



Editorial Office

Atmospheric Chemistry and Physics

November 14th, 2023

Dear Dr. Querol and reviewers,

Revised submission of “Global impacts of aviation on air quality evaluated at high resolution” to *Atmospheric Chemistry and Physics*

Thank you for your time and effort in considering our manuscript. We greatly appreciate the comments made by the reviewers, and have made significant efforts to address their concerns.

We recognize that a broad concern expressed by the reviewers regarded the robustness and interpretation of our findings. In response we have now performed an additional set of simulations, re-evaluating the response of surface air quality to aviation emissions but now including additional sensitivity scenarios and diagnostic data which have allowed us to dig more deeply into the root causes of the differences we first observed. These include:

- A full-length simulation using an updated version of the base model, to evaluate sensitivity to bulk changes in our understanding of atmospheric science and emissions;
- An additional series of tracer transport tests to provide quantitative evidence for claims regarding vertical mixing times between seasons; and
- Two simulations extending through northern hemispheric winter (the time of peak exposure to air quality changes resulting from aviation emissions) seeking to evaluate:
 - The degree to which NO_x emissions alone can be said to cause these impacts (removing the possibility of near-airport primary PM_{2.5} as being the driving cause); and
 - The relative contribution of landing and take-off emissions (below one kilometer altitude).

Due to the high computational cost of these simulations we have conducted our sensitivity analyses at C90 (~100 km) resolution, which our earlier simulations showed are sufficient to resolve ~98% of air quality impacts associated with aviation in a higher-resolution (C180, ~50 km) simulation. In each of these simulations we have archived and analyzed additional data in order to better understand and communicate the impacts of aviation on air quality.

In addition, and guided by the reviewers' comments regarding the need for more scientific insight into aviation emissions impacts, we took this opportunity to archive a greater variety of information from the additional simulations. This has allowed us to evaluate: the long term fate of aircraft emissions compared to their products; where, and the degree to which, ozone production and loss rates are changed by aviation emissions; and the potential contribution of differences in background atmospheric composition at different resolutions.

Finally, we realized during this review process that an additional post-processing step was applied internally when simulating the effect of eliminating aviation emissions over the North American region, but that this step had not been described in the manuscript. This step means that emissions were eliminated over US land specifically (at all altitudes) rather than in a simple block domain. This does not substantially affect the meaning of the manuscript, which is focused on the effect of resolution and the relative contribution of near-airport emissions. However, it does increase the relative contribution of in-domain emissions. Nonetheless the central point – that 72% of aviation-attributable impacts on US air quality come from out-of-domain emissions – is unaffected. We have also changed the manuscript discussion throughout to reflect this, more accurately referring to the “NA” domain as instead the “US” domain.

We believe that the changes we have made to the manuscript have resulted in a substantial improvement. In addition to these bulk changes, we have also made specific additional changes to address each of the reviewer

comments. A response to each reviewer comment is given below; the original comments are shown in **bold** with our responses shown in *italics*.

Referee #4

The authors use the GEOSChem model to run 1-year simulation with and without aviation emissions to attribute mortalities to the changes in ozone and PM_{2.5}. They use different resolutions and calculate the differences aviation related mortalities. In general the paper is well written, however, a deeper analysis that tries to explain differences in the modelling results due to resolution is lacking. Moreover, I have severe doubts concerning the method of an incremental approach that is applied in this case. Is the sum of all incremental approaches for all individual sources, such as industries, households, transport, aviation, etc. giving the total mortalities due to either ozone and PM_{2.5}? Thunis et al. (2019) and the references therein clearly show that this is not the case. They have a very strong argument that incremental approaches can not be used for source (and here mortality) attribution. I think addressing these points is crucial before the paper may become accepted.

We thank the reviewer for their comments and their thorough review. We have taken these concerns seriously and have committed significant resources to ensure that we are able to provide quantitative data which more completely analyze the underlying causes of our findings. We have also worked to ensure that the goal of our approach is not conflated with a broader assessment of source attribution, but rather focused specifically on the question of the avoidable impacts of aviation, what fraction of these impacts result from in- or out-of-region emissions, and the degree to which assessments of these quantities might be sensitive to model resolution. A detailed response to each comment is provided below.

I have some doubts whether the approach in calculating aviation related mortalities is applicable. Around line 160: "This is the quantity for which Turner et al (2015) determined a 12% increase (95% confidence interval: 6.0-18%) in respiratory mortality per 10 ppb increase in ozone exposure." To my understanding this analysis is based on situations where an enhancement of ozone can be attributed to an increase in an emission source. However, here, we have a constant emission of NO_x by aviation that competes with other sources of NO_x, such as industry, lightning etc. In that case Thunis et al. 2019 (and the literature cited therein) clearly pointed out that an incremental approach, with and without aviation emissions is not applicable. If the mortalities due to ozone are estimated with your approach for all individual emissions separately and added up would that give the same number as for all emissions? It is a highly non-linear system, so they answer is not (see also Tunis et a.). I think that the uncertainties due to the used method are much larger than due to different resolutions

We fully agree that an approach aimed at understanding the fraction of current air quality degradation which is attributable to different sources, sectors, or regions might find a different answer. However, our objective here is to quantify a) what the potential improvement of air quality might be if aviation emissions could be reduced to zero, b) what fraction of this benefit can be achieved in one region (the US) through actions taken only in that region, and c) the degree to which these quantities are sensitive to model resolution (and therefore the degree to which policy assessments relying on a coarse-resolution model might over- or under-estimate potential benefits). As discussed elegantly by Clappier et al. (2017), the choice of which approach to use is sensitive to the question being asked. For example, it is entirely possible to find that two different sectors could, if eliminated independently, reduce air quality mortality by 35%, but that performing each together could achieve a total of only 50% (as discussed in e.g. Thunis et al. (2021)). This does not change the fact that eliminating a single sector would still reduce total air quality impacts by 35%, but it does make it important for readers (in particular policy makers) to understand the role of non-linear interactions. We now explicitly call out this need for nuance and sensitivity to context in the discussion (lines 512-517).

The analysis of the impact of different resolutions lacks quite some analysis.

a) How are regional pattern e.g. around major airports changed? and why?

We now include a discussion of this effect as part of Section 3.1 (lines 283-293) which shows how concentrations of black carbon are increased local to US airports in such a way that simulation at C180 is more accurate, but that the changes in ozone and PM_{2.5} are less strongly localized (see Figure R 1 below, which is now also included in the SI as Figure S2). As illustrated by the relatively small differences in mortality between simulations at C90 and C180, simulations at these resolutions demonstrate qualitatively similar patterns of changes in PM_{2.5} and ozone but neither are dominated by proximity to large airports. The same is not true for black carbon (soot), which shows strong peaks near to major airports (e.g. the hot spot near LAX and Santa Monica Airport) which are poorly resolved

at coarser resolutions. However, even in this hotspot soot makes up less than 3% of total $PM_{2.5}$ exposure, meaning that the overall health impact calculated for aviation emissions is not significantly affected by the transition from C180 to C90. We do however now highlight that soot specifically is more accurately captured at C180, such that studies focused on near-airport air quality (as opposed to regional or global air quality) would benefit for very high resolution simulations which are not plausible for global analyses.

