Referee 1

General

This manuscript represents the latest in a series of interesting studies based on research flights above Russian source regions. It presents a complex dataset consisting of many parameters, but rightly focusses on the vertical profiling at key localities with expected variation due to local sources or long-range transport. It is an appropriate contribution to the special issue.

Some sections lack precision and contain statements that are too vague and need constraining. Often too much conjecture without supporting evidence.

Regarding unclear sections or vague statements, we had two lines of action:

1) The comments from referee 1 indicate directly several points of analysis requiring clarification. They have all been clarified one by one as described in the rest of this response. We give here a short overview of these points and will repeat the explanation in the rest of the text:

- Better defining the use of Mace Head as a reference station for this study with a selection of only marine sector measurements.
- A better definition of the background.
- Distinguishing between "reference" (stations used for comparing the measurements of the campaign) and "background" (value subtracted from measurements for comparison with simulated enhancements).
- Giving more evidences from previous campaigns on the CO_2 uptake by Siberian ecosystems observed during this campaign.

2) We also clarify certain points not specifically pointed by the reviewer 1:

- Results Section is better internally structured. We move from general statistics to average vertical profiles to individual vertical profiles to finally a comparison between measurements and simulated enhancements. Each sub-section is better introduced and provides a rational link to the following section.
- As suggested by the reviewer, the high variability for both CO_2 and CH_4 in certain flights observed in Sect. 3.1 (statistics) is now better explained. We mobilize data analysis from previous campaigns over Siberia and from the literature. We particularly focused our analysis on two type of flights: 1) both CO_2 and CH_4 with high variability and 2) and only CH_4 present high variability during the flight. This section is now focused on the description of the flight by flight variability.
- In Sect. 3.2 (average vertical profiles), more information on gradients of western CH_4 vertical profiles are given to support the presence of emissions by local wetlands.
- Also in Sect. 3.2, more information on vertical gradients of western CO_2 vertical profiles are given to support the depletion of CO_2 at lower altitudes by local ecosystems.

The surveys were in September. It is not clear why early September was chosen so some clarification is needed, but as a period of minimum sea ice and maximum thaw it makes sense to target this period, although not as a period of maximum emissions.

Concerning the period chosen: when conceiving the campaign, it was decided that the measurements would take place during the summer when the ocean is open. Now, having precisely September as a month for the campaign just results from logistical constrains. We acknowledge that September is not the period with maximum emissions but it remains still within the initial requirements. It is also interesting for having minimum sea ice and maximum thaw. This precision is added in the new version of the manuscript as "The period of the campaign was chosen to be during summer to have measurements when the ocean is open. While September is not a period of maximum emissions, it presents minimum sea ice (Fig. A1) and maximum thaw.".

Have the authors considered increased shipping emissions at this time, or make a statement that the inventory may now need to include such sources.

EDGAR inventory does have a category for shipping emissions that we include in our "fossil fuel" category. North of 60°N, shipping represent 1.5% of our "fossil fuel" category. Shipping emissions north of 60°N are increasing by 0.5% over the period 2016-2021 (EDGAR v7). This sector has monthly variability in the EDGAR version that we use here (v6, until 2018). We find that emissions are roughly constant over the year, the only seasonal feature being a limited decrease of 15% in December-January.

It is not clear why Mace head is chosen as a comparative site for this study. What is it supposed to show? Using the monthly average puts the value in the middle of the observed flight ranges. If the MHD background sector only had been averaged then the averages would have been closer to those at Mauna Loa and closer to the baseline for observations (higher CO_2 and lower CH_4). MHD can be a good background for air masses crossing Europe at 50-70°N, but only for air from the westerly sectors.

MHD is used a reference for background composition of incoming air masses. We agree with the referee that MHD marine background is appropriate. For the new version of the manuscript, we use the MHD marine background sector only as a reference. Following Biraud et al. (2000) the marine sector is a selection of data based on standard deviation, wind speed and wind direction to only have air from the westerly sector. In September 2020, 25% of hourly data are both within the marine sector and flagged as valid. This marine background selection results in a decrease of both average mixing ratio of CH_4 and CO_2 by 25 ppb and 1 ppm respectively.

