Reply to the review by the Anonymous Referee #2

Thank you very much for the helpful comments. We revised the text and figures significantly and took all comments into account. Please find in black the original comments from reviewer #2 and in red our reply.

This work aims to provide insight into the effect that aviation NO_x has on the climate through its contribution to tropospheric ozone and depletion of atmospheric methane. In particular, the authors aim to show how the specific weather pattern into which the emission occurs could affect the long-term chemical impact. They evaluate this using a trajectory-tracking approach with simplified tropospheric chemistry, embedded within a global chemistry-climate simulation. Using the time and magnitude of the maximum change in ozone as proxies for the total ozone-related climate impact, they find the overall radiative forcing is likely to change significantly depending on the season and local meteorological conditions.

The central question of the paper is both interesting and important. If robust, this work could be a significant advance in the field of understanding the interactions between aviation and the atmosphere. It could also provide valuable, actionable information on how to reduce aviation's climate impacts. However, I have concerns regarding the methodological approach, and regarding the metric used to quantify "climate impact".

As such, I believe that major revisions are needed before this paper is ready for publication in ACP. I list major concerns below first, followed by minor issues.

Thank you very much for your review and seeing the value of our work to the community. Your input helped to improve the manuscript significantly. In order to provide a higher consistency within this work, we adapted major parts of the manuscript. Please find below an overview on the changes, based on the referees' comments, which will be discussed in more detail further below:

1. Abstract: The abstract was rephrased to better represent the work and results of this work
2. The introduction was extended to cover the chemical processes and an elaborated discussion why the O_3 maximum (w.r.t. time and magnitude) is selected for the analysis
3. The model description now includes a description of the lagrangain model used
4. Section 3 was eliminated since the chemical system is now covered in the introduction. A subsection was added to the results, discussing the variability of the O_3 maximum and how typical temporal developments of early, mid and late O_3 maxima look like.
5. The altitude analysis was changed to now focus on the mean altitude and not the altitude difference

6. The influence of temperature, NO_x and HO_x in summer and winter are now moved to a single section discussing the O_3 production efficiency
7. In order to account for the rapid cycling between OH and HO_2, we now analyze the OH recycling probability when discussing the CH_4 depletion
8. The discussion now includes a section on how the lagrangain modelling approach influences the results
9. All figures were updated to show the relation for winter and summer and make them visually more appealing

Major comments

While innovative, it is not clear to me that the chemistry and physics simulated within the "tracked" air parcel will provide an estimate of aviation's impacts on the atmosphere which are accurate or consistent with the "parent" EMAC model.

1. Is it validated against any estimates using more conventional techniques? The authors cite Grewe et al (2017c) Section 4.3, but this does not appear to have an explicit comparison of the Lagrangian model's result for aviation impacts compared too ther studies. A direct comparison showing (e.g.) the ozone and methane response in EMAC when performing a conventional simulation of an increment in aviation NOx would be very helpful. Alternatively, if such a comparison is already present in (eg) Grewe et al (2017c), a quantitative evaluation (e.g. "agreement to within X% when simulating aviation emissions") would be helpful. If not already present in Grewe et al2017c I recommend that such an analysis be added to Section 5.

Thank you for pointing this out. The concept used in this study is designed to gain more insights in aviation effects, which conventionally can otherwise not be obtained, limiting the comparability to conventional approaches. However, the following four indicators support the consistency of the modelling approach:

1. The transport scheme is reasonably well established (Brinkop and Jöckel, 2019).
2. The chemical response to a local emission agrees well with earlier findings of Stevenson et al. (2004), who simulated the monthly mean response of ozone and methane to a NO_x pulse and which are very similar to the results from this approach (Grewe et al., 2014, their Fig. 9).
3. The use of a trajectory analysis to interpret either observational data or modelling data is well established.
4. A first verification of the resulting global pattern of the atmospheric sensitivity to a local NO_x emission by comparing to sparsely available literature data was promising (Yin et al., 2018).

We added an elaborated discussion on this matter to the discussion section of our manuscript.

Brinkop, S. and Jöckel, P.: ATTILA 4.0: Lagrangian advective and convective transport of passive tracers within the ECHAM5/MESSy (2.53.0) chemistry–climate model, Geosci. Model Dev., 12, 1991–2008, https://doi.org/10.5194/gmd-12-1991-2019, 2019.

Grewe, V., Frömming, C., Matthes, S., Brinkop, S., Ponater, M., Dietmüller, S., Jöckel, P., Garny, H., Tsati, E., Dahlmann, K., Søvde, O. A., Fuglestvedt, J., Berntsen, T. K., Shine, K. P., Irvine, E. A., Champougny, T., and Hullah, P.: Aircraft routing with minimal climate impact: the REACT4C climate cost function modelling approach (V1.0), Geosci. Model Dev., 7, 175–201, https://doi.org/10.5194/gmd-7-175-2014, 2014.