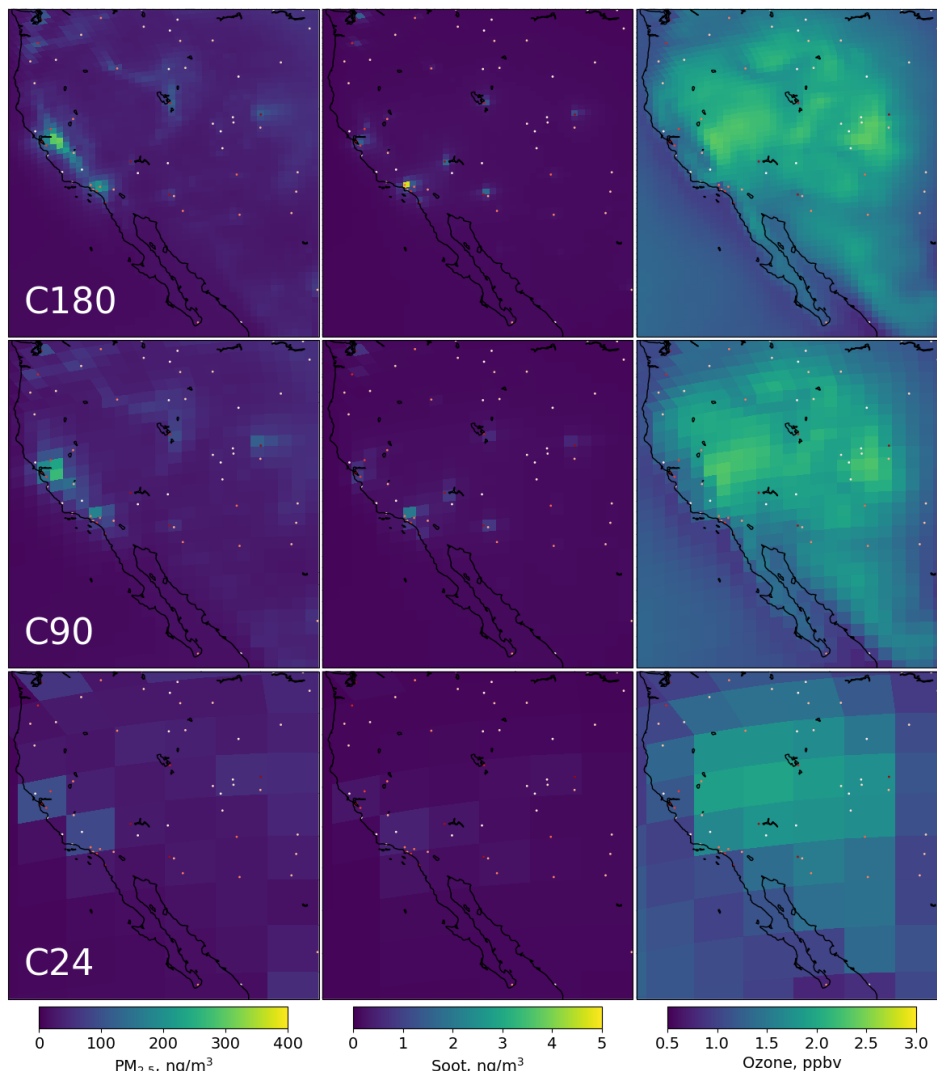


Figure R 1. Changes in annual mean $PM_{2.5}$ (left), soot (center), and MDA8 ozone (right) over the US west coast due to aviation simulated at C180 (top), C90 (middle), and C24 (bottom) resolution. Red dots indicate airports which served at least 100,000 passengers in 2015 according to the US Bureau of Transportation Statistics from T-100 segment data.

b) How is the transport changed due to resolution changes? Convective transport, lightning? Scavenging?

Bulk (advective) transport is in theory unchanged, as we are using a chemistry-transport model where meteorology is read in as a set of externally-imposed fields. However we still expect some effect due to (e.g.) the inability of a coarse model to resolve local eddies, including convective motions which may have only been resolved at fine spatial and/or temporal resolution. These phenomenon are discussed specifically in the context of this model in Yu et al. (2018).

For scavenging it is possible that there is some resolution dependence, although the cloud data (including cloud fraction, water content, and precipitation rates) are again pre-calculated in the meteorological data. As such we would not expect significant non-linearity. Unfortunately we do not have scavenging data available from our original simulations, which would otherwise have allowed us to evaluate the degree to which (e.g.) aviation-attributable NO_x is washed out differently at different resolutions.

Finally, lightning emissions are calculated internally based the cloud convective depth and flash density. These fields are archived at a single, common resolution in order to minimize differences in lightning emissions between resolutions. As such, no significant differences are expected between simulations at different resolutions. This is now stated in the Supplementary Information.

We now discuss these components of resolution dependence in in Section 3.3 (lines 384-401 specifically), which is dedicated to providing insights into both the base mechanisms by which aviation affects surface air quality and the reasons that these mechanisms might be different at different resolutions.

c) Where is actually ozone production and loss rates changed?

We now evaluate changes in ozone production and loss rates as part of Section 3.3. The relevant text is reproduced below:

As shown in Figure R 2, the vertical profile of ozone averaged from 15° to 45°N, in addition to odd oxygen production rates, is affected by aviation emissions during each season. Aviation-attributable surface-level ozone is maximized during winter and minimized during summer. This is true at all altitudes below 10 km, whereas background ozone (rightmost panel) is greatest during summer over the same range of altitudes.

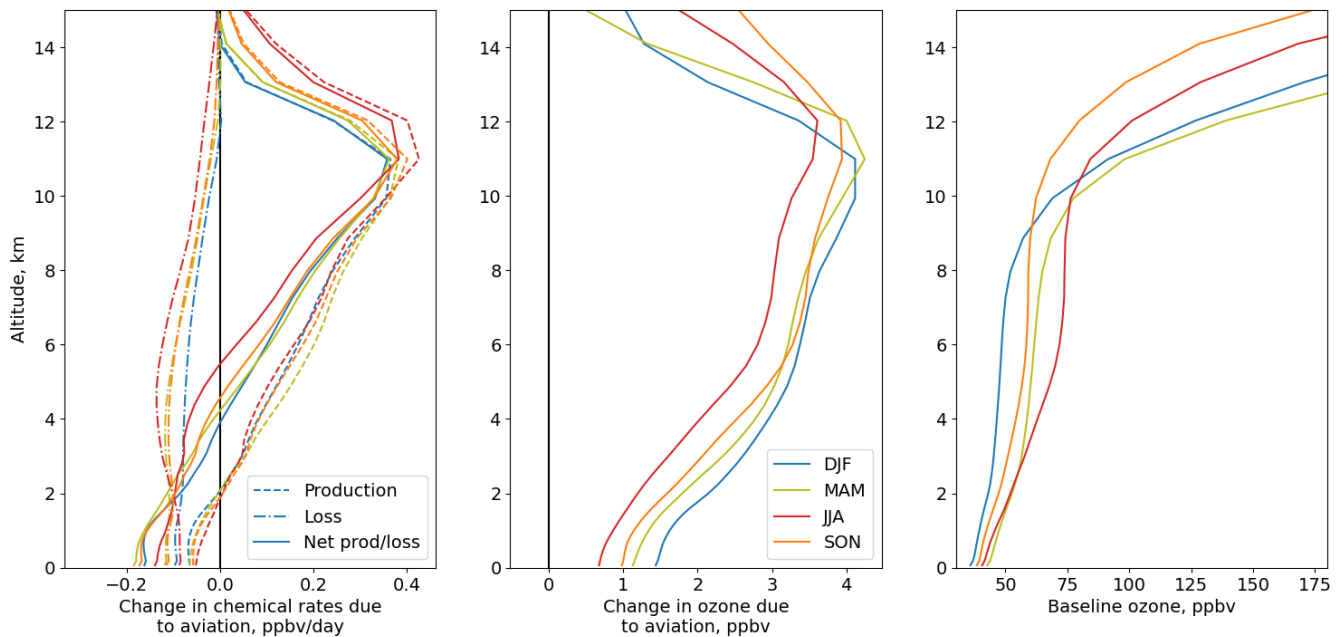


Figure R 2. Changes in production and loss rates of odd oxygen (left) and the net effect on ozone mixing ratios (centered) along with baseline ozone mixing ratios (right), separated by season. Values are for the average from 15°N to 45°N at C90 resolution. This is now Figure 6 in the main text.