The manuscript is well-written and clear for the most part, but there are many small scientific points requiring clarification outlined below, and some additional typographic errors highlighted directly on the attached manuscript. There are a few very minor corrections needed to the English that can be easily sorted out, with adjacent words ending in s. Either the first will end in s or the second in s depending if the phrase is referring to singular or plural observations. One such phrase has been highlighted.

Manuscript is revised for this kind of mistake. Thank you for the global feedback that is precious for the study.

Detailed Points

Line 31-32 – 'rise of 1.09 °C in the last decade (2011-2020) compared to the last decades of the 19th century (1850-1900)'. This iss confusing - do you mean that the average temperature for 2011-2020 was 1.09°C higher than the average for 1850-1900?

This is what we meant, this is rephrased in the new version as "In the last decade (2011-2020), global surface temperature was 1.09 °C higher than the last decades of the 19^{th} century (1850-1900)".

Line 44 – '(178 to 206 Tg CH4 yr-1)'. Exactly same range as agriculture - one is incorrect

Good ranges are updated in the new version as "Anthropogenic emissions represent 60% of global total methane emissions with agriculture and waste management (191 to 223 Tg CH_4 yr^{-1}) and fossil fuels exploitation (113 to 154 Tg CH_4 yr^{-1})".

Line 89 – 'specific signatures'. What is meant by this?

This sentence is rephrased in the new version for more clarity as: "We then focus on four individual vertical profiles to characterize atmospheric transport of three different regions.".

Table 1 – The UTC dates are a little confusing. Seeing the maps first it seems like there are 2 flights on 2 separate dates, rather than one flight in the middle of the day local time.

Table 1 presents both the local time and UTC for each flight. For the map we now display only the starting date of the flight to avoid the potential confusion pointed out by the reviewer. We think however it is important to keep UTC time as some flights crossed different time zones (especially as a reference to ulterior figures with time as parameter).

Figure 2 – The Picarro instrument rack appears to be installed below flight deck level, below which there is a label S, that is not explained in the caption. Anything specific about the positioning as this is the key instrument for this manuscript and is too new to be explained in the 2011 paper on plane layout.

The original figure used in the article was from a campaign planning document. Belan et al. (2022) provides an updated figure of the aircraft configuration. The Picarro is on the main deck and not below deck. We reproduce the figure from Belan et al. 2022 in the updated manuscript.

Line 139 – 'a ridge coming from'. Maybe should emphasize for the non-meteorologists that this is high pressure.

Precision is added as "and a high-pressure ridge coming from".

Line 144 – '60% of the measurements were taken above the Arctic Circle (> 66° N)'. Move this sentence elsewhere. It is splitting discussion of vertical measurements.

Sentence is moved at the end of the sub-section line 157.

2.2 Instrumentation – For the Picarro there is no mention of the pumping speed or the delay time between the air entering the 3m tube and the measurement. The offset could be important during vertical profiling through stratified air masses, if origins are different.

The maximum vertical offset induced by changing air residence in the tubing is 23.4m, calculated as follows. The length of the tube was 4.8m and the actual flow rate is 1.0 L/min. The tube used is synflex 1/4" so the inside diameter is about 4 mm. Air velocity is about 1.33 m/s so the residence time is 3.6 s. The aircraft vertical speed is 6.5 m/s so the offset is about 23.4 m. In our paper, we show fully resolved vertical profiles (1-second data) only in Fig 6. Then our study uses vertical profiles average in 500 m bins and, for the last part, measurements comparisons with simulated mixing ratios are 1-minute averages. Therefore we assess that this offset does not influence our results.

Line 160-161 – 'temperature (up to 15 °C hr-1)'. Seems to be a strange value. The aircraft can ascend to 10000m in 10 minutes during which time T will change by much more than 15°.

