Review Response to ‘Impact of future aircraft NOx emissions on atmospheric composition and climate: dependence on background conditions’

Author responses are in blue, with quoted updated manuscript text in italics, and where only small changes have been made these are shown in orange.

Reviewer 1:

General comments (overall quality)

This paper provides a worthwhile 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 and foreseeable attempts to mitigate climate impacts by altering flight profiles.

This paper provides a useful cautionary lesson that strong variability exists between different atmospheric models, limiting the usefulness of single-model studies, so encouraging care in the use of their results.

I recommend publication subject to some clarifications and corrections as listed below.

Thank you for your detailed reading and thoughtful comments that have enabled us to improve and clarify the manuscript following your suggestions.

Specific comments (scientific questions/issues)

L105 "The experiments are all timeslice runs, with present-day meteorology," could you clarify what you mean there? I understand timeslice simulations to use a repeating climatological year of sea/ice surface temperatures and likewise whatever emissions forcing is of interest. But "present-day meteorology" sounds like nudging to reanalysis winds -- for some specific repeated year or a range of years? -- and how would that work for the year 2050, or is it just 2050 scenario emissions feeding historical (2014-2018) simulation years where reanalysis winds are available? So to be clear what inputs are repeated in an annual pattern, which vary year by year, what nudging is done if at all, and what 5 years do the simulations run across, are they all really 2014-2018?

Yes, the CTMs (OsloCTM and MOZART) use reanalysis data to run, and the CCMs (EMAC and LMDZ-INCA) are run in quasi-CTM mode - i.e nudged to reanalysis data. Have added a couple of sentences to clarify this and which inputs are from which year:

‘The experiments are all timeslice runs, with one year of spin-up and then run for five years (four years for OsloCTM3). Present-day meteorology from 2014-2018 is used for all runs (including for the future scenarios), taken directly from the reanalysis datasets for the CTMs and run (nudged) using a quasi-CTM mode for LMDZ-INCA and EMAC (see Table 1). All other inputs e.g. emissions, methane concentrations, correspond to the scenario and year (present day or 2050).’

L131 "The ozone column (in Dobson Units, DU)" I assume that means DU per kernel level; "column" might be interpreted as *total* column (from ground to space), which might be of interest but doesn't go into the kernel-based calculation as I understand it.

That's correct, have added *'per kernel level'*

L140 "change in methane lifetime from each simulation ($\Delta\tau$ CH4)" and L150 (equation 1): having looked at this myself recently, to make it clear I think it needs to be the _relative_ change in CH4 lifetime, a unitless quantity. Otherwise the units don't balance in equation 1. It goes back to Holmes (2011) who says "_relative_ change in CH4 lifetime" (my emphasis). If the calculations were done with an absolute change (say in units of years), it wouldn't work.

Yes indeed dTCH4 is the relative change in methane lifetime (%), have clarified in the text *'based on the modelled relative change in methane lifetime (%)'*

L144 "fnon-steady is a factor to correct for the fact that due to its long lifetime methane steady state is not reached". I'm having difficulty relating that to the referenced paper Grewe and Stenke (2008), could you spell out the use of this factor in more detail? (It becomes important later around L330 and again around L383, so we want a really good understanding of it.)

The methane non steady-state factor is defined as the ratio

fnonsteady(t) = dCH4nonsteady(t) / dCH4steady(t)

The methane change at steady state is provided as described in the paper by:

dCH4steady(t) = dTCH4(t) * feedback * CH4(t)

The transient change in methane mixing ratio dCH4nonsteady(t) can be calculated by solving for instance Equation (3) of Grewe and Stenke (2008) or based on a chemistry model as described by Lee et al. (2021). Based on the methodology described by Grewe and Stenke (2008), we derive non steady-state factors for SSP1-2.6 and SSP3-7.0 future scenarios of 1.06 and 0.90 respectively.

