Author response to Referee 1

We kindly thank the referees for their relevant comments, suggestions, and corrections.

The responses are organized as follows: the reviewer's comment is in blue, our answers are in black, and the changes proposed for the revised manuscript are in italics (black for modified sentences, grey for unchanged sentences that have been copy-pasted here to remind the context). The numbering of the pages and lines corresponds to the preprint, not to the modified document. Last, we answer "Done" to all the comments suggesting a modification that we completely agree with, and that do not require any clarification in this document.

Before addressing the review itself, we have to highlight the update of the results since the preprint submission, as we made a small error on the radiative forcing calculation, more precisely on methane feedback: we computed the latter accounting for the decrease of methane lifetime due to reaction with OH, but we forgot to include the stratospheric loss and soil deposition.

In the abstract, the previous statement:

"The NO_x net ERF is systematically positive and ranges from 7.3 to 22.1 mW m^{-2} among the different models (14.1–22.1 mW m^{-2} without the least sensitive model)."

has been changed into:
"The NO_x net ERF is systematically positive with a model mean of 18.3 mW m^{-2}, ranging from 9.4 to 24.5 mW m^{-2} among the different models. This net NO_x forcing is reduced by 35% and 43% accounting for the negative forcing arising from the formation of nitrate and sulfate particles, respectively."

Radiative forcings related to NO_x in Table S6 have been corrected as follows (the short-term NO_x effect on ozone has also been updated, but it is smaller than 0.2 mW m^{-2}):

ERF (mW m^{-2})	EMAC-NO_x	LMDZ-INCA	MOZART3	OsloCTM3	GEOS-Chem
Former results					
Short-term ozone	27.7	43.0 [40.6]	42.0	34.0	56.0
CH_4 direct effect	-13.2	-17.1	-12.9	-12.9	-24.8
Long-term ozone	-5.37	-6.96	-5.24	-5.24	-10.1

Stratospheric H_2O	-1.74	-2.25	-1.70	-1.70	-3.26
Total CH_4	-20.4	-26.4	-19.9	-19.9	-38.2
Total	7.37	16.7	22.2	14.1	17.9
New results					
Short-term ozone	27.6	43.0 [40.6]	42.2	33.8	56.0
CH_4 direct effect	-11.9	-15.4	-11.5	-11.6	-22.0
Long-term ozone	-4.83	-6.24	-4.68	-4.73	-8.93
Stratospheric H_2O	-1.56	-2.02	-1.51	-1.53	-2.89
Total CH_4	-18.3	-23.6	-17.7	-17.9	-33.8
Total	9.36	19.4	24.5	15.9	22.2

We also added further information on the calculation of radiative forcings, in Section 2.3:
"To derive the radiative impact of aviation-induced atmospheric composition changes, we use concentration-based kernels to calculate the stratospherically adjusted ozone RF and the instantaneous top-of-the-atmosphere RF due to aerosol–radiation interactions (Skeie et al., 2020; Samset et al., 2011). To perform the RF calculations, ozone and aerosol data from all models were interpolated to the kernel resolution (2.25° x 2.25° and 60 vertical levels). For the calculation of the ozone column for each model, the air mass from OsloCTM3 was used, following the method in Skeie et al. (2020). The ozone RF calculations from the kernel have been found to compare favorably against offline radiative transfer model calculations in LMDZ-INCA and OsloCTM3. More generally, kernel-based estimates of ozone RF have been found to agree with those from full radiative transfer in previous applications (Lund et al., 2021). The change in methane volume mixing ratio and the associated radiative forcing are calculated based on the modelled change in methane total lifetime for each simulation. It assumes, in addition to the methane oxidation by OH, a stratospheric sink characterized by a 120-year lifetime and a soil sink characterized by a 160-year lifetime. This calculation is combined with the methane feedback factor (referring to the CH_4 feedback on its own lifetime) and an emission non-steady state factor, method described in Berntsen et al. (2005), Hodnebrog et al. (2012) and Terrenoire et al. (2022). The methane reference mixing ratio is fixed to 1834 ppb. The methane feedback factor (f=1.45) is taken as the model mean from a recent model intercomparison (Sand et al., 2023). We use a non-steady state factor to correct for the fact that due to its long lifetime, methane steady state is not reached, such as assuming steady-state to derive the radiative forcing overstates the response (Grewe and Stenke, 2008). This non-steady factor was recently recalculated by Bellouin et al. (in prep.) to be 0.680 for present-day conditions. From this methane mixing ratio change, the methane RF is calculated