We were here talking of the ambient temperature variation around the sensor (which is below 15 °C hr^{-1}). This sentence is based on manufacturer fact sheet. As this information can be misleading and doesn't bring useful information, it will be removed in the revised manuscript.

Line 208-209 - 'minimal mole fraction for each flight '. Not clear what this is. Is this a single value? Does the minimum not change with time of day and the extent of mixing at lower altitudes? What happens if the flight passes through a stratospheric wave at high altitude and records a lower mole fraction. Needs some more explanation.

The referee points out the lack of clarity in the definition of the background, suggesting to revise and better define our backgrounds. For the updated version of the manuscript, we define the background as follows: 1) we only keep CH_4 values in air that have a corresponding O_3 mixing ratio < 70 ppb, in this way we eliminate as much as possible influences from stratospheric air. 2) Then we take the 5th percentile as the background value. (Previously we used minimum mixing ratio.) Background mixing ratio are calculated flight by flight as previously. This background is used in Section 3.4 where it is subtracted from measurements to calculate measured enhancements.

More generally, we now better differentiate in the manuscript between "background" (as defined here) and "reference" measurements obtained from remote stations such as MHD and Mauna Loa.

2.4 Methane flux inventories – Why is each flux category sentence a separate paragraph? Would be clearer as a list.

It is updated as a list in the new version (see lines 219 to 238)

Line 238 – 'analysed separately'. Does this include a different selection of background mole fraction?

This does not include different backgrounds. Actually, in this Section 3.1 we do not remove "background" mixing ratio from measurements, we simply compare our measurements to "reference" measurements in remote stations. As discussed above, we now clarify this by consistently differentiating between "reference" and "background" measurements.

Line 244 – 'Mace Head (1979 ppb).' Is this a monthly average? Mace head has a well-defined background sector. If there is high pressure bringing continental European emissions to the site the monthly average will be significantly above the Atlantic background.

It was indeed the monthly average but without filtering the air masses coming to the station. The new value for marine sector only is updated in the new version as "1956 ppb" (for additional explanations, please refer to the general comment above on Mace Head).

Figure 4 - Very low mole fractions of CH_4 on 2 flights - see earlier comment about stratospheric air.

A statement about stratospheric air will is added. Saying: "Flights T1 and T6 registered the lowest CH_4 mixing ratios of the campaign (respectively 1812 ppb and 1777 ppb with an aircraft altitude above 10000 m), coming from stratospheric air masses which results in strong gradients for these two flights.".

Lower CO2 at Mace Head than flights or Mauna Loa. Is biospheric uptake due to continental outflow also the reason.

This suggest the following possibility that we do not investigate in detail in the paper: air outflowing from Eurasia to the Pacific without much further biosphere/atmosphere exchange (possibly through the Arctic) ensure that flights and MLO have comparable CO_2 concentration. The excess CO_2 compared to MHD could then arise from loading airmasses with anthropogenic CO_2 as they pass over populated Europe and western Russia (as suggested by the shown back trajectories, see e.g. fig 10).

Another possible explanation is that the marine sector selection being based only on wind at the time and location of the measurement, may have been significantly influenced by continental uptake and is therefore below the hemispheric reference value at MLO.

Lines 266-269 - Need to explain the two flights with large CH4 variability in the CH4 section 3.1.1

We revise our analysis on high CH_4 flights. We now further discuss the flights with the highest 3rd quartile (western and eastern Arctic coasts). Lines 254 to 256 now read : "Flights in the western Siberian Arctic have the highest 3rd quartile of the campaign (L1, L2 and T2) then followed by the flights in the eastern Arctic (L3, L4.2 and L4.3). Large mixing ratio are expected to be encountered in the boundary layer, linked to regional sources, these will be discussed in the Sect. 3.2 by using vertical profiles."