Have added a more in-depth description of this in the methods section:

"The methane feedback on its own lifetime ($f_{feedback} = 1.45$) is taken as the model mean from a recent model intercomparison (Sand et al., 2023). $f_{non-steady}$ is a factor to correct for the implicit assumption in using $f_{feedback}$, which assumes a steady state is reached (see discussion in Grewe et al., 2019; Lee et al., 2021). Due to its long lifetime, and slow feedbacks on lifetime via OH, a methane lifetime corresponding to steady state is not reached in any given year, and the methane concentration change resulting from NOx perturbations also depends on the recent history of the emissions of NOx, and of OH concentrations. For increasing NOx emissions, assuming steady state to derive the radiative forcing overestimates the methane concentration response (e.g. Grewe et al., 2019; Myhre et al., 2011), and vice versa for decreasing emissions. Based on the method described by Grewe and Stenke (2008) (solving their Equation 3) we

recalculated these factors for 2050 conditions for both SSP1-2.6 and SSP3-7.0 future scenarios and derived non-steady factors of 1.06 and 0.90 respectively.”

L160 "For long-term ozone we use a normalized forcing of 0.180 W/m2/CH4 ppbv and for stratospheric water a normalized forcing of 0.058 W/m2/CH4 ppbv." What is the source of those values, do you have a reference?

The reference is in the previous sentence, have updated to make it clearer. The values here are from the Supplementary Table 5 in Sand et al (2023), O3 ERF per CH4 flux and Strat. H2O ERF per CH4 flux, converted to units of W/m2/CH4 ppbv: *‘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, Table S5): for long-term ozone we use a normalized forcing of 0.180 W/m2/CH4 ppbv and for stratospheric water a normalized forcing of 0.058 W/m2/CH4 ppbv’*

L185 "The models show two grouped responses: a stronger ozone response at a higher altitude (LMDZ-INCA, MOZART3, ~200hPa), or a weaker, lower peak ozone response (EMAC and OsloCTM3, ~300hPa, see Fig 2b)." Actually zooming in on Fig 2b I'd say all three of LMDZ-INCA, MOZART3 and EMAC peak at 220-250 hPa; it is only OsloCTM3 that peaks at around 340 hPa, with a notably broader altitude distribution.

Have updated to correct this: *‘The models show varied responses: a stronger ozone response in LMDZ-INCA and MOZART3, and a weaker response in EMAC and OsloCTM3 (Fig 2b). OsloCTM3 also has a lower and broader peak (~340 hPa) than the other three models (which peak around 220-250hPa.’*

L192 "in the upper troposphere, where the NOx is emitted, ... while in the lowermost stratosphere, the increase is likely a combination of less transport across the tropopause, and less ozone destruction" Well that depends on latitude, a lot of flights will actually cruise in the lower stratosphere. Fig 1a shows present-day aviation emissions centred around 11 km altitude but many will reach 12 km. North of around 30 deg, those altitudes will be in the LS. And a large proportion of flights will be further north than that (Fig 1b). This is something to bear in mind in other places, i.e. aviation emissions are largely at a fixed range of cruise altitudes (9-12 km), which can be in the UT or LS depending on latitude and season.

There was some confusion in this sentence about whether the transport referred to the NOx emissions or ozone, which was misleading and has been updated. This comes directly from the results in Cohen et al. (2025b) (towards the end of Section 3.1, Figs S1-S3). Edited text here to more closely reflect those findings: *“Results from present-day simulations using the same model ensemble suggest that in the upper troposphere, increased local photochemical production of ozone causes the short-term ozone increase, while in the lowermost stratosphere, the increase is due to a combination of lower ozone net flux to the UT (due to smaller ozone gradient), and less LS ozone destruction via OH (Cohen et al. 2025b).”*

L253 "The models all show lower ozone concentrations in the northern mid latitude upper troposphere... under an SSP1 background..." I think it's more complicated than that because we have two deltas going on. Firstly we have deltaO3 = O3 (normal NOx emissions) - O3 (20%

reduced emissions), which is positive. Then (say for LMDZ-INCA in the UT) we have an increased deltaO3 (more red) for SSP1 background minus the same thing for SSP3 background. So that's like a deltadeltaO3! So when you finally say "show lower ozone concentrations" that sounds like a simple change in concentration (e.g. from 200 to 150 ppbv), but actually it's a change in sensitivity (how much O3 concentration changes with NOx). So maybe Fig 4a actually tells me something more like "The models all show enhanced ozone production from aviation NOx under an SSP1 (vs SSP3) background in the UT, and reduced production in the LS under SSP1 (vs SSP3) background".