using the simplified equation from Etminan et al. (2016). The indirect long-term ozone and stratospheric water vapour RFs are calculated based on the change in methane mixing ratio adopting the normalized forcings from a recent model intercomparison (Sand et al., 2023). For long-term ozone we use a normalized forcing of 0.180 W m^{-2} $ppbCH_4^{-1}$ and, for stratospheric water, a normalized forcing of 0.058 W m^{-2} $ppbCH_4^{-1}$.
The calculated ozone ***and methane*** *RFs are converted to ERFs based on the efficacies provided by Lee et al. (2021)* ***(1.370 for the short-lived ozone forcing and 1.180 for the methane direct and indirect forcings)****. For aerosols, the kernel includes rapid adjustments for BC, thus representing the ERF, while the ERF/RF ratio is assumed to be 1.0 for the scattering aerosols due to lack of other information (Lee et al., 2021).”*

Another clarification has been added at the end of the paragraph on aerosol RFs commenting Fig. 11:
“[...] The RFs shown in Fig. 11 and discussed in this section are converted to ERFs to enable comparison with other studies in the following paragraphs (see Sect. 2.3). ***It is worth noting*** *that in our modelling set-up, the nitrate particles forcing is not only related to aircraft NO_x emissions driving the formation of ammonium-nitrate particles (NH_4NO_3). It is also affected by the oxidation of sulfur dioxide (SO_2) into sulfate particles due to NO_x-induced OH formation. This additional sulfate then collides with ammonia to form ammonium-sulfate ($(NH_4)_2SO_4$) particles, thus competing with nitric acid to react with ammonia and thus to form nitrate aerosol. Terrenoire et al. (2022) show that the nitrate forcing is more negative (or less positive) when one accounts for the aircraft NO_x effects on aerosols. With this limitation in mind, we note that the nitrate forcing would reduce the net NO_x forcing from a model mean forcing of 18.3 mW m^{-2} to 11.9 mW m^{-2}.”*

And in the conclusion:
“The net aerosol ERF_{ari} varies between -6.5 and -17.8 mW m^{-2} (multi-model mean of -11.6 mW m^{-2}), ***with the formation of sulfate and nitrate particles that reduces the net NO_x forcing by 43 % and 35 %, respectively.”***

Concerning the introduction, we also point out the division of the second (and long) paragraph into two smaller paragraphs: one more general for the non-CO_2 emissions, and the second one more specific on the consequences of aviation NO_x.

Last, in the end, six figures have been added into the Supplement, both from the reviewers’ suggestions (Figs. S6, S9, S11–S13), and from our own initiative, as we found it would be useful for the reader (Fig. S7).

—----------

General comments

- This paper provides an important addition to the existing literature concerning the modelled effects of aviation NOx pollution, which has high practical and policy relevance given the likely evolution of air transport.

- This paper does not itself add the strength of validating model predictions against observations. However, the referenced (published) companion paper (Cohen, 2025) does address how 4 of the 5 models compare against IAGOS observations. GEOS-Chem is the exception there though, which is often an interesting outlier in this paper.

- The overall climate forcing effect of aviation NOx is still not well-constrained by this study, as the authors discuss, but it usefully adds results and analysis from further models. The authors should make clear what is new in this study in terms of models, date ranges and emissions covered, etc to distinguish it from previous work.

We totally agree with this important suggestion, made by both reviewers.

As written in the response to Referee 2, we added some additional remarks, as follows:

- In the introduction, in the third (formerly second) paragraph, we clarified two points addressed by our study, about the estimates from Lee et al. (2021). One concerns the formation of aerosols due to aircraft NO_x, and the other one concerns the heterogeneity of the studies reviewed by Lee et al. (2021), which is a strategy different than ours as it synthesizes 60 estimates taken from 20 studies.

"[...] Since methane is an ozone precursor in the troposphere and a water vapor precursor in the stratosphere, its increased sink in the short term decreases the production of these two species during the decade following the emission, thus increasing the cooling term due to methane destruction (e.g. Myhre et al., 2011). Through the production of tropospheric OH, aircraft NO_x also promotes the formation of sulfate and nitrate (NO_3) particles (because of an enhanced oxidation of SO_2 and NO_x), thus acting as an additional cooling factor (Brasseur et al., 2016; Terrenoire et al., 2022). Based on the existing literature, Lee et al. (2021) estimated the net aviation NO_x impact to be an effective radiative forcing (ERF) of 17.5 [0.6–28.5] mW m^{-2} for the year 2018, resulting from a positive ERF from short-term ozone of 49.3 [33–76] mW m^{-2} and a negative ERF from long-term methane decrease of -34.9 [-65 – -25] mW m^{-2}, thus highlighting high uncertainties for both processes. However, Lee et al. (2021) do not account for the NO_x effect on aerosol formation, as they point out the limited number of studies on this topic and the large associated uncertainties. Also, the assessment by Lee et al. (2021) is based on an important set of modelling studies which have been harmonised regarding time periods and aircraft emissions in order to account solely for inter-model differences in both chemistry and radiation. In this study we revisit the aircraft NO_x perturbations on chemistry and climate based on the latest generation of five global chemistry-climate/transport models and on a commun modelling protocol regarding surface and aircraft emissions and time period covered. This further reduces the need for harmonisation of the model results as a post-treatment and allows to focus on the inter-model differences in the treatment of chemistry and dynamics."