Lines 298-299 – 'Eastern flights present an enhancement of +10 ppb to +20 ppb compared to Mauna Loa and Barrow mixing ratios respectively at 35 ppb and 30 ppb for the same altitude range'. Not clear. Can probably rephrase as 'compared to respective Mauna Loa and barrow mixing ratios of 35 and 30 ppb.

It is rephrased correctly as mentioned above.

Lines 308-309 - You are using the CO data that has very big errors, without showing the errors. Is 4 ppb significant if the instrument was not working well?

The phrasing in the submitted manuscript was misleading and we apologize for that. We are not comparing the two CO values (that corresponds to this 4 ppb difference the referee mentioned) in the manuscript. It is rephrased as "Highest average mixing ratios at low altitudes for CO are also measured during flights T6 and T7 with similar values of 103 ppb and 107 ppb respectively."

Section 3.3 – Is it possible to defined the boundary layer height for these flights and on

Figure 6. The CO_2 and CH_4 profiles give a strong indication of where they are likely to be, but were sondes released from the aircraft to give this information and confirm the height. Currently the BL information is quite general.

Unfortunately no, the plane didn't have any dedicated sensors that were measuring BL height. We can only rely on meteorological and composition sensors.

Fig 6a could have the same CO_2 scale labels as b and d given that the range is the same.

Figure is simplified.

Figure 7 caption – Need to specify if local or UTC times and be consistent in usage throughout.

It is UTC, the precision is be added.

Line 352 – 'coming from the BL'. What evidence is there for this coming from the BL when figure 7 shows total column footprints?

The simulated footprints can also be represented with a threshold altitude of 500 m. This shows air masses above Northern America indicating the origin from the BL. This useful figure is added in appendix C. We add in the text lines 387 to 391: "This O_3 depletion mentioned in Sect 3.2. (-24 ppb compared to immediately lower layers) occurs in the air mass whose history is shown in Fig 7b. The same backward transport simulation with a threshold altitude of 500 m (appendix Fig. C1) shows an air mass coming from the BL in Northern America then crossing the Atlantic Ocean and Northern Europe (only visible on the total column footprint).".

Figure 8 - Both a and c have the vertical axes labelled as CH_4 when it should be CO.

Axis labels is corrected.

Lines 383-384 – 'and the closest one being Kenai Alaska liquid natural gas terminal'. Are the trajectories / particle dispersion for this flight time appropriate for a contribution from this source?

We don't have any temporal information on the emissions of the LNG terminal, we just know that trajectories crossed it. In absence of information on the intermittency of the emissions, we considered the emissions as steady.

Line 390 – 'depending on the geographical by source type and location'. Something missing.

It is corrected in the next version as "3.4.1 Contributions to measured CH_4 by source type and by location"

Figure 9 – It would be useful to have the 1:1 line on the graph to highlight the underestimations of the model.

The 1:1 line is added in the next version.

Figures 10 and 11 - The altitude is shown by the grey line but there is no scale from it. We can make presumptions from previous figures, but it needs some indication in the caption, e.g. bottom of graph = 0m, top = 10,000m, or similar.

It is corrected in the next version as "(bottom at 12 m and top at 8963 m)" (this quote refers to Fig. 10, same type of indication for Fig. 11).

Line 535 and associated earlier results discussion - Need to clarify latitudes for ground CO2 sinks in September. Most studies of the boreal and Arctic zone CO2 suggest that for latitudes >50°N the vegetation is a net source of CO2 by September, not a sink (e.g. Welp et al., ACP, 2016)

We argue that in September, Siberian ecosystems (>50°N) can be a sink for CO_2 especially in western regions. Previous YAK-Aerosib campaigns realized in late-August/early-September present vertical profiles with a low CO_2 concentration close to the ground jointly with a positive strong vertical gradient when altitude is increasing. In Paris et al. (Bams, 2010 https://doi.org/10.1175/2009BAMS2663.1), the panels (b) and (c) of Figure 5 present such vertical profiles for the most western flights at this time of the year (i.e. flights 5, 8, 9 and 12). Paris et al. (ACP, 2010, https://doi.org/10.5194/acp-10-1671-2010) demonstrated based on the relation between measured CO_2 concentration and airmasses' residence time in the lowest 300 m that for September campaign, the longer the airmasses resided over local areas (boreal and sub-arctic Siberia, >50°N), the stronger the CO_2 uptake by Siberian ecosystems (see Fig. 12, dark blue triangles for YAK2 Local).