Agreed, some more care needed describing the double difference - have updated as suggested *'The models all show reduced ozone production from aviation NOx under an SSP1 (vs SSP3) background in the northern mid latitude upper troposphere, and enhanced ozone in the LS under SSP1 (vs SSP3) background (Fig 4a).The relative magnitude of these changes varies between models, giving a different net response.'*

L263 "LMDZ-INCA simulates lower NOx concentrations in the UTLS region in the SSP1 background" well yes, but most dramatically in Fig S4 above 200 hPa (lower-latitude higher tropopause), so mostly above aircraft cruise altitudes (max ~180 hPa or 41000 ft). Though the NOx is still showing weakly lower around typical 200-300 hPa cruise altitudes so I guess still relevant to aviation.

Yes, the lower NOx concentrations in LMDZ-INCA are most pronounced at higher altitudes but as you say, there is also a weak decrease at cruise altitudes. Have added this: "*LMDZ-INCA simulates lower NOx concentrations in the UTLS region in the SSP1 background (mostly higher up but also weakly at typical cruise altitudes)*"

L276 "From Cohen et al., 2025b, we know that LMDZ-INCA transports aviation emissions to the UTLS region leading to accumulation there" I had a quick look at Cohen 2025b and failed to identify what you were referring to there. Fig 4 in that paper shows the NOx response to aviation emissions which does indeed peak in the UTLS region at around 50-70 deg, but then that's where it's deposited by aircraft at cruise altitudes, I don't see it being transported anywhere very different for LMDZ-INCA there. But maybe I've missed the part of that paper you're thinking of?

Have deleted this sentence and focused on the background sensitivity in this paragraph

L281 "This means that the model diversity in ozone RF response (Fig 4b) is smaller than the diversity in ozone mixing ratio response (Fig 4a)." Well maybe, but as they have different units, we can't really tell that from Fig 4 (but you might from your own calculations). If we had proportional/percentage, rather than absolute changes, in O3 concentration response and O3 RF response I guess we could then compare them. I suppose I'm just concerned that the implication is that we can draw that conclusion from the plots mentioned, when I think we can't.

The purpose of this was to highlight that the stratospheric differences in ozone concentration do not really affect the ozone RF, have updated this: *"This means that the modelled ozone RF response (Fig 4b) is mainly determined by the ozone concentration changes in this region (and not so much affected by ozone concentration differences in the upper stratosphere, Fig 4a)."*

L352 "The differences are largely attributable to the differences in aviation NOx emissions between these scenarios." Until I got to that sentence, I wasn't clear that the stats in the preceding sentence referred to SSP1 aviation in SSP1 background and SSP3 aviation in SSP3 background, given that you've covered all 4 permutations of background + aviation emissions thus far.

Have added *'(for the standard scenarios with consistent emissions and background).'*

L375: "as the ozone response to NOx is highly altitude dependent." and perhaps the O3 RF dependence on altitude is still greater (from the Skeie kernel), i.e. even if the amount of ozone generated by a flight didn't vary too much with altitude, the RF effect could still vary noticeably.

Thanks for highlighting this, have updated to *'the ozone response to NOx (and the resulting forcing in particular) is highly altitude dependent.'*

Tables S1, S2: good to have these, they gave me something to check the conclusions text against. But what about the corresponding 'present day' values?

Have updated Table S3 with present day ERF values from Cohen et al. (2025b)

Technical corrections (typos etc)

Abstract L20 "using three models" but later we have 4 models, e.g. L69 "Here we use MOZART3, LMDZ-INCA, OsloCTM3 and EMAC models..."; but when I get to section 4 I realise only 3 are used for background sensitivity experiments ("excluding EMAC"). Could allude to this in the introduction where the 4 models are introduced, though it's not wrong as it is.