- In the introduction, at the end of the fourth paragraph, formerly the third, which reviews the literature:

"[...] processes not accounted for, as in Brasseur et al. (2016). ***The latter study also shows the impact of aircraft on gaseous and aerosol species, but not necessarily with the same models, thus limiting the interpretation of the results. Last, the aviation impact on nitrate aerosol is relatively new in the literature, and most of all, there is no model intercomparison regarding its perturbation due to aircraft emissions."***

- In the introduction, at the end of the 5th (formerly 4th) paragraph, we chose a more affirmative formulation regarding our explanation for the spatial pattern of the ozone perturbation:

"The objectives of this study are (1) to provide an overview of the methodology of the harmonized multi-model study, (2) to present aviation-induced changes in the concentration of reactive species and the extent to which models agree, as well as specific differences between individual models, including an evaluation of the linearity of aviation induced effects. ***Thanks to the ancillary variables provided by two of the models, our intercomparison is the first to suggest an explanation for the pattern of ozone changes in response to aircraft NO_x.*** *Finally, (3) to provide estimates of the radiative effects of the simulated aviation-induced ozone, methane, and aerosol changes."*

This sentence replaces the former one:
"Based on ancillary variables provided by some of the models, we suggest additional explanations for the model results."

- In the summary and conclusions section, first paragraph:

"In light of increasing air traffic, we perform and document a new multi-model assessment of the atmospheric composition response to aviation NO_x and aerosol/aerosol precursors emissions, and the associated radiative forcing of climate, ***notably providing the first model intercomparison on the impact of aviation on nitrate aerosol. The purpose is to refine estimates of the impact of aviation by limiting differences between the models in their implementations, for a better understanding of the intermodel variability in the results.*** *We present a model intercomparison involving five state-of-the-art chemistry-transport models (CTM) or chemistry-climate models (CCM). For this study, each participating model provides a set of present-day runs, including at least one reference run with all the anthropogenic emissions, and one perturbation run without aviation emissions, using the same anthropogenic and biomass-burning emission inventories.* ***The models used to estimate the impact of aviation-induced aerosols are also applied to the gaseous phase, thus maximizing the consistency between the estimates for gaseous and aerosol species. For all models, the effective radiative forcings are calculated with the same radiative code, the same grid, and the same feedback factors."***

- *Same section, second paragraph:*

"With the same common protocol applied to the models, our net NO_x ERF estimate and range, as well as the relative distribution between individual contributions from ozone, methane, and water vapor changes to the net NO_x ERF, remain similar to previous studies (normalized to present-day emissions)."

- There is probably less existing literature on the climate forcing from aerosols resulting from aviation pollution, so that novelty should be highlighted.

We agree with Referee 1. Additionally to Referee 2's suggestions on the abstract regarding aerosols too, we added more context:

"This work shows encouraging results regarding our confidence in aviation NO_x-induced ozone response because of a good model agreement. To a lesser extent, some similarities in the results regarding aerosols are also encouraging, given the few existing model intercomparisons on this topic. However, results also highlight [...]"

Specific comments

- Abstract L28: "because of a better model agreement." better than what, previous comparisons? Or just "good model agreement"?

"Good model agreement" is a better formulation, we changed it in the abstract. Thanks for the suggestion.

- Section 2.1 L104: "the pairs of mixing states correspond to the hydrophilic-hydrophobic dichotomy" I understand that organic particles may start hydrophobic but become hydrophilic, but is this "dichotomy" a well-known term here? I'm not an aerosol expert but don't recognise it. Others may benefit from more explanation here too.

No, the term "dichotomy" is not a technical term. We clarified as follows:

"For each model described in Table 2, the pairs of mixing states correspond to the hydrophilic or hydrophobic state (with hydrophobic particles that can evolve into hydrophilic through ageing processes), and the three mixing states for EMAC-aer also include a mixed-particles category. It applies to both black carbon and organic carbon."