Overall – check for consistency of use of mixing ratio. I think I saw at least one mention of mole fraction.

Only "mixing ratio" is kept for the consistency.

Reference List – Structure and ordering looks good, but I did not check for consistency with text citations.

Supplementary Information – This is useful. Is Figure B1 discussed in the text? This comes back to my earlier comment and the increased utilisation of this route by shipping at this time of year and the associated emissions.

Figure B1 is discussed several times in the original time. For the next version, a new mention to it is added earlier in the Sect. 2.1 Campaign description lines 102 to 104. The position of the figure is updated subconsequently. Further explanations can be found above in the general comments about “the period of the campaign” and “the shipping emissions”.

Referee 2

General comments

The paper tries to capture CH4 and CO2 characteristics in Russia with several observation flights in September 2020. Using tracer gases (CO and O3) and footprint analysis, the authors investigated the cause of variation for the gases. Observation over Eurasia is limited, particularly aircraft observation; thus, the data is valuable for this research field.

However, although the data could be used to validate any transport model output, a snapshot of one month's data is difficult for meaningful analysis for GHG research.

I need clarification on the new findings. In some chapters, statements are too ambiguous and contain too many guesses without probable evidence.

Besides the validation of models, inversions and satellite retrievals, large scale aircraft campaigns have demonstrated their ability to lead to new meaningful information about atmospheric composition and surface atmosphere exchange. We would argue that they trade their relatively short duration for an extended 3D spatial coverage combined with high precision instrumentation.

To clarify our new findings, our analysis of the campaign data revealed that:

- CO_2 and CH_4 show a large variability over Siberia, especially in the lower and upper troposphere.
- A strong large scale vertical gradient, extending far from the source area, is linked to strong wetland CH_4 emissions and CO_2 biospheric uptake in Western Siberia.
- Long range transport from Europe and North America plays a significant role in mixing ratio variability.
- Model simulation generally underestimate observed CH_4. Oil and gas emissions dominate regional anthropogenic CH_4 emissions but their emission inventory is likely underestimated.
- Throughout our northernmost flights, our data is compatible with Arctic Marine CH_4 emissions equal to or higher than the Weber et al., 2019 inventory.

In addressing both reviewers' comments, we carefully reworked our manuscript to eliminate all ambiguous statements and to provide more evidence supporting our conclusions. Concerning the clarity, many general and precise points have been improved following suggestions by Reviewer 1 (see response to Reviewer 1). Many statements have benefited from more elaborate discussions such as the study on CO_2 and CH_4 variability, the origin of regional emissions or the CO_2 uptake by regional ecosystems. Finally, we also clarify guesses pointed out by the referee such as on the CO_2 enhancement at low altitude close to Bering Strait of the absence of wildfires in the contributions of the simulated CH_4 enhancements.

Specific comments

L378-379: Then, what is the reason for CO2 enhancement?

The reason for CO_2 enhancement is now given as follows "Figure 8c shows a scatter plot of CO against CO_2 for the whole flight L4.1 (including all six profiles). The correlation coefficient under 2000 m is significant but slightly lower (R2 = 0.33, p = 10-11) than correlation coefficients between CO_2 and CH_4 shown on Fig. 8d (R2 = 0.64, p = 10-37), suggesting that the CO_2 enhancement may be driven by a mixture of emissions including local combustion processes.". Further explanations are given in the rest of Sect. 3.3.3, they will be detailed in the next comment just below.

L379-380: Any explanation is needed for the different tendencies between low and high.