Added to Table 2 caption *'Not all models ran all experiments; no background sensitivity experiments were run in EMAC.'* and line 97 in Methods *'EMAC (Pozzer et al., 2011, did not run background sensitivity experiments)'*

L64 "quantify [the] climate impacts" missing 'the'?

Added 'the'

L144 "(Sand et al. ,2023)." space/comma issue

Fixed

L166 "present-day" shouldn't be hyphenated as a noun (and at L170 has spurious space after hyphen)

Fixed throughout

Fig 3: to be honest I find it hard to keep in mind what the lighter vs darker colours and hatching vs no hatching mean; maybe it's because hatching="darker" doesn't fit with SSP1 (being weaker emissions than SSP3), when dark solid colour goes with SSP3 emissions? I know it's a pain changing figures, and it's not wrong. But if you change it, I would vote for showing hatching(=kind of darker) for "dirtier" SSP3 background emissions, not "cleaner" SSP1 background emissions. (Later: you are consistent though in using hatching for SSP1 background again.)

This is now consistent throughout (fixed Fig caption below), with lighter colours for SSP1 emissions, and with hatching or dashed lines for SSP1 background. Have also changed the hatching to white=weaker background which is more intuitive.

Fig 3: in fact the caption currently says "SSP1 and SSP3 (darker and lighter colours respectively)" which I think is the wrong way round.

Fixed, *'SSP3 and SSP1 (darker and lighter colours respectively)'*

L343 "Unger et al. 2013)" spurious parenthesis

Fixed, added bracket

L352 the "m-2" is currently broken across a line at the end, maybe meaning it is the wrong sort of dash/minus sign?

Fixed, changed to correct minus sign

Reviewer 2:

General comments

This is a welcome and interesting study that adds important new model results on the topic of aviation NOx and climate. My main concern is that the exact model experiment set-up is not sufficiently well described (in particular how methane is handled – and the consequences of fixing it). The authors conclude that it is important to let methane run free (I completely agree) so that its response can be better modelled. But in the meantime, these experiments where it is fixed need to be better explained so that the total impact of NOx emissions can be analysed and compared between studies. Some early studies that did allow methane to respond are not discussed, and that feels like an important omission. If this can be amended and the manuscript clarified as requested below, then I would be much more likely to be supportive of final publication in ACP.

Thank you for your time in reviewing this paper and bringing up some important points to clarify, which are addressed in the updated manuscript and are outlined further in the response below.

Specific comments

L86: Somewhere in the Methods section, it should be explained how methane is handled. I believe it is a fixed lower boundary condition, and the methane responses to NOx are estimated indirectly. This should be clarified at an early stage in the paper.

Yes, all models use a lower boundary condition for methane, added this to Methods: *'All models use a lower boundary condition for methane, i.e. a fixed concentration at the surface corresponding to present day or SSP1/3.'* and in the RF calculation section *'Since a methane lower boundary condition is used in these models, an offline method is required to estimate the* *resulting change in methane mixing ratio, ΔCH4 (and the associated radiative forcing, RFCH4) from the emissions perturbations in these experiments.'*

L105: Related to previous comment – if you only run for 4 years (and analyse last 3 years), then my question is: how is methane handled? (And the answer is: it is fixed).

See previous comment response

L115 Table 2: Document somewhere (probably in the is table) what the aviation NOx emissions, and the background (total) other NOx emissions, in each scenario/experiment are (in Tg N/yr).

The aviation emissions for the present day, SSP1 and SSP3 scenarios were in the text and have now added the background emissions too. They are in TgNO2/year as this is the default provided in CEDS/CMIP6. *"The aviation NOx emissions in the present day, SSP1 and SSP3 scenarios are 3.78, 2.86, 5.80 TgNO2 per year respectively, and the background emissions are 152, 64 and 164 TgNO2 per year respectively."*

Have also added Table S1 in the supplement to show all NOx emissions for the scenarios: *"Table S1: NOx emissions (TgNO2 yr-1) used in the different scenarios, for background and aviation. Corresponds to Table 2 in the main text but with emissions quantified."*

L141: Even if the "method" related to the "non-steady state factor" are described in Berntsen et al. 2005 (etc.), some summary explanation of the method is required here. This is clearly some fudge-factor to cope with having fixed methane, but needing to calculate a methane response.

