic attribution of missing sources of COS is that there are no NOAA observations in the Tropics and its upwind regions to constrain the TM5 inversions.

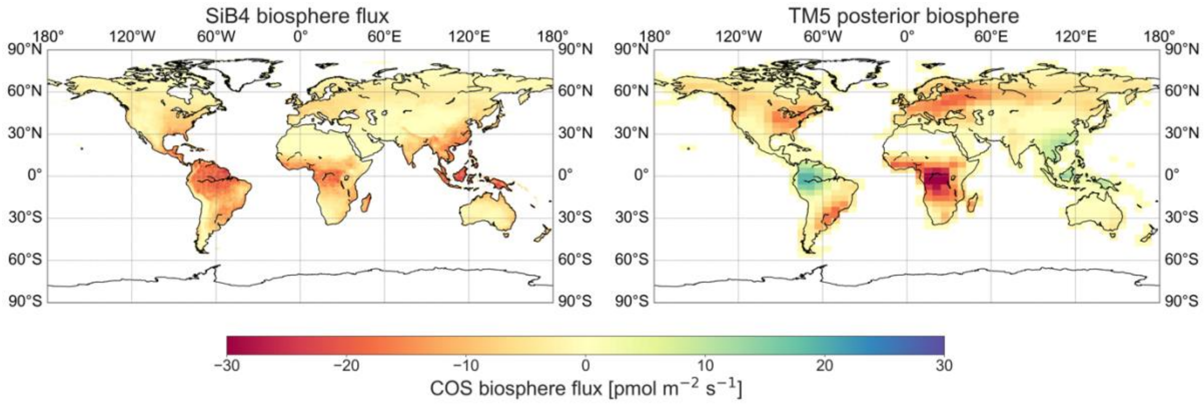

Figure S17. Global distribution of the COS vegetation flux as simulated by SiB4_var_Ogee (left) and the posterior biosphere flux as presented by Ma et al. (2021).

5. I think that Figure 1 is redundant given Table 3 and I would remove it.

We have moved Figure 1 to supplementary information.

6. I found the notation V_{max} pretty unusual. I had to go back several times to Equations 2 and 3 to check the definition. I would recommend to use something like $V_{c,max25}$, which is pretty standard and tells the important information, i.e. it is for carboxylation and at 25 °C.

Corrected as suggested

References

Glatthor, N., Höpfner, M., Baker, I. T., Berry, J., Campbell, J. E., Kawa, S. R., Krysztofiak, G., Leyser, A., Sinnhuber, B. M., Stiller, G. P., Stinecipher, J. and Von Clarmann, T.: Tropical sources and sinks of carbonyl sulfide observed from space, Geophys. Res. Lett., 42(22), 10082–10090, doi:10.1002/2015GL066293, 2015.