Stevenson, D. S., Doherty, R. M., Sanderson, M. G., Collins, W. J., Johnson, C. E., and Derwent, R. G.: Radiative forcing from aircraft NOx emissions: Mechanisms and seasonal dependence, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2004JD004759, 2004.

Yin, F., Grewe, V., Frömming, C., & Yamashita, H.: Impact on flight trajectory characteristics when avoiding the formation of persistent contrails for transatlantic flights, Transportation Research Part D: Transport and Environment, 65, 466 – 484, doi: https://doi.org/10.1016/j.trd.2018.09.017, 2018

2. More detail on the Lagrangian model would be helpful. For example, what is the total air mass of the well-mixed box? How is diffusion treated? This is important because of the role of non-linear chemistry (see e.g. Kraabøl et al 2002), which could result in suppressed ozone production when concentrations are very high (i.e. early in the plume's development).

We agree that more information on the Lagrangian model is needed such that the reader understands how the transport processes (identified to be important for the resulting ozone and methane change) are simulated. Therefore, a detailed description of the submodel ATTILA was added to the base model description. Note also that the Lagrangian transport scheme has been evaluated in detail by Brinkop and Jöckel (2019).

Brinkop, S. and Jöckel, P.: ATTILA 4.0: Lagrangian advective and convective transport of passive tracers within the ECHAM5/MESSy (2.53.0) chemistry–climate model, Geosci. Model Dev., 12, 1991–2008, https://doi.org/10.5194/gmd-12-1991-2019, 2019.

The approach used appears to quantify the direct effect that aviation NO_x and H_2O could have on the climate through increases in short-term ozone and decreases in methane, on the basis that both are greenhouse gases. However, one of the major effects of aviation NO_x is a long-term reduction in tropospheric ozone, driven by the methane loss (see e.g. Holmes et al 2001). Is this accounted for here? If I have understood the trajectory-tracking mechanism correctly it does not include feedbacks, so I would not expect it to be accounted

for. If so, it would be useful to include an estimate of the total magnitude of the missing long-term ozone loss.

This effect is known as primary model ozone (PMO). Within REACT4C, PMO was not explicitly simulated, as in most other studies, since this change in ozone occurs far beyond the 90 days simulated. Simulating it explicitly would be too computationally expensive and unfeasible for this modelling approach. Instead a constant scaling factor was applied to the climate impact due to changes in methane, to account for PMO. Due to this simplified approach, PMO cannot be taken into account in this study. We added a paragraph to section 2.3 to explain these circumstances.

The analysis is predicated on the idea that the time and magnitude of the peak change in ozone is a significant indicator of overall climate impacts. However, it is not clear to me that this is the case, and I could not find a quantitative justification or citation to this effect in the paper. The closest I found was the assertion that, in REACT4C, a higher ozone concentration change "generally" leads to a larger RF (p7, line 12). Why use these metrics instead of (e.g.) the total integrated ozone perturbation over the time of the simulation?

In general, we agree that the most obvious variable for addressing the climate impact is the integrated ozone perturbation or the resulting RF. We have realized that our introduction guided both reviewers into a wrong direction on the objective of this paper. Apologies for that. Identifying the resulting climate impact is not the objective of this paper. In this work, we are mainly interested in how transport processes effect the resulting ozone and methane change. By using our dataset, we identify that two characteristics (time and magnitude of the ozone maximum) differ most under varying weather conditions. To improve the introduction and to avoid such a misinterpretation, we added an additional figure (Figure 1 in the revised version), a table (Table 1 in the revised version) and an additional paragraph to the introduction clarifying the purpose of this manuscript. In the figure we show that two emission regions next to each other lead to different ozone perturbations, characterized by a different time and magnitude. Since, the emission occurs under different weather conditions (in and west of a high pressure ridge), we open the question if these differences can be explained by the different weather conditions experienced by the air parcels.

At the end of section 2.1, it is stated that "50 trajectories [are initialized] at 6, 12 and18 UTC", but this is then immediately followed by "in the present study, only 12 UTC is considered". I'm confused – why have the first statement? Also, what is the error which we can expect from only including one time point? Won't this result in same geographical biases?

6, 12 and 18 UTC was only simulated for WP1 within REACT4C. Grewe et al. 2014 demonstrated that the results are more sensitive to the horizontal than the temporal representation. It was thus decided to only simulate 12 UTC for all other weather patterns, due to limited computational resources. We therefore can only include 12 UTC within this study. We updated this paragraph to include this additional information.