See response below

L145 "…steady state is not reached…" I am unclear what this means, since methane is prescribed, so it won't respond in the experiments. What is meant by "methane steady state" in this context?

See response below

L146 Please explain what "overstates the response" means, and what non-steady state factors of 1.06 and 0.90 mean. I'm guessing if f_non-steady is >1 then assuming steady-state understates (rather than overstates) the response?

Response to the above three comments together:

The methane non steady-state factor is defined as the ratio

fnonsteady(t) = dCH4nonsteady(t) / dCH4steady(t)

The methane change at steady state is provided as described in the paper by:

dCH4steady(t) = dTCH4(t) * feedback * CH4(t)

The transient change in methane mixing ratio dCH4nonsteady(t) (necessary for comparing to other RFs) can be calculated by solving for instance equation (3) of Grewe and Stenke (2008) or based on a chemistry model as described by Lee et al. (2021). Based on the methodology described by Grewe and Stenke (2008) we derive non steady state factors for SSP1-2.6 and SSP3-7.0 future scenarios of 1.06 and 0.90 respectively. This approach is widely documented and discussed in the references included below.

Have added a more in depth description of this in the methods section:

"The methane feedback on its own lifetime ($f_{feedback} = 1.45$) is taken as the model mean from a recent model intercomparison (Sand et al., 2023). $f_{non-steady}$ is a factor to correct for the implicit assumption in using $f_{feedback}$, which assumes a steady state is reached (see discussion in Grewe et al., 2019; Lee et al., 2021). Due to its long lifetime, and slow feedbacks on lifetime via OH, a methane lifetime corresponding to steady state is not reached in any given year, and the methane concentration change resulting from NOx perturbations also depends on the recent history of the emissions of NOx, and of OH concentrations. For increasing NOx emissions, assuming steady state to derive the radiative forcing overestimates the methane concentration response (e.g. Grewe et al., 2019; Myhre et al., 2011), and vice versa for decreasing emissions. Based on the method described by Grewe and Stenke (2008) (solving their Equation 3) we recalculated these factors for 2050 conditions for both SSP1-2.6 and SSP3-7.0 future scenarios and derived non-steady factors of 1.06 and 0.90 respectively."

L150 Does equation [1] dimensionally balance? Please clarify the units of each term. I'm confused because I thought delta-CH4 was a change in mixing ratio, and the f factors were dimensionless. But then also the change in lifetime must have units of time, so something is wrong…

Addressed in a previous comment from Reviewer 1: Yes indeed dTCH4 is the relative change in methane lifetime (%), have clarified in the text *'based on the modelled relative change in methane lifetime (%)'*

L155 Equation [2]: be explicit about the units. Are concentrations in ppbv, and RF in W/m^2?

Have clarified units *'From this methane mixing ratio change (in ppbv), the methane RF (in W m^{-2}) is calculated'*

L161 Presumably this conversion of RF -> ERF is also highly uncertain? Please comment on this further source of uncertainty.

This was discussed further down (in the discussion) but has been moved to the Methods section: *'These ratios are used due to absence of better information – the ERF/RF ratios are assumed to be equal to efficacy factors from a single study and are therefore associated with a high uncertainty (Ponater et al., 2006).'*

L166 It is not entirely clear what "short-term" refers to. In the introduction, you mentioned that "long-term" referred to 5-10 year timescales. So, is short-term <5 years? Or are you equating long-term with the methane-related response, and short-term with everything else? Please clarify.

Long-term refers to the methane-induced response, which is calculated offline here (due to the methane LBC), and short-term therefore refers to the response modelled online, via changes in NOx concentrations and therefore ozone precursors. This has been clarified in the text (and previous reference to 5-10 years deleted).

L 183 Related to previous – here you do define short-term with respect to methane – please say earlier.

Addressed above

L185-187 I wonder if part of the explanation of the model differences is related to model vertical resolution (EMAC has most levels; LMDZ has fewest levels).