We attribute the difference to two causes. First, the change in odd oxygen production rates in the free troposphere is consistent throughout the year, while odd oxygen loss rates are reduced during the winter. This results in longer lifetimes for aviation-attributable ozone without a decrease in production.

The second factor is transport. Figure R 3 overlays the aviation-attributable change in ozone with the mean pressure (vertical) velocity in Northern Hemispheric winter (DJF) and summer (JJA). As discussed by e.g. Williams et al. (2019), the contribution of stratospheric air to surface ozone is controlled by both the degree to which upper tropospheric air is enriched by stratospheric ozone and the rate of vertical mixing from the upper troposphere to the surface. The former factor is not relevant to aviation’s influence, which as discussed above shows relatively little variability between seasons in this latitude range although aviation-related increases in ozone are greater north of 45°N. The latter factor of increased transport rates, however, does appear to be more significant. Downwelling motions in Northern midlatitudes during winter months allow ozone to descend from cruise altitudes to reach the surface. When combined with reduced ozone loss rates, the result is an increase in the amount of aviation-attributable ozone reaching the surface during winter.

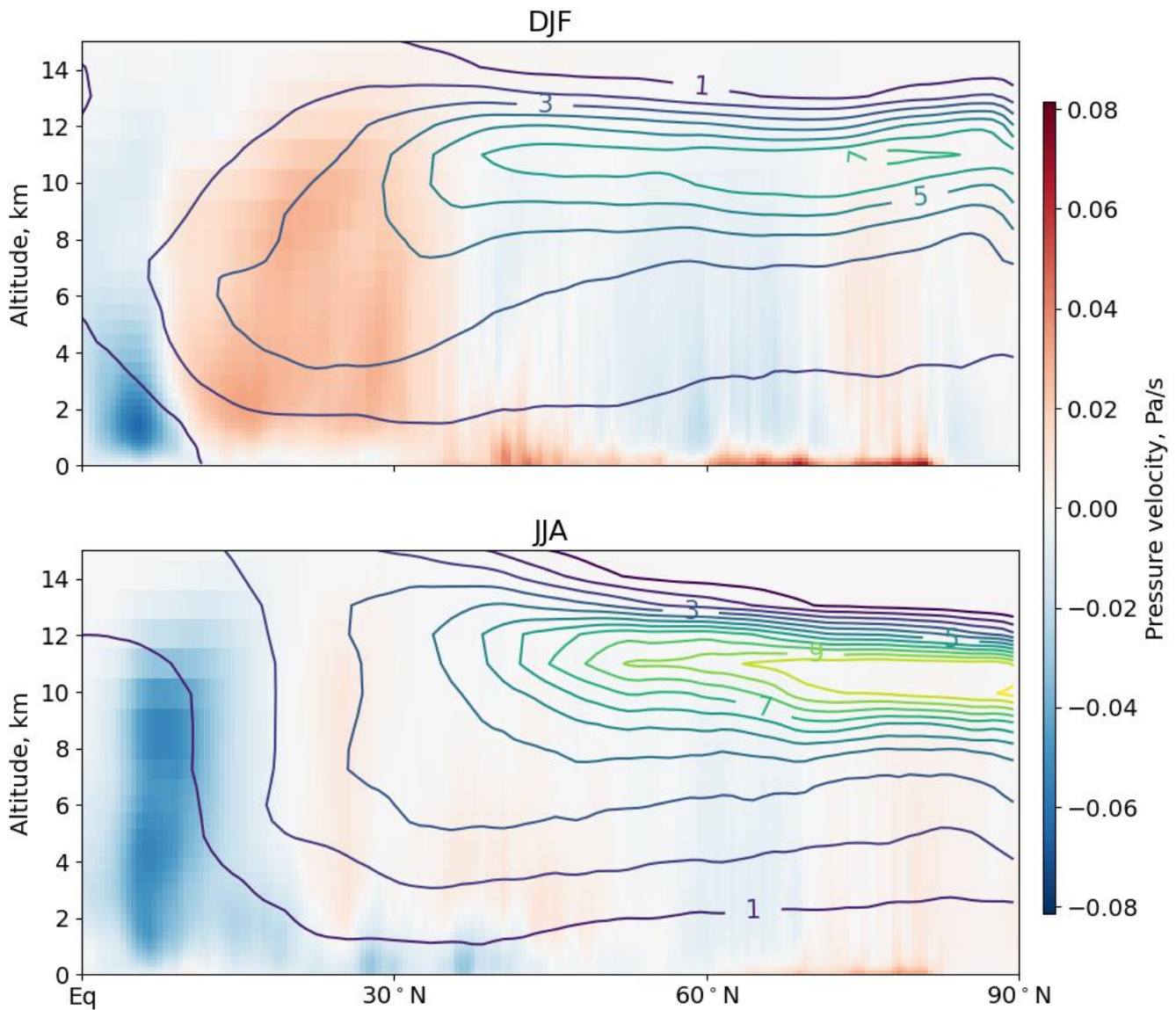


Figure R 3. Seasonal mean pressure velocities (shading, Pa/s) and aviation-attributable changes in ozone (contours, ppbv). Blue colors indicate an upwelling motion, while red colors indicate a downwelling motion. Each ozone contour corresponds to a change of 1 ppbv. This is now Figure 7 in the main text.

I think the intercomparison with other studies mostly relies on the same model, right? What about other aviation simulations for NO_x and aerosols?

Several prior studies of the impacts of aviation on global premature mortality have indeed used GEOS-Chem, including (for example) Barrett et al. (2010), Eastham and Barrett (2016), and Quadros et al. (2020). Yim et al. (2015) found comparable results when using the CMAQ model in local regions, but arguably their work still relied on GEOS-Chem for boundary conditions. However, there are other studies which found similar results without using GEOS-Chem. The global simulation performed in Vennam et al. (2017) found changes in surface-level pollutant concentrations which were within the range of the aforementioned studies. Similarly, the multi-model intercomparison by Cameron et al. (2017) found agreement among multiple chemistry transport models (including GEOS-Chem) regarding the magnitude of the change in surface-level ozone and PM_{2.5} resulting from aviation emissions. Studies by Lee et al. (2013) and Phoenix et al. (2019) used the Community Atmosphere Model (CAM) versions 3 and 5 respectively to assess the impacts of aviation on surface air quality, finding again that aviation emissions cause increases in surface ozone of the order of 1 ppbv during winter with the dominant contribution from cruise-altitude NO_x. These studies are now cited in the manuscript on lines 441-450.

