Answer to Referee report #2

We thank the referee for valuable comments and suggestions. Our answers are given below. The original referee comment is repeated in **bold**, changes in the manuscript text are printed in *italics*.

In this study, the authors presented detailed analyses of five chemical species (PAN, C_2H_6, HCOOH, CH_3OH and C_2H_4) measured by the Gimballed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA) instrument during the Transport and Composition in the Southern Hemisphere Upper Troposphere/Lower Stratosphere campaign (SouthTRAC) conducted in over the South Atlantic in September-October 2019. In addition to the in-situ measurements, a back trajectory model (HYSPLIT) is used to examine the origins of the pollutants. The Copernicus Atmosphere Monitoring Service (CAMS) model simulations are also used to examine the transport pathways. The enhancements in those five chemical species, which were captured during each flight were found to have varying degree of agreement with the CAMS model results. This study presents a compelling result by utilizing a valuable set of data and the global and trajectory models. I would like to suggest a few minor changes which might add richness to this work.

We thank the referee for the encouraging statement.

General Comments:

- **I would like to suggest adding a little more background on the five chemical species chosen in this work. What do they have in common? Why were those selected? How much understanding do the community has in terms of their sources, sinks and their chemical lifetime?**
  We thank the referee for this suggestion. In the introduction, we added a statement about the selection and the commonalities of these specific species (see answer to the specific point below). Further, we now discuss sources, sinks and lifetimes (which are all presented in Tab. 1) in the introduction, too.

- **Adding some information about the measurements of those species by satellites would be helpful, if possible. Are there any references comparing the satellite measurements and the model simulations? Do other models have difficulty simulating those species accurately? Adding a few relevant references would help understanding the general aspect of those species.**
  In the introduction, we mention satellite measurements in nadir and limb geometry for the discussed species. We now extended this part of the introduction with examples for usage of satellite data in modelling studies.

- **Does the CAMS model perform well in general? I would like to see a statement about why the CAMS model is used here. Is the goal to evaluate the model or to improve the model? If the improvement is the goal, a more specific direction would be needed possibly in conclusion.**
  In the discussion of Sec. 5, and in the conclusions, we write that CAMS performs well for the species PAN, and we address issues with the other trace gases, together with educated guesses why these other trace gases perform not as good as PAN.

In the introduction, we added a statement why the CAMS model is chosen for this study, and the goal of this study: *The CAMS reanalysis uses a state-of-the-art atmospheric chemistry model for data assimilation, which is publicly available and widely used for air quality and pollution related studies (e.g., studies citing Inness et al., 2019). In this work, we aim to evaluate the CAMS reanalysis in the remote upper troposphere above the South Atlantic, a sparsely measured region. With our comparisons we further aim to give recommendations for improving the CAMS model with respect to emissions and atmospheric lifetimes for the studied species.*

- **It is stated throughout the study that the degree of agreement between the measurements and the model varies depending on the species. I would like to suggest adding more thoughts or references to make the findings valuable. If the agreement is not good, how can we improve it in the future?**
  We extended the relevant parts of the manuscript according to the specific points raised by the referee below. For the improvement of CAMS, we make suggestions (variation of emission strengths in the emission inventory, review of atmospheric gain or loss processes).

Specific Comments:

P1, L14: Are PAN, C_2H_6 and HCOOH longer-lived than CH_3OH and C_2H_4? I am curious why the agreement between the measurements and the model is better for PAN only.
As shown in Tab. 1 later in the manuscript, PAN, C_2H_6 and HCOOH have upper tropospheric lifetimes longer than weeks, while CH_3OH and C_2H_4 rather have lifetimes of few days. As explanation why the agreement between the measurements and the model is better for PAN only, we suggest later in the abstract *model deficiencies in the representation of loss processes and emission strength.*

P2, L23 & 24: I recommend listing examples of 'some of these traces gases' and 'some pollution trace gases' here.
We added carbon monoxide and nitrogen dioxide as examples for ozone precursor biomass burning gases, together with additional references (Bozem et al., 2017; Bourgeois et al., 2021). Further, we name now VOCs for aerosol formation, because the contribution of single trace gases to secondary aerosol formation are still subject to current research. Again, we give additional references (Hobbs et al., 2003; Akherati et al. ,2020).