The vertical model resolution contributes to the inter-model variation, as shown for water vapour transport and lifetime in Van't Hoff et al. (2025) (Fig 3 and Fig A5 https://acp.copernicus.org/articles/25/2515/2025/#&gid=1&pid=1). However, this isn't the only reason for the differences, or a full explanation e.g. ozone response is smaller in OsloCTM3, but OsloCTM3 has more levels than LMDZ-INCA. Have added a sentence in the text: *"Varying vertical resolution is a contributing factor to the vertical distribution of the model response (e.g. van 't Hoff et al., 2025), with LMDZ-INCA and EMAC having the fewest and most levels respectively."*

L191: Clearer to say overestimate UT and underestimate LS

Updated to '*They found that the models overestimate UT and underestimate LS ozone concentrations'*

L200 Figure 2. (a) I think a better axis label would be something like "Change in the (short-term) O3 burden due to aviation". (b) Suggest add zero line (axis). (d) Clarify caption – normalised by change in annual total NOx emissions?

Have updated these as suggested

L216 Do you also look at the long-term O3 (via CH4) response?

This section focuses only on the short-term ozone response, which is modelled directly in these experiments, so we are able to do detailed analysis on the spatial distribution and background conditions. The methane–induced ozone response is estimated and discussed in the next section, mainly to enable comparison of net NOx estimates with other studies. The calculation of long-term ozone RF relies on the estimation of methane concentration change (via methane lifetime change), and a constant relating methane concentration difference to ozone RF, which has limitations as discussed.

L222 Document the TgN changes (see comment on Table 2).

Addressed in previous comment, have added *"(see Table S1 for emissions)."* here.

L223 Are these global annual mean?

Clarified as *'annual mean global short-term ozone RF'*. Also updated Fig caption

L229 Clarify aircraft emissions

Added *'aircraft emissions'*

L242 Do you know for sure that the differences due to the model grids being difference are definitely "small"? Perhaps replace small with "some", unless you have tested this and shown they really are small.

small -> some

L250 Figure 3. Clarify these are annual mean responses. Can you clarify how seasonal variations in response are averaged, and that seasonal variations do not contribute in any way?

The model output is monthly, and the annual average is a mean of all the months. There is some seasonal variation in the background sensitivity, in particular with LMDZ-INCA showing positive or negative sensitivity depending on the season. Have added these figures to the supplement and added some text: *"Here we show the annual mean response, but there are also some seasonal differences in the sensitivity, e.g. LMDZ-INCA in summer with SSP3 emissions shows a negative sensitivity to SSP1 background, consistent with OsloCTM3 (see Fig S1)."*

For 3(b) can you comment on how the number of levels contributes to the sharpness of the response in these vertical profiles? (LMDZ-INCA, with the fewest levels, appears to have the broadest response).

See response to comment above also related to vertical resolution - here have added: *"The lower vertical resolution in LMDZ-INCA may contribute to its more (vertically) broad ozone response."*