- Introduction L84: "the recent inventories used by CMIP6 for anthropogenic surface and aircraft emissions", but CMIP6 data only runs to the end of 2014 (or 2015?); the referenced McDuffie 2020 work extends CMIP6 for a few years as CEDS(GBD-MAPS). See also L255.

We clarified, as for the next comment.

- Section 2.2 L255: "The historical anthropogenic emissions are taken from the Community Emissions Data System inventory CEDSv2 (McDuffie et al., 2020; O'Rourke et al., 2021,

regarding NOx, SO2, and BC emissions)." Having a look at McDuffie 2020 they only go to 2017, while the O'Rourke 2021 reference is for CEDS v_2021_04_21 which seems to go out to 2019. As some of the models run to 2018 or 2019, does that mean that the emissions for some species (those not in your O'Rourke list) were extrapolated or otherwise estimated to cover 2018 and 2019? A brief explanation of that would be good if so.

Thanks to Referee 1 for pointing out this confusing inconsistency. To clarify: at the time we made the simulations, the latest available dataset was the CEDS emission files from Hoesly et al. (2018), ending in 2014 (included). Hence the use of the emissions taken from the future scenarios of Gidden et al. (2019). Concerning the three mentioned species (NOx, SO2 and BC), it was unclear: those are the three species taken into account for aircraft emissions.

We modified the paragraph as follows:

"The historical anthropogenic emissions are taken from the Community Emissions Data System inventory CEDS (Hoesly et al., 2018). Regarding aviation emissions (represented here by NO_x, SO_2, and BC), these files resolve aviation emissions with 25 vertical levels (or injection heights) from the surface up to 15 km."

- Section 2.2 L260: "For these years (2015–2018)," but from Table 2 the GEOS-Chem model ran for 2019, so should it be 2015–2019 here? Or maybe GEOS-Chem used different emissions here too? (Later: more on GEOS-Chem emissions being different in section 2.3 L304, maybe worth a cross-reference)
We added a mention of 2019 in GEOS-Chem:

"For these years (2015–2018), the differences between the scenarios remain small (less than 6 % for NO_x), as are the differences with the year 2014 in the CEDS inventory, given that the scenarios data sets have been harmonized with the historical data sets to ensure a consistent evolution before and after this transition year (further information for the year 2019 in GEOS-Chem is available below)."

- Section 2.2 L280: "We rescaled the EMAC-aer perturbations using the NOx emissions from CEDS for the period 2014–2018 (and the same rescaling factor for every species)," I'm not concerned that you had to do something special there, but don't really understand how that rescaling worked. E.g. for SEN100 I would expect you just zeroed out the aviation emissions in use by this model. So what is the "rescaling" of the perturbations? (Later: maybe related to section 2.3 around L294? In which case could cross-reference that section.)
We did not notice that we were mentioning the rescaling process before its definition, which is redundant and confusing. We therefore removed this sentence, as it is explained in detail a few lines after.

- Section 3.1 L391: "Due to stratospheric intrusions, an extension of the mean ozone perturbation is visible at low latitudes, downward and equatorward." Is this really about

stratospheric intrusion? If we had a global intrusion of that shape and size, surely it would be a named phenomenon... maybe it is partly. But we can already faintly see enhanced NOx in the same area in Figure 4. So maybe it is largely created there in situ. After all, many flight routes following approximate great circles will climb and descend from lower-latitude airports but then (for most of the flight) cruise at higher altitude and further north, giving the same kind of "diagonal" pattern on these latitude-altitude plots. (Later: at L479 you make such a case for aerosols, though with actual near-surface emissions there too.)
We agree with Referee 1 on this comment. We changed our statement as follows:

"An extension of the mean ozone perturbation is visible at low latitudes ***(as for NO_x, to a lesser extent),*** *downward and equatorward,* ***which can be linked to the aircraft trajectories and to subsidence motions."***

- Section 3.1 L392: it might be interesting to comment on the somewhat different distribution of OH, which is most enhanced inside the UT compared to NOx and O3 in the LS. I guess because there is much more H2O below the tropopause. (You come back to OH at L412+ but without venturing an explanation for its different distribution.) (Also return to OH around L444, could mention further discussion "below" at least here.)