P2, L28: It would be helpful to add a reference at the end of this sentence or rephrase this as 'their potential influence on climate may increase over time'.
We adapted your suggested change together with suggestions from the other referee.

P2, L29: I recommend making changes to this sentence. For instance, 'and may have other sources in addition to pyrogenic emissions.
We changed this sentence according to the other referee's suggestion.

P2, L30: Why those five gases were chosen and what do they have in common?
We added to the manuscript: *These trace gases have been selected for this study, because they all are potentially emitted from biomass burning events, because they have a large range of upper tropospheric lifetimes (from a few days to several months), and*

because they are part of the GLORIA (Gimballed Limb Observer for Radiance Imaging of the Atmosphere) measurements and of the CAMS (Copernicus Atmosphere Monitoring Service) model output. In addition, these trace gases are measured by various infra red satellite sounders (see below) but in coarser spatial resolution than the GLORIA measurements.

P2, L36: 'Filamentary structures' have been mentioned throughout the manuscript. It would be helpful to have a definition or description of it.
We clarified that we mean mesoscale structures with horizontal extension of up to hundreds of kilometers by the term 'filamentary structures'.

P2, L37: I recommend modifying the sentence 'Biomass burning events are typically represented by emission data sets in atmospheric models'. I think emission inventories are one of the factors determining how the model represents the biomass burning events. In fact, emissions, chemistry, and transport all make contributions to the model performance.
We clarified the beginning of this paragraph: *Atmospheric model simulation of such pollution trace gases is challenging: For good model performance, knowledge about pollutant emissions, chemistry and transport are necessary. Location, time and emitted species of biomass burning events are typically represented by emission data sets in atmospheric models.*

P2, L43: Adding more explanation about 'atmospheric processes' would be useful here. Does this refer to a chemical reaction or a physical process?
We changed 'atmospheric processes' to 'chemical reactions and physical processes' and give now an additional example.

P2, L48: Is there a website or a reference for the SouthTRAC campaign?
We added a link to the SouthTRAC website. An overview paper is only available for the gravity wave part of the campaign (which is not the focus of this paper). This reference (Rapp et al., 2021) is cited in Section 2.1, where the flight campaign is introduced in more detail.

P6, L118: Does this mean that only the horizontal motions will be analyzed here? Can we still trust the horizontal motions from the trajectories when the vertical motion is not accurate?
Now, we clarify: *For this reason, the vertical motion of the HYSPLIT trajectories is not discussed in detail here , and it is not tried to retrieve the origin of the measured air masses, but rather the location, at which the air masses reached upper tropospheric altitudes.*
The original formulation that the horizontal motions are not analyzed may have been misleading. However, the vertical motion is limited by the meteorological fields, as mentioned earlier in this section. In particular, fast upward transport may be not represented in the meteorological fields. This means, that the air masses of the trajectories may have entered upper tropospheric altitudes at any point along the trajectory. In section 4, we discuss the trajectories very carefully due to this limitation in the vertical transport. According to the comment of referee 1, we further now avoid the formulation "origin of air masses" and rather speak of the "location at which the air mass reached the upper troposphere".

P6, Section 3.1: It would be necessary to include references for FIRMS, MODIS and ERA5 in this section.

We added the FIRMS website, and these references for MODIS and ERA5:
* Giglio, L.: MODIS/Aqua+Terra Thermal Anomalies/Fire locations 1km FIRMS V006 and V0061 (Vector data), 10.5067/FIRMS/MODIS/MCD14ML, NASA EarthData, 2000.
* Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on pressure levels from 1979 to present., 10.24381/cds.bd0915c6, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2018.