For the correlation between CO_2 and CH_4 at low altitude, the explanation already present in the submitted manuscript is now broadened, see lines 420 to 429. As the anti-correlation between CO_2 and CH_4 at high altitude, it is now explained in two separate sections: 1) in Sect. 3.3.3 lines 430-435 as "The anti-correlation between CO_2 and CH_4 above 2000 m (Fig. 8d) could be a residual from active sources for CH_4 and active sinks for CO_2 during the summer. While the air in this region has been relatively well-mixed as shown by the flat vertical profiles of CO_2 and CH_4 for flights L4.1, L4.2 and L4.3 (Fig. 5) and the reduced value ranges for CO_2 and CH_4 on the scatter plot of Fig. 8d (compared to the value ranges on Fig. 8b), mixing of airmasses affected simultaneously by CO_2 sinks and CH_4 sources in the past might be incomplete resulting in this residual anti-correlation. This will be better investigated in the Sect. 3.4.1 by using the simulated mixing ratios for flight L4.1." and in Sect. 3.4.1 lines 493 to 496 as "Overall at altitudes above 2000 m, most contributions are missed by the model showing that the air in the free troposphere has not been into contact with the surface for more than 10 days. This supports the anti-correlation observed in the Sect. 3.3.3 between CO_2 and CH_4 for flight 3.1 that is a residual of the active sources and sinks during the past summer."

Overall, we want to bring to the attention of the editor and the referee, that we largely reworked on the Sect. 3.3.3 that was called into question in this comment and the previous one to bring a finer analysis with more evidences.

L380: What is the meaning of 8.5 ppb CH4/ ppm CO2?

It means that the CH_4 concentration increased by 8.5ppb for an increase of 1ppm for CO_2. It is phrased more explicitly in the new version as "The regression slope is 8.5 ppb of CH_4 per ppm of CO_2.".

L458: Please rephrase the sentence.

Sentence is rephrased as "Over these two flights, we can observe that the peak-to-peak amplitude of simulated mixing ratios is higher than the peak-to-peak amplitude of measured mixing ratios indicating that some contributions are likely missing in our model (e.g. the second simulated peak of Fig. 10a, between 08:07 UTC and 08:17 UTC is missing some contributions).".

L462-463: Why no simulated wildfires contribution?

The wildfires have been simulated but the contributions are insignificant compared to the other ones over the whole campaign. This point is made clear in the new manuscript as "Wildfires have influenced previous campaigns during specific episodes, especially close to the sources, as reported in Paris et al. (2008) and in Antokhin et al. (2018). Here the wildfires are simulated as well, but are not visible since the contributions are negligible compared to the other ones over the whole campaign.".

Technical corrections

Figure 1: Letters, mainly yellow, are too small.

All letters are increased and an outline is added for better readability.

Figure 2: Numbers on the aircraft are too small. I need help finding numbers 15-20. The indication for 2-6 is not necessary.

The original figure used in the article was from a campaign planning document. Belan et al. (2022) provides an updated figure of the aircraft configuration. We reproduce the figure from Belan et al. 2022 in the updated manuscript.

Figure 4: The data from Barrow can be shown here as well.

The data of Barrow is shown on the figure and the text is updated accordingly (see lines 253 and 271).

L277: "CO2, CH4" -> "CH4, CO2"

It is corrected.

L288: "1960 ppb" -> "1990 ppb"?

It is corrected.

Figure 8: Y-axis for (a) and (c) must be CO [ppb].

It is corrected.

Figure 10, 11: The axis for altitude is lacking.

Top and bottom altitude for each figure are written in the description as "(bottom at 12 m and top at 8963 m)" (this quote refers to Fig. 10, same type of indication for Fig. 11).

e.g. L399, L514: Subscript for CH4 does not done.

It is corrected.

L569: "Anokhin" -> "Antokhin"?

This reference is outdated and is removed in the new version manuscript. The two references now used for the aircraft of the campaign are Antokhin et al., 2018 and Belan et al., 2022 (see line 99)