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.

Great, thanks. I'm re-reviewing changes against my comments only, I have not re-read the entire manuscript again from scratch but trust that reviewer 2 will check their changes similarly.

I've highlighted my responses in bold where I think some action, or at least checking, is still required.

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).'

OK thanks

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'*

OK

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 (%)'*

Sorry to be super-picky, but I wouldn't specify *percent*, because we don't want anyone to be confused with a possible factor of 100; it's just a unitless proportion

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."*

OK thanks, I haven't been through the maths, but understand and appreciate the useful qualitative description

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'*

Ah OK, thanks

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.'*

OK thanks except new typo thing: should have space before units in "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)."*

OK thanks, that's fine now. Originally it did say "in the upper troposphere, where the NOx is emitted" and that's what I was objecting to really, because aircraft cruise in the LS and not the UT if you are reasonably far north.

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.'*

Nearly there but one of us has got the polarity wrong! I reckon (clearest in LMDZ-INCA) that the ozone production is *increased* (enhanced) in the UT and *reduced* (diminished) in the LS.

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)"*

OK thanks

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

OK

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)."*

OK that's a nice understandable message now.

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).'*

OK

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.'*

OK

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)

OK thanks

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)'*

OK

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

Added 'the'

OK

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

Fixed

OK

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

Fixed throughout

OK everywhere now

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.

OK great thanks

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)'*

OK

L343 "Unger et al. 2013)" spurious parenthesis

Fixed, added bracket

Obviously not a biggie, but this one (missing parenthesis I should have said) not actually fixed in tracked changes/revised version

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

OK but my apologies, missed this last time, earlier in the same sentence you have the units "mWm-2" without a space or dot after the "mW", inconsistent with elsewhere. Hardly worth mentioning!

Reviewer 2: [removed for my re-review pass as reviewer 1]