The mention for further discussion below has been added as follows:

"In spring, the models generally exhibit a dipole structure with positive values centered near 40°N, mostly below the tropopause as well, and a negative response at high latitudes in the LMS ***(further discussion on chemical mechanisms responsible for the perturbation patterns is available below).*** *The negative values are more pronounced with OsloCTM3, ..."*

We also add a clarification that H_2O is not the only source of OH:

"The net OH response is higher with LMDZ-INCA (as expected from Fig. 6), characterized by a stronger positive response in the mid-latitudes and a weaker negative response in the high-latitude LMS. This positive OH response is consistent with the H_2O tropospheric background being the greatest in LMDZ-INCA, ***though the hydroperoxyl radical (HO_2) is another source of OH in the UT, which we did not investigate here.*** *Last, [...]"*

- Section 3.1 L409-410: reading the text and looking at Figure 5 side by side, two of the quoted numbers seem mismatched by a couple of tenths of a ppbv (6.7 vs 6.9, 13.0 vs 13.2). The differences may not be scientifically significant but they should agree, if only to make relating the text to the plots more straightforward, and anything >0.1 difference cannot just be a difference in rounding here. Perhaps the text was written against an older version of the plots or vice versa?
We thank Referee 1 for pointing out this inconsistency which is, indeed, probably due to a

last update made in the figures but not in the corresponding paragraph. All the values have been corrected now, and the statements adjusted, as follows:

"Some differences are well visible between the models. Contrasting with the ozone perturbation generally in the altitude range of 8–12 km, the OsloCTM3 ozone response is located lower in the troposphere, in the range ~5–10 km. In June, these lower altitudes in ozone perturbation are characterized by the strongest NO_x perturbation in the LMS, with a zonal mean above ***120 ppt*** *at all latitudes beyond 45° N. By contrast, the other models* ***do not reach 100 ppt (LMDZ-INCA, GEOS-Chem), or very locally (EMAC-NO_x, MOZART3).*** *In terms of mixing ratio, the maximum value of the ozone response ranges between 3.0 and* ***6.7 ppb*** *in January, with three models relatively similar to each other (EMAC-NO_x, LMDZ-INCA, and MOZART3, within a range of 4.2–4.6 ppb). During the seasonal peak, the maximum value exhibits more discrepancies, with two models having lower maxima* ***(4.1 and 6.9 ppb*** *for OsloCTM3 and EMAC-NO_x respectively), two models having higher maxima (14.2 and* ***13.2 ppb*** *for MOZART3 and GEOS-Chem), and one model having intermediate values (10.5 ppb for LMDZ-INCA)."*

- Section 3.1 L420 "substantially weaker... -38 %" weaker than what, MOZART3?

We realise that it was not clear. It is explained now:

"With MOZART3, the maximum ozone response during JJA remains similar between the two studies, and 16 % lesser in DJF ***in the current study****. With EMAC-NO_x and OsloCTM3 however,* ***still in the current study compared to Søvde et al. (2014),*** *the maximum ozone response is substantially weaker in both seasons (-38 % and -47 % in summer, and -39 % and -51 % in winter)."*

- Section 3.1 L403-L422: this is a long paragraph that starts with O3 (fig 5), switches to OH (fig 6), then backtracks to O3 (fig 5). It might be a smoother experience for the reader if you moved the OH discussion to be a new, following paragraph.

Done.

- Section 3.1 is pretty long overall, and without any lower-level headings you can't cross-reference easily between parts of it. Maybe split it up into subsections e.g. for the different species discussed?

The species are regularly compared together, almost in each paragraph, which makes it hard to separate them into different subsections. Instead, we divided Section 3.1 into subsections linked to the kind of analysis we make:

3.1.1 Seasonal cycles in the UTLS (linked to Fig. 1: Hovmöller diagrams)

3.1.2 Geographical distribution in the UTLS (linked to Figs. 2 and 3: maps)

3.1.3 Zonal cross sections (linked to Figs. 4–6)

3.1.4 Involved chemical mechanisms

3.1.5 Dependence on the chemical background

- Section 3.1 L452 "stronger in models with lower NOx" I agree, also lower NOy more generally so could mention that too

We added an anecdotic comment noticing it, as follows:

"It appears that the responses in ozone and OH in the UTLS are generally stronger in models with lower NO_x (and lower NO_y) UTLS background, though this is not the only factor controlling O_3 and OH sensitivities to NO_x emissions."

- Section 3.1 L453 "The OH response increases with H2O background" the GEOS-Chem point is a real outlier here; I see no correlation if I include that as an equally valid datapoint. You do mention in the next sentence that it is much higher but I'm left not really knowing the cause of that. If GEOS-Chem is excluded as a special case, then arguably you do have correlations with background CH4 and O3! So we need a good reason to include it, or not, when considering OH at least.