L251 …higher changes in ozone concentration due to aircraft…

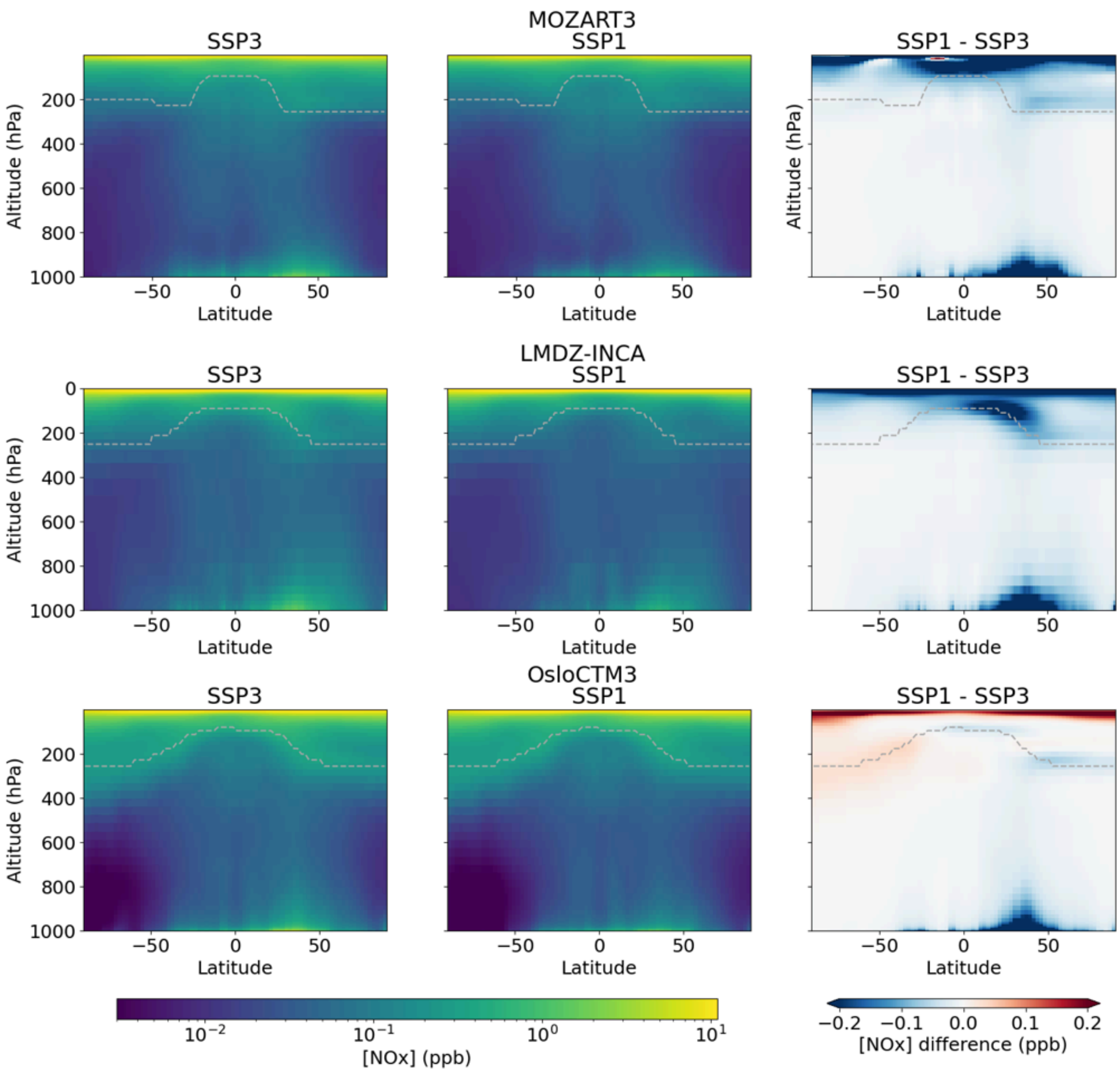
Updated to *'a higher ozone concentration change or RF due to aircraft emissions, i.e. a higher ozone sensitivity'*

L253 ..lower changes in ozone…

Have updated this according to above comments from Reviewer 1 *'The models all show reduced ozone production from aviation NOx under an SSP1 (vs SSP3) background in the northern mid latitude upper troposphere, and enhanced production in the LS under SSP1 (vs SSP3) background (Fig 4a).'*

L261 Figure S4 – I think I'd rather see background zonal mean NOx in SSP1/3 rather than the difference; and a log scale may be useful as I guess there are big differences between changes near the surface and in the UTLS.

Zonal mean NOx background for SSP1 and SSP3 below for all models - I'm not sure this adds any information as they all look very similar/any differences are much smaller than the range to be visible in the colour bar (even with log scale). The focus here is on the difference in background conditions and the sensitivity to that.



L270 Figure 4. I was quite confused by this figure – is it the difference between scenarios of the difference in O3 due to aircraft emissions? Please clarify if it is this.

Updated the figure caption to add *'Difference between scenarios (SSP1 vs SSP3 background) of the response to a 20% reduction in aviation emissions'.* This is also clarified in the text in response to comments above (L253)

L277 "a higher background response" – Clarify what is meant. Is it a higher response to varying the background, or a larger change in the background?