Eq. 1 and 2 (in addition to the above comments): I am wondering, although knowing that this approach has been used frequently in the past, if this eq. is actually well representing the aviation induced mortality. We have many sources for PM2.5 and ozone such as households, industry etc. My understanding is that the mortality from all sources is M_Base. Hence the sum of all mortalities dM_i from individual sources i (aviation, households, fires, ...) should give M_Base.

$$M_Base = \sum_i dM_i = \sum_i M_Base * (RR_i - RR_Base) / RR_Base = M_Base / RR_Base * \sum_i (RR_i - RR_Base)$$

However, in general, the sum of all changes of the surface concentration and exposure is not base concentration or exposure, respectively. (see e.g. Thunis et al. 2019 and the literature therein). Thunis et al clearly indicated the inaccuracy of incremental approaches. Emmons et al. (2012) showed a differences of a factor of 2-4 for a surface source in using contribution and incremental approaches. Local conditions the differences might be even larger. And the nonlinear function in eq. (2) is adding to the discrepancy. Hence, to my understanding, here the mortality changes by switching off aviation is investigated and not the impacts from aviation. Please make this clear in the text and title.

We agree with the assessment that we here perform a calculation of the number of mortalities which could be avoided if aviation emissions were to be brought to zero. We also agree that, for studies which seek to attribute all air quality impacts to different sectors, other approaches (such as tagging) may be more appropriate. However our intent here is to quantify the degree to which avoidable mortalities result from current-day aviation emissions, and therefore the benefit which might be achieved if those emissions could be eliminated (either through policy means or technological means, such as the post-combustion emissions control discussed in Prashanth et al. (2020)). To make this as clear as possible, we now discuss this nuance both in the introduction (lines 81-83) and the discussion (lines 512-517), stating unequivocally that the calculations performed here relate to the context of the benefit which is theoretically achievable by acting on aviation emissions in isolation. In the discussion in particular we now give the following statement:

"This work describes the potential benefit of a rapid reduction in aviation emissions, and the degree to which different modeling approaches can accurately capture the expected outcome of such a reduction while accounting for non-linearity in the response. Approaches which are intended to perform an attribution of current-day air quality or mortality impacts between different sectors, regions, or species using methods such as tagging (Emmons et al., 2012; Butler et al., 2018) may find different results. As discussed by Clappier et al. (2017) and Thunis et al. (2021a, b), the relevance of these results to planned policy will therefore depend on the context and objectives of the policy."

Eq. 1/2 Are you evaluating the exposure and mortality on the basis of, e.e., daily values and then averaging or are you using the mean values in the first place? How large are the differences between the two approaches? So day-by-day variations in the transport pattern and e.g. chemical responses might be large enough to considerably change the exposure as a mean.

We calculate mortality based on the annual-average value of PM_{2.5} in µg/m³, and the annual-average value of the maximum daily 8-hour average ozone concentration in ppbv. This is necessary as the epidemiological data supporting the relationship between exposure and mortality correlates an increase in long-term average (chronic) exposure with mortality risk. We do not consider the effects of short-term exposure, for which a different set of concentration response functions and epidemiological data would be required. We now explicitly state this distinction in the methods (line 174-176).

213 "This provides additional evidence" I wouldn't call a missing correlation as an evidence. I suggest to rewrite by using a statement such as "not contradicting to ..."

We agree with this assessment. The wording has been changed to "This is consistent with the hypothesis that..." (line 255).

228 "The greater resolution has two effects: physical phenomena are more finely resolved .." While I agree in general with the statement, I think there is more analysis required to understand the impact of resolution on physical processes. Are you sure that the simulation is more realistic? How is the vertical transport changed? Do you have an analysis of convective up- and downward transports that are changed. Is the dynamical lifetime affected, and if, by which process. Are there measurements to allow a judgement on the quality of the higher resolved processes? E.g. 222Rn? Are natural processes and sources changed? e.g. lightning NOx? I think a much deeper analysis is required to allow a more detailed judgement.

We now include a discussion (Section 3.3) which focuses on understanding first the reason for the changes we observe (see discussion in response to the comment regarding ozone production rates), but also the likely effect of a change in resolution on process representation. This includes not only a concise review of relevant literature with regards to the representation of convection and transport at different resolutions with this specific model (GCHP), but also the results of additional tracer transport simulations which we have conducted.

With regards to changes in natural sources, lightning emissions are calculated from flash rates and convective depths which were archived at the native resolution of the meteorology specifically in order to avoid changes in resolution affecting emissions. We also use pre-calculated emissions for natural sources of dust, biogenic VOCs, and sea salt. Emissions of iodine from sea salt are still calculated online but are assumed to have a negligible effect on aviation-attributable changes in air quality. These are detailed in the Supplemental Information although we would be happy to move this description to the main text if preferred.

246 "The relative contribution of NA aviation emissions" please rephrase and see above the comment based on Thunis et al.

We have now clarified that these calculates reflect the expected change if aviation emissions are brought to zero (lines 303-304).

Sect. 3 in general: The simulation is based on one year. The year 2015 is known for intense heat waves. Is there any information available on the robustness of the results, e.g. annual variability, etc.

We now cite in the main text (lines 481-485) estimates of meteorological sensitivity from three previous studies of the effect of aviation emissions on surface air quality, each of which evaluated the effect of the meteorological year. Barrett et al. (2010), Eastham and Barrett (2016), and Quadros et al. (2020) each evaluated the effect that a change in meteorological year would have. Barrett et al. (2010) found that the choice of year and meteorological product could change estimated mortalities by up to 21%. Eastham and Barrett (2016) found that, across five different years, global exposure to PM_{2.5} and ozone could vary by up to 5% due to the specific choice of meteorological year. Meanwhile Quadros et al. (2020) simulated how a change in the simulation year (reflecting changes in both background emissions and meteorology) from 2005 to 2013 would affect exposure, finding a net increase of 6.6 to 12% depending on the region of interest. We cite the range 5 to 21% as a conservative estimate.

Intercomarison: As far as I see most citations refer to the use of GEOSchem. right? What about other models? In general amore detailed explanation is missing.

We focus our comparison on previous GEOS-Chem-based studies as this allows us to more exactly separate the roles of different changes. However, we do now include a comparison of our estimated changes in surface concentrations against the multi-model intercomparison performed by Cameron et al. (2017) and against the global results from Vennam et al. (2017). We do not include a comparison against the nested results from Vennam et al. (2017) as this work suggests that the exclusion of changes in large-scale atmospheric composition in that work was the primary cause of their simulated reduction in ozone and PM_{2.5} concentrations when performing a fine-resolution calculation. However, we do now refer to simulations of aviation's air quality impacts performed using other models (notably CAM 3, CAM 5, CMAQ, GEOS, GATOR-GCMOM, and GISS Model-E2) (lines 441-450).