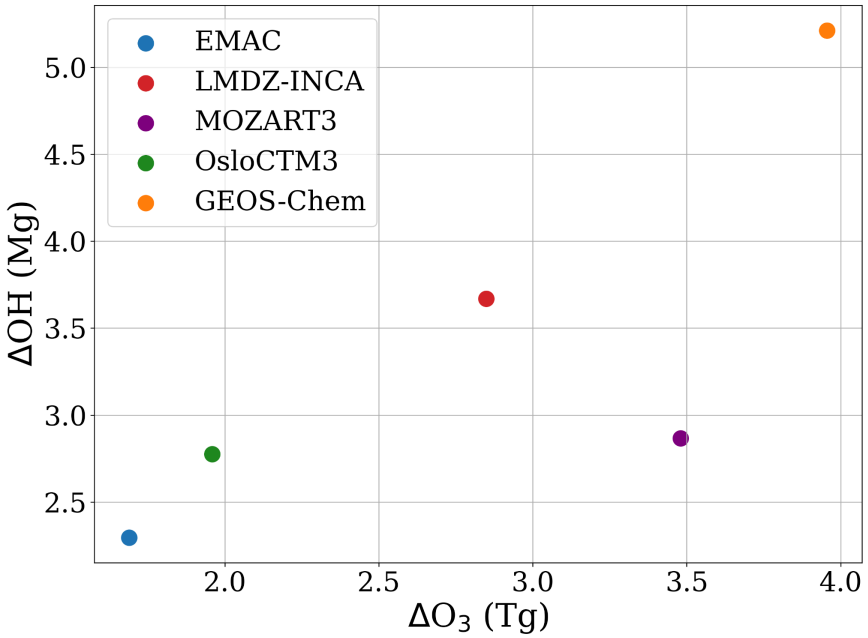
We added a note on the GEOS-Chem "exception", possibly due to substantial differences in the spatial distribution of NO_x emissions. Following Referee 2's request, we added maps showing partial columns for NO_x emissions, and it appears that GEOS-Chem has more balanced emissions between West and East hemispheres, whereas the other models have much more emissions in the Atlantic area (including North America, North Atlantic and western Europe) than in Asia (mainly East Asia, South Asia, and Siberia).

"The OH response is correlated with the ozone response, and with the H_2O background if not for GEOS-Chem. It is higher with GEOS-Chem notably because this model specifically does not show an OH negative response in the polar LMS. One possible explanation comes from the spatial distribution in NO_x emissions that differs between Flightradar24 and CEDS inventories as mentioned in Section 2.2, leaving part of the emitted NO_x evolving in different chemical regimes between GEOS-Chem and the other models. On the opposite, the OH response is lower with EMAC-NO_x as both the ozone response and the H_2O background are relatively low. The OH response is comparable between MOZART3 and OsloCTM3, because of their strong negative response in the LMS, as seen in Fig. 6. The net OH response is higher with LMDZ-INCA (as expected from Fig. 6), characterized by a stronger positive response in the mid-latitudes and a weaker negative response in the high-latitude LMS. This positive OH response is consistent with the H_2O tropospheric background being the greatest in LMDZ-INCA."

- Section 3.1 L454 "OH response... is correlated with the ozone response." I completely agree, but we can't really see that from figure 7, because there is no plot of delta(OH) against

delta(O3). I plotted it myself with approximate values from figure 7 for each model and get a nice correlation (just MOZART3 a bit of an outlier from an otherwise straight line, as it happens, though I wouldn't draw any great significance from that). I suggest you add such a plot.

We agree with Referee 1 that this deserves a figure to make this statement more evident. We added the following figure in the Supplement, as the new Fig. S6.



- Section 3.2 L480: "and due to subsidence"; I'm no expert on this but perhaps that should be couched with some uncertainty, as you haven't really quantified whether there is more BC close to the ground because that's where it was emitted or because it sinks. (In principle maybe you could by a comparison with the input emissions distribution fed into the models.) Out of interest a quick web search gave a figure of 0.5 km/month as a subsidence rate in https://doi.org/10.1126/science.aax1748 (but for the stratosphere?). (The BC lifetimes you later give around L492 might then suggest it disappears too quickly to subside much.)

Thanks to Referee 1 for this investigation. The paper cited here (Yu et al., 2019) apparently concerns stratospheric dynamics, which does comprise subsidence but with different characteristics than in the troposphere. But to be less affirmative, we added “possibly” before “due to subsidence”.