Grewe, V., Frömming, C., Matthes, S., Brinkop, S., Ponater, M., Dietmüller, S., Jöckel, P., Garny, H., Tsati, E., Dahlmann, K., Søvde, O. A., Fuglestvedt, J., Berntsen, T. K., Shine, K. P., Irvine, E. A., Champougny, T., and Hullah, P.: Aircraft routing with minimal climate impact: the REACT4C climate cost function modelling approach (V1.0), Geosci. Model Dev., 7, 175–201, https://doi.org/10.5194/gmd-7-175-2014, 2014a.

Many of the conclusions seem to treat correlations as causal links (e.g. section 4.3).Some conclusions – such as that "during summer the O_3 formation is limited by the background NO_x concentration, whereas in winter low HO_2 concentrations limit the total O_3 gained" – do not seem to be sufficiently supported by the data. It would be helpful to see a more explicit justification for why this must be the mechanism.

In order to approach this, we performed major parts of the analysis again. In addition, parts of Section 4.1, 4.2 and 4.3 were merged into a single section (now Section 3.3). In this section we now discuss which factors influence the efficient production of O3. In the case of the O_3-NO_x-HO_x relation, NO_x and HO_x are analyzed in parallel. The mechanistic why NO_x and HO_x limit the production of O_3 in summer and winter was added.

More generally, much of the analysis is not very quantitative – such as page 14, lines 3-5which states that a visual inspection of a correlation makes it "evident" that winter pat-tern 5 looks "almost the same" as for summer pattern 3. I strongly recommend that the authors redo this analysis in a more quantitative fashion and remove conclusions which cannot be both quantified and mechanistically explained in a way which is supported by data.

Thank you very much for pointing this out. In general, we agree with the comment. Some of the analysis is not obvious to the reader, based on the figures provided. We therefore expanded most of the analysis performed in this study and added results for summer and winter. In this particular case, the distribution of HO_x is now presented for WP1, WP5 and SP3 in an additional figure. The reason why HO_x is not the limiting factor for WP5 is now addressed by using this figure and mean and median values.

Some of the analysis used for chemistry is difficult to interpret, and seems to ignore the cycling nature of certain key species. For example, on page 13, it is stated that "35%and 42% of background OH is produced by HO_2 reacting with O_3 and NO, respectively". However, OH and HO_2 are expected to be cycling rapidly. As such, "production" of OH in this fashion is hard to interpret, since it usually matters more to consider what the sources of HO_x are. I am therefore skeptical of the claim that H_2O is only a minor source of OH. I recommend that the authors rephrase this discussion to be about the OH/HO_2 ratio and the production and loss of HO_x than to talk separately about OH and HO_2.

This is a good point and an important fact missing in our analysis so far. Two measures were performed to account for the cycling between OH and HO_2. In Section 3.3 (earlier Section 4.3) the analysis was performed with respect to HO_x instead of only HO_2.

Additionally, Section 3.4 (earlier Section 4.4) was completely redone. In order to account for the cycling between both species, we now analyze the OH recycling probability.

Minor issues

P2 l22: I went back and checked the claim that Gilmore et al (2013) showed that "during summer the climate impact is up to 1.5 times higher and only half in winter, when compared to the annual meaning". I do not think this is true. They did show that the ozone production efficiency was 50% higher than the average in summer, but this is largely compensated by changes in ozone lifetime. The overall change in ozone production rate is only about 10% above the annual mean in summer, and correspondingly about 10% below the mean in winter (see Figure 1 of said paper).

This is indeed correct. Thank you for pointing this out. We intended to write about the O_3 production efficiency which is about 50% higher in summer. We changed this section in the manuscript to represent their findings appropriately.

The claim that "the standard deviation of background concentrations [of ozone and NO_x] are generally considered to be higher than changes induced by aviation . . . making them hardly detectable" (p16 l30, and paraphrased on page 14, lines 17-19) is based on a single, outdated study. Wauben et al 1997 uses a 1995 aircraft NO_x inventory, now 20-25 years old. Since total aviation NO_x emissions have likely more than doubled since then (e.g. Wasiuk et al 2016), I think this claim either needs a more recent and robust backing or it should be removed.

We could not find a more recent peer-reviewed study addressing this problem. Based on their findings we still think that even a doubling of the NO_x emissions will not be detectable on a global scale and that their statement is still valid. We therefore keep this part within the discussion.

The manuscript has several spelling and grammar mistakes. For example, in several locations (e.g. p2, l34) the authors use the word "adopt" when I suspect "adapt" is intended, and on p12 line 19 the word "exited" should be "excited". I would recommend another sweep through the manuscript to fix these and other typos.

The errors pointed out are changed in the manuscript. In addition, the manuscript was checked for further errors of this type.

What happened to WP2 in Table 1?

WP2 was not taken into account in our analysis due to technical problems. These problems lead to an incomplete dataset. We added this information to section 2.3.