Referee #1

After reviewing this manuscript, I do not think in its present form it is appropriate for ACP. The manuscript runs a chemical transport model at different resolutions to determine the change in near surface ozone and 2.5 micron Particulate Matter (PM_{2.5}), and then presents 'health effects' as premature mortality. I think too much of the uncertainty and analysis is due to epidemiological uncertainty which really cannot be assessed in ACP. If the paper were to focus on the physical modeling it would be fine: but there is zero uncertainty estimated there. I think the manuscript needs to be a much better analysis of the physical modeling for ACP. If the focus is to be on the health effects and uncertainties, it should go in a health focused journal. I do not like the use of headline numbers of thousands of deaths, when there is no uncertainty in the abstract, values have changed by a factor of 5 from previous estimates, and some of the work is based on a single health study. As a physical scientist, I am not able to assess whether these are valid or not, and hence I do not think ACP is appropriate. So ideally, this would be put in a health oriented journal. If the authors wish it to be in ACP, I think it should focus on the changes to Ozone and PM_{2.5}. Specifically there needs to be a better assessment of uncertainty, which probably means using more than one meteorological year.

We appreciate this perspective and have worked to address it. Firstly, we believe that there is unfortunately a need to find a common metric for impacts from different pollutants which necessitates the translation of exposure, which is species-specific, into health risks, which are not. The use of premature mortalities has also historically been supported by ACP, as evidenced by the ongoing and frequent publication of studies in ACP which quantify impacts from PM_{2.5}, ozone, or both in their abstract, including several studies since 2021 alone (Zhan et al. 2023; Nault et al. 2021; Geels et al. 2021; Sun et al. 2021; Zhang et al. 2021; P. Wang et al. 2023; Tarín-Carrasco et al. 2022). The specific concentration response functions which we have used have already been applied in the context of aviation (Quadros, Snellen, and Dedoussi 2020) and we do include evaluations of our results – and implications – if one assumes different concentration response functions. However we recognize that this needs to be more clearly presented in the paper, and now highlight in the abstract that this result is sensitive to the choice of CRF (line 28). We also now clearly state the 95% confidence interval in impacts resulting from uncertainty in the health response in both the abstract and discussion and close our abstract with a statement regarding uncertainty in the physical response to aviation emissions (line 29).

We have also made a concerted effort to improve and extend our assessment of physical uncertainty. We have performed an additional set of simulations in which we use the same set of aviation emissions with a different model version, including substantial new updates to the treatment of tropospheric chemistry (see Section 2.3). Using information from this new simulation set and from literature data we now explore the degree to which different models, different versions of the same model, different meteorological years, and different model resolutions can affect estimated exposure to both ozone and PM_{2.5} and therefore health risks, with a new section (3.4) dedicated to the question of physical uncertainty.

Finally, we have substantially expanded our overall analysis. Previously we included two sections in the results: a combined analysis of the effect of resolution and in-region vs out-of-region emissions (3.1), and a comparison to previous work (3.2). We have now expanded this into five sections. Section 3.1 is dedicated only to quantifying the effect of resolution change and includes an evaluation of the role of near-airport emissions. This has allowed us to separate out the role of in-region vs out-of-region emissions in Section 3.2. Section 3.3 introduces a new analysis of the chemistry and fate of aviation emissions by season, based on the new C90 (~100 km) sensitivity simulations which were performed in response to this review. It also includes an analysis of the potential causes of differences between simulated impacts, based on both the results of our model simulations at different resolutions and information available from the literature, and discusses the potential causes of our findings regarding in-region vs out-of-region emissions (lines 360-372). Section 3.4 investigates two additional sources of physical uncertainty, while also placing our observed effect of a change in model version alongside literature evidence for model-based uncertainty. Finally, we have expanded Section 3.5 (formally 3.2), the comparison to previous work, to include a larger set of studies using models other than GEOS-Chem to evaluate aviation's effects on surface air quality. These are described in more detail through our point-specific responses below, but we hope that the reviewer agrees that we have significantly deepened the relevance and depth of our manuscript thanks to their suggestions.

I also do not think the 'different resolutions' is particularly interesting since the meteorology is just interpolated to drive the model, and because the scale of resolutions is still within the range of using exactly the same parameterization methods. This might also not be a wise focus for the manuscript as it is not particularly strong.

We respectfully disagree with this assessment but have worked to make the argument more compelling in the manuscript itself. One of the motivations of our work was the realization that in this model (as in other chemistry transport models), the change in resolution can significantly affect the result of the calculation specifically decrease there being no change in underlying physical parameterizations. For example, the treatment of convection within the model is based on that of the parent GCM – however the parent GCM is resolving a much larger fraction of convection in its large scale transport fields. Since this is not recovered when the meteorology's resolution is degraded, convection can be under-represented in coarse CTMs (see eg. Yu et al. (2018), now discussed on lines 384-393). However atmospheric scientists have limited resources and cannot afford to perform all their simulations at the parent model resolution; for example, the GEOS-Chem High Performance CTM used in this research requires around 35,000 CPU hours per simulation year if running at C90, but around eight times that per simulation year at C180. Finding that, in contradiction to previous work, a C90 simulation is adequate to resolve ~98% of impacts resolved at C180 (but that a much cheaper C24 simulation is not) enables researchers to more efficiently explore the causes of – and potential solutions to – aviation's environmental impacts. In light of the comments here we have endeavored to make this clear by deepening our investigation of the differences in physical modeling between simulations at different resolutions (Section 3.3).

Specific comments:

Page 1, L27: What is the uncertainty on these numbers? It seems that the uncertainty is entirely due to epidemiological factors? How much is due to ozone changes? Any of it? I am not convinced that this is useful at all given that air quality is highly non-linear and the extremes are not well reproduced by any scale of models 50-400km. Are the ozone numbers different? Suggest that things be focused around ozone and NO_x quantification rather than the epidemiology.

We agree that extremes are not well represented by global-scale models. However, the change in surface ozone concentration due to aviation emissions is diffuse, as shown in Figure 2 of the main text, and analyses such as Pungler and West (2013) have concluded that estimates of health impacts from exposure to ozone and premature mortality can be captured in models at these resolutions. We now perform a deeper analysis, however, of the response of atmospheric composition (rather than just health outcomes) to emissions in the model. In particular, the new section 3.3 uses a series of sensitivity analyses to quantify how ozone changes at different altitudes, the role of chemical lifetime compared to vertical transport, the degree to which this varies between seasons, how much of the air quality response is due to cruise altitude emissions versus landing and takeoff, the role of NO_x compared to other aviation emissions, and (where possible) the sensitivity of these factors to model resolution.

Page 2, L56: Why so much lower? Can you summarize?

Based on the results from this study, we believe that the results from Vennam et al. (2017) reflected the fact that their nested simulations employed fixed boundary conditions. The use of fixed boundary conditions would mean that aircraft emissions being advected out of the domain could not influence surface conditions. Since our results show that US emissions cause increases in ozone which are hemispheric in nature, this oversight in Vennam et al. (2017) likely resulted in the loss of the key mechanism by which US aviation emissions could affect US surface air quality. This is highlighted on lines 501-506 of the main text.

Page 2, L62: Global total from aviation?

Correct – we now specify this explicitly in the text (line 65).

Page 3, L72: can you explain how you can equate climate impact with air quality impact? This seems not to be a scientific question.