Updated to *'a higher sensitivity to background'*

L279 Do you mean less transport of surface emissions to the UTLS?

Meant to refer to the contrast between these two models and LMDZ-INCA - that they don't show the strong ozone response in the LS. Have updated this to clarify that we're referring to stratosphere-troposphere exchange. *"OsloCTM3 and MOZART3 do not show the strong positive ozone response in the LS region, which may be due to more similar background NOx in the UTLS region in SSP1 and SSP3, and/or due to less stratosphere-troposphere exchange (than in LMDZ-INCA)."*

L283 Would it be sensible to normalise the responses? (to the magnitude of the change in aviation emissions)

Have updated Fig 4c and d, normalised to the magnitude of change in NOx emissions, which shows a higher sensitivity to background with SSP1 aviation emissions in two of the models. Updated the sentence in the text: *"In terms of vertical profiles, the SSP1 emissions experiment gives a larger ozone RF response to changing background in MOZART3 and LMDZ-INCA (Fig 4c,d), showing that in these models, the sensitivity to background NOx is higher with lower initial aviation emissions, consistent with a more NOx-limited regime."*

L304 Clarify by 'long-term' ozone, do you mean the ozone related to changes in methane?

Yes, have added *'long-term (methane-induced) ozone'*

L309 Figure 5: By "long-term" do you mean time-integrated to infinity (or 100 years) or what? (Maybe infinity and 100 years are the same).

As above, have added '*(methane-induced)'*

I suggest remove "Net" from the y-axis label (as net only applies to the dots, and it is in the legend).

Removed *'Net'*

If error bars were able to be included on the net terms, would these all span zero? (I suspect so).

Yes, some of the models give net negative forcing already for SSP1 and these are very sensitive to the underlying assumptions - with uncertainty coming from both components already included e.g. methane forcing, and additional uncertainty due to elements not included here such as NOx impact on aerosol forcing (discussed further down in the text). Have added a sentence to outline this: *"The high uncertainty and strong sensitivity of the NOx terms to underlying assumptions (discussed further below) mean that the uncertainty in the net NOx RF is large, and net negative values of forcing cannot be excluded, as also shown in more comprehensive assessments e.g. Lee et al., (2021)."*

L330 The mysterious "steady-state factor" turns out to be rather crucial to your end results… hence the need to clarify what this is earlier.

Yes, indeed, this is very critical as discussed here and in previous work (Lee et al 2021, Skowron et al. 2020, Myhre et al. 2011, Grewe et al. 2019). Please refer to Lee et al. (2021) appendix for a description of the method (currently being updated in Bellouin et al. in prep). Have clarified this in the text in responses to previous comments above.

L338 and L385 I totally agree that methane emission-driven models are important. With this in mind I find it odd that some of the early work on these topics that did use methane emission-driven models (e.g., Wild et al., 2001; Stevenson et al., 2004) is not discussed.

Thanks for highlighting this, had added some text here: *"Early studies using chemistry transport models to study aviation NOx impacts did include methane emissions and therefore a had more interactive ozone-OH-methane representation, including the short and long term ozone response, and the methane response (Stevenson et al., 2004; Wild et al., 2001). Both studies found an overall small net positive RF from aviation NOx perturbations. Stevenson et al., (2004) highlighted a complex seasonal variation in the tradeoff between the short and long term responses, which affected the sign of the net NOx RF."*

Technical comments

L84-88 (and throughout): capitalize "section" references.

Done

Table 1: The in-table references (Price et al., etc.) are missing from the reference list.

Added all table references

L343 Unger et al. (2013)

Added bracket

References

Stevenson, D. S., R. M. Doherty, M. G. Sanderson, W. J. Collins, C. E. Johnson, and R. G. Derwent (2004), Radiative forcing from aircraft NOx emissions: Mechanisms and seasonal dependence, J. Geophys. Res., 109, D17307, doi:10.1029/2004JD004759.

Wild, O., M. J. Prather, and H. Akimoto (2001), Indirect long-term global radiative cooling from NOx emissions, Geophys. Res. Lett., 28, 1719–1722.