- Section 3.2 L480: "The LMDZ-INCA and GEOS-Chem models show similar responses, with a maximum in April." Looks like it is higher for both in May not April, if my reading of the colours in Fig 8 is correct.

Though it is less visible with GEOS-Chem, the figure shows a higher maximum in April, at cruise levels at the mid-latitudes, with a small orange-red area.

- Section 3.2 L523: Terrenoire (2022) could get a mention too in this paragraph about previous aerosol perturbation work.

It is true, but Terrenoire et al. (2022) uses LMDZ-INCA with the same version as in our study, which would bias the comparison with the literature. We forgot to mention that this paragraph concerned the studies using a different model than the ones in our study, it is done now.

"Among the studies using different models than this paper, Unger et al. (2013) present the same spatial pattern [...]"

- Section 3.2 L535: "are more significant" maybe "are more significant by mass" say, as the BC might be more significant (say) in health terms?

Done.

- Section 4 L556/Figure 11: I presume the plots shown are a global average, but it would be good to state this explicitly. Also these ones are not month-specific, so are they annual averages? (I guess so from the "in most other months" comment about GEOS-Chem.)

The text has been updated to *'Fig 11 shows the **global annual** average vertical profiles'*, as in Figure 11 caption.

- Section 4 L587: I just want to check I've understood the units correctly in "a range of 1.22–1.26 % (TgN yr-1)-1 in methane lifetime". So for example if my no-aviation CH4 lifetime was 10 years, it would decrease at the low end by (1.22/100)x10 = 0.122 years for each Tg of N aviation emissions? No need to clarify anything unless that interpretation is incorrect. Though should make it clear it is a decrease (could make numbers negative to emphasise that, as they are in Table 5). Also Table 5 has (-)1.25% not 1.26% for EMAC-NOx, another rounding difference I guess.

Yes, that interpretation of the units is correct. We have updated to emphasise that it is a decrease: *"a range of -1.22 to -1.26 %"*. Lastly: we apologize for confusion, the right value is -1.26%.

- Section 4 L590: looks like an error in "GEOS-Chem sensitivity is ~190 % greater", should be ~90% (=(2.36-1.23)/1.23)x100, taking 1.23 as the approx average of the lower 3 model values)

Done.

- Section 5 L659: "and a moderate perturbation (> 40 ppt) confined to the northern midlatitudes in winter" looking back at Fig 2 I'd say that's also true only for the high-traffic areas, which isn't quite clear here

Yes, it was implicit, as the 3 high-traffic areas are located in the northern mid-latitudes. We added “all of them in the mid-latitudes”, in order to make it clear.

- Section 5 L667: "mean net aviation NOx ERF of 15 mW m-2" but at L592 it is "mean value of 16 mW m-2", i.e. 15 vs 16, again just small differences but they should be consistent, especially with this headline number

Once again, thanks to Referee 1 for pointing out this inconsistency. We homogenized it with a mean value of 15.6 mW m^{-2}, and a range of [7.3, 22.1] mW m^{-2}.

Edit: with the update of ERFs mentioned at the beginning of this document, the new phrase is finally: “*Overall, we estimate a positive net aviation NO_x ERF in these model experiments, with a multi-model mean value of 18.3 mW m^{-2} and a range from 9.4 to 24.5 mW m^{-2}.*”

- Supplement S3 Figure S6 [now Figure S8]: as the emissions inputs are the same apart from GEOS-Chem (explained L274), I wonder why the MOZART3 ENOx has a notably weaker northern latitude band, whereas the others have similar intensity in two latitude bands?

The MOZART3 model uses annual emissions with monthly weight factors to simulate a seasonal cycle, though this cycle shows less magnitude than the monthly emissions used in the other models. The NO_x levels in MOZART3 thus appear to be ‘flatter’ and more evenly spread over time compared to what is observed in other models, but the global totals and other NOx figures do not show anything unusual about MOZART3 in this regard.

- Supplement S4 Table 3: I find the "Total CH4" values are ~1.54x the "CH4 direct effect" values. At L316 it says "the methane feedback factor from Sand et al. (2023)". But Sand et al have a model mean CH4 feedback factor of 1.45 not 1.54. Best check this, and also clarify in the table legend what "Total CH4" means.

We clarified “Total CH_4“ in the caption of Table 3 (now in Supplement S6):

“The row ‘Total CH_4’ corresponds to the sum of all ERFs related to CH_4, i.e. CH_4 direct effect, long-term ozone, and stratospheric H_2O. The row ‘Total’ corresponds to the sum of the two rows showing the short-term ozone ERF and ‘Total CH_4’.”