We believe that this is a policy concern which requires scientific input. Decision makers such as the FAA and ICAO cannot perform monetization and prioritization of emissions control strategies (a political concern) without accurate quantification of both the climate and air quality impacts of aviation emissions (a scientific concern). This is of particular relevance now due to ongoing debate regarding whether to prioritize reduction in NO_x emissions (motivated by air quality concerns) or improvement in fuel efficiency (motivated by climate concerns), with a key deciding factor being uncertainty regarding the true air quality impacts of NO_x. We cite the need for accurate quantification as part of the motivation of this paper but we do not claim to provide a mechanism for comparison between climate and air quality impacts.

Page 3, L79: All of these resolutions are using basically the same type of model set up. What would be the impact of using a more detailed treatment that would explicitly resolve the non-linear nature of exposure? Are there parameterizations for this from regional air quality models? That would be more interesting.

We agree that it would be interesting to also evaluate how and why regional air quality models find different results than global-scale models. However, one of our goals in this work is to understand why a single, consistent global model might find substantial differences in impact when only a single parameter – the grid resolution on which chemistry and transport are calculated – is changed. The fact that we find a 20% increase in estimated impacts as a result of this change alone suggests that it is an important factor in understanding disagreement between prior studies, and should be considered alongside factors such as differences in the parameterization of different physical processes between models. We now state this explicitly in the discussion on lines 495-497.

Page 3, L81: Please explain what the coded resolutions refer to the first time it is used.

We now do not introduce the “CN” resolution names until they can be properly explained.

Page 3, L94: Are the in and out emissions impacts linear? E.g., if you take Global - NoNA + NoNA - Off, do you get the same answer as Global - Off?

Since we only perform three simulations (with global aviation, with all aviation except that over the US, and without any aviation), we cannot directly evaluate the linearity of the response. This would require an additional simulation in which we only include aviation over the US (say, “NAOnly”) – at which point we could evaluate the degree to which the effects of “NAOnly – Off” differ from the effects of “Global – NoNA”.

Page 4, L98: why do you say CN for resolution and then use C?

The N in this case is intended to correspond to the number, hence the phrasing “Model resolutions are therefore denoted as CN where N is the number of grid elements along each edge of the cube” (lines 101-102).

Page 4, L104: What happens is you try to run at much higher resolution (say 5-10km). Can you do this, even for a limited time to gauge the physical impact on O3 and NOx? Using the same chemical mechanism. Versions of the MERRA-2 system have been run at these resolutions.

We did investigate the run time required to do this, as we agree it would be a valuable experiment. However, the combination of the high computational cost of the chemical integrator in GEOS-Chem and the need for at least a minimal spin-up period made this prohibitive. For context, a simulation at ~100 km resolution (C90) on 192 CPU cores takes 12 hours to complete one month. Even with perfect scalability and assuming no decrease in the model time step, a simulation at 10 km globally would require 100 times the resources (i.e. 230,400 CPU hours) to complete one month. The previous work with MERRA-2 at these resolutions was completed with an older version of GEOS-Chem using NASA resources, and required 3.5 million CPU hours to complete a year (Hu et al. 2018). We unfortunately do not have access to the necessary resources to consider such a simulation at this time.

Page 4, L119: Are the same MERRA2 data used and just averaged or binned to lower resolution? How high resolution does the MERRA-2 data go? What if you used one of the GEOS ‘nature runs’ at 3 or 7km?

We perform area-conservative regridding (averaging) over MERRA-2 data on each model level, preserving the full vertical resolution (72 levels) but reducing the horizontal resolution as necessary (Eastham et al. 2018; Martin et al. 2022). As discussed above we unfortunately do not have the computational resources to perform a simulation at 7 km resolution. However we have applied what resources we have available to produce the sensitivity simulations discussed above which we believe have helped to provide valuable additional insight into how and why aviation affects air quality, the degree to which resolution affects the simulated answer, and the role of in-domain versus out-of-domain emissions.

Page 6, L144: How large are the biases? If they are larger than the signal, then could the non-linearities impact the results?

It is hard to say whether biases in the base simulation could cause a systematic effect in the results, hence the caveat provided here. Since aviation emissions cause a relatively small change in surface concentrations, it is difficult to develop strong empirical constraints on which the specific ability of a model to simulate aviation-

attributable change in surface conditions can be evaluated. We now attempt to clarify this point in Section 3.4 while also providing additional context for some of the sources of uncertainty which could not be quantified.

Page 6, L153: Given that the model is uncertain and you are dealing with small perturbations, can you give an uncertainty estimate for (a) the difference in concentration due to aviation and (b) the uncertainty in your mortality estimates?

We now explicitly state in the abstract that any model-based estimate of the physical response of the atmosphere to aviation emissions is inherently uncertain (lines 28-30). However, as specified above we now include both an assessment of the degree to which coarse-resolution modeling introduces errors (end of Section 3.3) and limited exploration of potential sources of uncertainty (Section 3.4), alongside an expanded comparison of our results against those from other studies including those using other atmospheric models (Section 3.5).

Page 6, L165: You base the ozone values on one study? That doesn't seem appropriate. There must be others you can refer to?

We agree that more studies are needed to investigate the impacts of chronic exposure to ozone. We use Turner et al. (2015) as our core study as it remains the most recent epidemiological study of the health impacts of chronic exposure to ozone, was previously identified as a key study for understanding the potential global burden of disease resulting from ozone exposure (Malley et al. 2017; Seltzer et al. 2020) as well as aviation specifically (Quadros, Snellen, and Dedoussi 2020), and was itself based on a high-quality cohort (that from the American Cancer Society Cancer Prevention Study-II). The other major study of health impacts resulting from chronic exposure to ozone that we are aware of is Jerrett et al. (2009), which is based on the same cohort and appears to be superseded by Turner et al. (2015). This reflects the relative lack of epidemiological data for ozone exposure, which is of critical importance for this topic given that aviation appears to have a nearly unique influence on northern hemispheric ozone. Accordingly, we now state (lines 437-438) that “[g]iven that there are relatively few studies of the health impacts of chronic exposure to ozone compared to exposure to PM_{2.5}, this uncertainty is a key area of future research”.

Page 7, L172: Here you describe the uncertainty in mortality rates: why is this not propagated through, especially to the abstract.

This was an oversight, and uncertainty bounds are now given in both the abstract and the discussion. We also explicitly state in the abstract that the uncertainty quantified is only due to uncertainty in the health response.

Page 7, L182: Again, these uncertainty estimates need to be in the abstract.

This is now the case.

Page 7, L185: The uncertainty interval for PM2.5 is ±8%? How different is that than a previous estimate? That seems ridiculously low. Also: what is the uncertainty due to the physical model difference in PM? This uncertainty is only due to epidemiology right?

We agree that this was misleading as it reflects only one potential source of uncertainty. We now explicitly state that this is the confidence interval due to uncertainties in the health response, and include a specific discussion (Section 3.4) of other sources of uncertainty in the atmospheric (rather than health) response to aviation emissions.

Page 7, L186: The ozone mortality estimated uncertainty is ±30%. Yet the difference from previous work is 560%! Is that entirely due to differences in the epidemiological assumptions?

Yes – as outlined above, we have endeavored to make clearer to the reader the difference between epidemiological and physical uncertainty.

Page 7, L191: Figure 2 is hard to interpret due to use of the default python color scale. Maybe get rid of anything which is not significant (make those points white, not a dominant blue): you need a definition of what is a significant difference and what is just noise. I suggest the variability either within a year, or between multiple years could give you a standard deviation here.