P7, Figure 1: This is a very nice set of figures. However, the boxes with various colors make the figure a bit complicated. It would be helpful to add the names of the gases where the maximum exists. For instance, add 'C2H6' in the pink box in Fig. 1c. This can also be considered for Fig. 2.
We thank the referee for this suggestion, but we are not sure how to adapt this suggestion into the figures. For the magenta box, it would be easy to call it the "C_2H_6 box", but for the yellow box, it could be called "PAN box", "HCOOH box" or "CH_3OH box" as all of those gases have strong maxima within this yellow box. In addition, there are more boxes than presented trace gases, so it is not possible to find a unique name for each box, if you are limited to the names of the trace gases.

P10, Figs. 3a & 3c: It would be useful to mark the initialization locations in these plots. For instance, add larger dots on the location with the same color with the trajectories.
We agree with the referee that it is not easy to see where the trajectories start. Unfortunately, marking the starting points with larger dots in the same color does not help to make the already busy plot easier to read. Instead we added all tangent points along the flight track, color coded with the grey scale color bar from Figs. 4-5 and changed the figure caption accordingly.

P11, L197: Have there been any studies showing the CAMS performance on simulating PAN?
As noted in Sec. 2.2, Wang et al. (2020) compared CAMS reanalysis PAN to aircraft measurements over the Arctic, North America and Hawaii and found an agreement between model and measurements. For higher altitudes, Wetzel et al. (2021) and Johansson et al. (2020) indicate an underestimation of PAN by the model above the North Atlantic, and within the Asian Monsoon, respectively. Further, within the Asian Monsoon, it is suggested that emission sources are missing. However, for CAMS surface PAN, the Wang et al. (2020) reference indicates a good performance.
We add those two references (Wetzel et al. (2021) and Johansson et al. (2020)) to Sec. 2.2, and add to Sect. 4:
CAMS surface PAN was shown to agree with measurements elsewhere (see Sect. 2.2 and Wang et al., 2020).

P13, Section 5.1: It would be helpful to add some insights on the different degree of agreement between the measurements and the model depending on each species. Is it related to lifetime of the species? Or surface emissions? Why does the model overestimate CH_3OH?
We now refer in the beginning of section 5 to the end of the section, where the different degree of agreement between measurements and model depending on the species are

discussed. Further, we extended this discussion subsection, according to the comment below.

P15, L287: This is one of the most important findings in this work. I would recommend spending more time on the discussion. Are the sources of C_2H_6 and HCOOH underestimated in the models and well known? If CH_3OH and C_2H_4 are overestimated in the model, could that be related to the surface emissions only? A few references on this subject might be useful to include here.
We extended the discussion as recommended by the referee. Now, we explain in more detail the influence of emission strength from the GFAS emission inventory, and we discuss the influence of too weak or too strong atmospheric loss processes to simulated VMRs.

P16, L314: Does 'which has been also observed' refer to the underestimation of C_2H_6 in the Northern Hemisphere as well?
We split this sentence into two, in order to make it clearer: *Structures of CAMS C_2H_6 are overall in agreement with GLORIA for both discussed research flights, but absolute VMRs are underestimated. This underestimation of C_2H_6 has also been observed in the northern hemisphere by Wang et al. (2020).*

P16, L319: It would be helpful to add a sentence after this. Could this overestimation be related to overestimation of surface emissions or missing sink reactions? Or could this mean that the lifetime estimation is inaccurate?
We extended this part according to the referee's suggestion: *CH_3OH instead is overestimated by the model, for both, peak and background VMRs. Structures measured by GLORIA are, however, reproduced by CAMS. This indicates that surface emission locations are simulated correctly, but emission strength might be overestimated, or a missing or underestimated atmospheric sink may cause this simulated overestimation in CH_3OH. Such missing atmospheric sink would also influence the estimated atmospheric lifetime.*

P16, L326: This paragraph discusses a very important point. I would recommend adding a bit more specific information about the emission inventories. For instance, adding a few different emission inventories and discuss how they underestimate or overestimate specific species might give clearer idea about the future improvements. The current paragraph discusses this issue as a general issue but not specific to this study.
We have extended the discussion part of Sect. 5, according to a previous comment, to discuss the influence of the emission inventory in more detail. Further, we extended the conclusions, according to the referee comment.