We have also clarified the meaning of the methane feedback factor (Section 2.3, completed with further information on the ERF calculations), as it relates to the changes in methane lifetime due to changes in methane itself, and not the effects on long-term ozone and stratospheric water vapour (which is another quantity).

- Supplement S4 Table 3: The "Total" row has your headline NOx net ERF values. I note these values include "CH4 direct effect" and not "Total CH4", otherwise they would all be

lower (0-15 mW/m2 instead of 7-22 mW/m2). Again at L316 it mentions using the methane feedback factor (which I think gives the "Total CH4" numbers), but those feedback-adjusted values do not seem to be used in the headline total ERFs. For the uninitiated (including me), could you explain the use of feedback-adjusted (or not) CH4 ERF I guess in the main text?

In Table S3, the row “Total CH4” corresponds to the sum of methane direct effect and its two indirect effects: long-term ozone decrease, and stratospheric water vapour decrease. Same for the former question, it has been clarified in the table caption.

Technical corrections

Many of this are very small grammatical points.

- Abstract L19: "and the uncertainty on their Effective Radiative Forcing (ERF) of climate" should be "uncertainty in their..." (in not on), and might be better to say "the uncertainty in their climate Effective Radiative Forcing"
Done.
- Abstract L28: "However, results also highlight" -> "However, the results also..."
Done.
- Introduction L41: "black carbon (BC) - or soot - particles" to be super-picky those should be em dashes not hyphens
Done.
- Section 2.4 L104: "mixted" presumably "mixed"
Done.
- Section 2.1.1 L158: "used to EMAC-aer" should be "used in EMAC-aer"

Done.
- Section 2.1.3 L189: "organic halogen compound" presumably "compounds"
Done.
- Section 2.2 L243: "without any aviation emission" should be "emissions"
Done.
- Section 2.3 L288: "derive 5-year averages for each month" might be clearer to say "calendar month" here, as I assume you mean you're making an annual climatology, e.g. averaging the 5 different Januarys together, etc?
Done. And yes, we confirm Referee 1’s interpretation.
- Section 2.3 L291: should be "factor of 5"
Done.
- Section 3.1 L328,L334,L337/Table 4: in the text we have "9.9 TgO3" (three times) but the table has "10.0" for GEOS-Chem, so I think the same number has been rounded in different ways with the same number of decimal places. Easier to relate the two if they are presented identically.
Done. We totally agree, and updated the text according to the table.
- Figure 1: maybe just a preprint thing but in the current PDF the quality isn't too good, looks like it has been compressed coarsely as a JPEG or something. But up to the journal what they

expect really!
We thank Referee 1 for noting it and for leaving us the choice about it, which we forward to the journal.
- Figure 1: I realise all the other stacked plots show the different models vertically going down the page, but this one shows models going horizontally across the page. It's not wrong, and I know it takes a lot of effort to change figures, but if you needed to rework it anyway for other reasons you might consider if it would be nicer to be consistent and have models going down and species going across here too.
We agree it can be confusing, but we found this configuration practical to compare the seasonality of the different species together.
- Section 3.1 L390: redundant "for ozone" in "For ozone, the perturbation even peaks ... for ozone"
True. We removed the first iteration.
- Section 3.1 L430: "for the three models" maybe "for all three models"?
Done.
- Section 3.1 L432: should make clear the direction of the methane lifetime change depending on the NOx background
This sentence has been clarified as follows:

*"Table S2 also shows that the **aviation-induced decrease** in methane lifetime is **enhanced** by a factor of 5% [...]"*
- Section 3.2 L480: this paragraph is pretty long, maybe start a new one for SO4?
We agree with Referee 1 that the paragraph is quite long, but only two sentences are dedicated to SO_4 only, as for BC only. All the rest of the paragraph (both before and after) treats BC and SO_4 together, so cutting the paragraph in 2 would be less consistent. But it is as long as the next paragraph, which equilibrates the text.

- Section 3.2 L485: "but also in the whole free troposphere. The maximum shifts..." it isn't totally clear to me whether the second sentence here is talking specifically about EMAC-aer still, or all the models. From the plots I think it is just EMAC-aer. Could clarify that e.g. by doing "but also in the whole free troposphere; (in this model?) the maximum shifts..."
Done (we confirm that this concerns EMAC-aer only).

- Section 3.2 L490: "can be caused by a number of factors, as the BC lifetime" intended "such as the BC lifetime"?

Done.

- Section 4 L638: "Last, " should be "Lastly, "

Done.

- Supplement L33: "in bracket" -> "in brackets"

Done.