Due to the prohibitive computational cost of our simulations we only have one year of simulation results at C180, although we do now include discussion of (for example) the uncertainty associated with different meteorological years based on prior evaluations in the literature. However, since these results are generated using a chemistry

transport model, chaotic climate feedbacks are not present in our results. As such there is not a convenient definition of significance with regards to the change in concentrations.

Page 7, L191: Why would the O₃ changes from aviation AT THE SURFACE be largest in the W. US and Tibetan Plateau? Does that really make sense? I suggest there is a transport problem with the model due to a terrain following coordinate and anomalous cross-isentropic transport to elevated topography. This is a common problem with transport schemes in global models.

This is a well-established result for aviation emissions, present in (for example) Barrett et al. (2010), Vennam et al. (2017), Phoenix et al. (2019) and most results presented in the multi-model intercomparison of Cameron et al. (2017). A reproduction of the result from the hemispheric-scale model used by Vennam et al. (2017) is given below, showing the elevated response over the Western US and Tibetan Plateau.

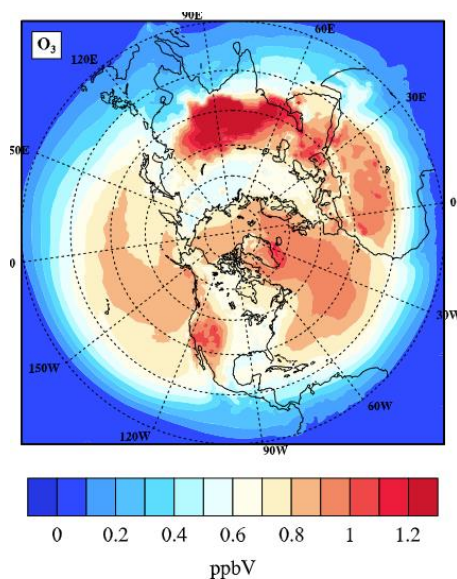


Figure 1. Surface ozone change calculated by Vennam et al. (2017) showing the same distinctive pattern of surface ozone change.

The effect of stratospheric ozone intrusions increasing surface ozone in the Western US has also been extensively discussed in the literature outside of the impact of aviation (Lefohn et al. 2011; X. Wang et al. 2020; Lin et al. 2012), and the same is true for the Tibetan plateau (Yang et al. 2022). We concur with those prior researchers that, rather than this being a pure model artifact, it reflects the fact that these locations are more strongly influenced by air from the upper troposphere and lower stratosphere which will in turn have been more strongly affected by aviation.

This is one of the interesting components of aviation's impacts which we believe makes it worthy of study, as the impact of the aviation sector on near-surface atmospheric composition is essentially unique compared to other industries.

Page 7, L193: So these particulates are secondary aerosols produced where there already is air pollution?

This is exactly correct.

Page 8, L205: This seems anomalous in the model. Does observed near surface air in the Western US have higher background O₃ than elsewhere? Representative of transport of higher altitude air to the surface?

Precisely. A particularly good discussion of this is given in Lin et al. (2012, 2015), where they show that surface ozone in the Western US is strongly affected by high altitude ozone, especially during the springtime. As a result, aviation's effect on surface ozone is also maximized in such locations and during the spring – because the high altitude air where aviation has the largest effect on ozone is being effectively carried to the surface. We now cite Lin et al. (2012) in the main text at this location (lines 236-237).

Page 8, L218: This begs the question: are the epidemiological effects linear? Does a 10 ng/m³ change in PM_{2.5} have the same impact if the background is 10ng/m³ (100% increase) or 100ng/m³ (10% increase)?

No, the epidemiological effects are not linear. For both ozone and $PM_{2.5}$ we use a non-linear concentration response function as detailed in Section 2.2. We therefore account for the baseline concentration in our calculation, as detailed in equation 1 of the main text. However the change in concentration associated with aviation is sufficiently small (of the order of 1% of baseline concentrations, as discussed by Lee et al. (2013) and Phoenix et al. (2019)) that there would not be expected to be a significant nonlinearity of the concentration response function between 0% and 100% of current aviation emissions.

Page 9, L220: Given the importance of secondary particulate matter, how uncertain is this production in the model? It seems as if this is one of the most uncertain elements.

We agree that the model's representation of secondary particulate matter is a key uncertainty and now highlight this in Section 3.4. On one hand, GEOS-Chem incorporates state-of-the-art modeling of secondary inorganic aerosol thermodynamics and partitioning, including nitrate aerosol, as represented by the ISORROPIA-II model (Fountoukis and Nenes 2007) with continuous updates and improvements from a community of dedicated researchers such as the implementation of nitrate photolysis (Shah et al. 2023), hydroxymethanesulfonate chemistry and its relationship to sulfate aerosol (Moch et al. 2020), and the entrainment and removal of NO_y in clouds (Holmes et al. 2019). On the other hand this remains an area of active and ongoing research, and uncertainty in overall nitrate aerosol production in particular remains high. We now call this out specifically on lines 420-426.

Page 9, L238: are these absolute or relative values? If relative they are small, but if absolute they are large. And it's confusing given the the previous percentages if they are relative.

All differences are relative. We now state (line 266) that "All differences are given relative to the aviation-attributable change calculated at C180 unless otherwise stated".

Page 10, L252: see earlier comment: are these linear with the total effect over the US?

As discussed in response to the prior comment they are not; we account for differences in the baseline concentration from cell to cell.

Page 11, L260: Figure 5: similar color scale problem to figure 2: you should blank regions without significant changes.

Since we are using a chemistry transport model, we do not have a meaningful way to quantify significance (and do not have the resources to run additional years at the resolution shown here).

Page 12, L281: So you get exactly the same answer? That's suspicious. Are you using the same exposure and same modeling tools to do this? What is the difference in their ozone and $PM_{2.5}$?

In this case we believe that it is a combination of a genuine result (more fuel burn is likely to mean more impact, after all) and coincidence. Quadros et al. (2020) do use a closely related modeling tool, but in their case they apply a nested approach rather than using a globally consistent resolution. They also use the same exposure assessment approach. There are certainly differences however, as demonstrated by the fact that our application of an updated version of the same model can result in a 15% change in estimated mortalities (explicitly called out in Section 3.4). Quadros et al. (2020) also do not find an identical split between ozone and $PM_{2.5}$, and this is now stated in Section 3.5 (line 431).

Page 12, L287: I think this needs to be in the main text.

Given the increased length of the article following these responses, we are reluctant to move more material to the main text. As such we have opted to keep the more detailed evaluation of sensitivity to CRF in the Supplemental Information for now.

Page 12, L299: how much is attributable to any physical differences in the change in $PM_{2.5}$ and Ozone? Initial background state and model biases?

Since both models in this comparison are chemistry-transport models rather than climate models, we assume that the role of the initial background state is negligible. However we do now include a comparison of the physical response calculated in our model to that from a range of other studies (lines 441-450), showing that the estimated

change in surface-level ozone and PM_{2.5} is within the range reported from other studies (per TgN emitted). We also have a new section dedicated to the underlying physics and chemistry of aviation air quality impacts, including potential causes of the changes we observe between resolutions and between in-domain and out-of-domain emissions (Section 3.3).

Page 13, L322: PM_{2.5} emissions are not due to NO_x are they? Maybe you need to remind the reader that NO_x emissions are what change ozone since NO_x is not mentioned in the results section.

We are grateful to the reviewer for highlighting the fact that we did not explicitly point out in the results section the role of aviation NO_x. We now include a brief recap of the fact that aviation NO_x has been shown in previous studies to be the dominant contributor to changes in surface-level ozone and PM_{2.5} (Whitt et al. 2011; Prashanth et al. 2022; Eastham and Barrett 2016) on lines 329-332. We have also performed a short (6 month, covering the critical winter season) sensitivity simulation in which we include all aviation emissions except for NO_x. We find that the aviation-attributable wintertime increase in ozone and PM_{2.5} is decreased by 108 and 103% respectively if aviation NO_x is not emitted. This is now stated on lines 373-375 of the main text.

Page 13, L328: I think this is disingenuous since you have really just limited the resolution differences to interpolation of input data. A proper assessment of resolution would alter the balance between parameterized and resolved quantities, but you do none of that and test only a limited range of resolutions and then apply them to a much higher resolution (1km) population data set. Which seems a big scale mismatch.

It is true that there is a large scale mismatch between the resolution at which these global assessments are performed and the resolution of the population dataset. We now include a statement in the discussion caveating the fact that our findings are specific to the question of how differences in resolution might affect global-scale assessments within a single model (lines 495-497).

Our objective is to help researchers and policy makers understand the different elements which cause differences between investigations of the same phenomenon, one of which is that different models are run at different resolutions. We also hope to help them to allocate resources, as our findings indicate that a researcher interested in understanding the global-scale impacts of aviation on surface air quality could run eight simulations at ~100 km for the cost of running one at ~50 km and still get results which match to within 2.5%.

Page 13, L332: Why would inconsistent models yield a smaller difference? Is that chance? I would expect a larger difference perhaps?

We realize that the emphasis in this paragraph incorrectly implied that inconsistency in model would inherently result in a smaller difference. We have rephrased the paragraph to instead emphasize that the causes of differences between our results and those in Yim et al. (2015) are different to the causes of differences with the results in Vennam et al. (2017). We also now emphasize that the differences found by Yim et al. (2015) which they associate with resolution could also be explained by model inconsistency, given the range of results reported by Cameron et al. (2017) in a model intercomparison. By contrast, the results of Vennam et al. (2017) are not within this range, and appear to be better explained by the use of a fixed boundary condition. We now explore this issue in more detail in the results on lines 360-372.

Page 14, L335: This is difficult to follow. Are you saying that Venam et al had non US aviation emissions in both? Wouldn't this tend to reduce changes 'neglecting the larger response'? So why is the response larger? I don't follow the logic.

This paragraph has been rewritten as described above.

Page 14, L351: You lost your uncertainty range again. Please put it here and in the abstract.

We apologize for this oversight. Uncertainty ranges are now included in the conclusions and abstract.

Page 14, L352: How much uncertainty results from the different meteorology in 2005 v. 2015?

We now cite previous estimates of the effect of meteorology on exposure to ozone and fine particulate matter resulting from aviation emissions.

Referee #5

This manuscript presented a modeling analysis of global health impact associated with ozone exposure due to aviation emission. Impacts of anthropogenic emissions on air pollution have been thoroughly investigated over inland areas, yet the contribution from aviation remains poorly documented. The study applied a solid modelling tool and focused on an interesting topic. But there are two critical main issues need to be addressed before the acceptance could be considered.

We appreciate the constructive feedback and have endeavored to address the two critical issues raised by the reviewer.

First, the manuscript proposed a very interesting question at the introduction section but didn't mention it in the discussion or conclusion. Line #57 mentioned: "whether the air quality impacts of aviation are dominated by local sources or are the result of larger atmospheric changes.", and line#63 mentioned "there questions urgently need to be resolved". So the introduction indicated this is one of the question that would be at least discussed in this study but unfortunately no such discussion was mentioned later. Section 3.2 provided a detailed comparison of the estimated mortalities between this study and previous studies, but the most significant difference was due to using a more epidemiological data, while the proposed question remain unsolved.

We agree that the original manuscript did not spend enough time addressing this central question – of the degree to which local versus global aviation influences surface air quality. We have separated out and expanded a new section of the results dedicated to this question (Section 3.2, "In-region versus out-of-region emissions"). This section first focuses the reader on the results which show that emissions from US aviation contribute roughly proportionally to surface-level ozone across the Northern Hemisphere, rather than having an effect which is mostly isolated to the region of emission. We now also include a discussion exploring why this is the case, which focuses on the simulations in which aviation emissions over the US are excluded. By subtracting those results from the simulation in which all aviation emissions are included, we observe how atmospheric composition is affected by US aviation emissions only. For the purposes of this response we have performed some additional analysis (below); although this is not included in the main text, it informed our discussion on lines 360-372 of why in-region emissions do not dominate aviation emissions impacts.

The change in mean ozone between 15 and 45°N for local winter and summer shows two key features (Figure R 4). First, we find that the overall cruise-altitude increase in ozone over the US is actually greater during summer than during winter. However, we also find that it is more localized in both longitude and altitude.

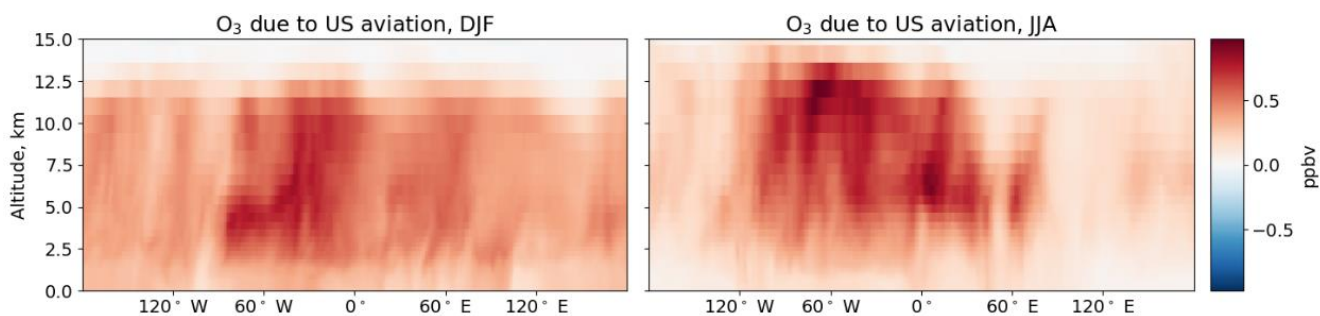


Figure R 4. Changes in mean ozone between 15 and 45°N (area-weighted) during northern-hemispheric winter (DJF, left) and summer (JJA, right) due to aviation over the US. Results are from the sensitivity simulations conducted at C90, for which additional data were archived.

As discussed in the response to Referee #4, this is likely due to the wintertime combination of longer lifetimes with faster vertical transport from cruise altitudes to the surface. The net result is that aviation has a nearly unique signal. The regional influence decreases by 38% from winter to summer, such that the effect of US aviation on US ozone decreases from +0.35 ppbv in the winter to +0.22 ppbv in the summer. Meanwhile, the effect of US aviation on the rest of the world is decreased by 79% over the same period, with the population-weighted mean exposure to ozone falling from +0.27 ppbv in winter to +0.056 ppbv in summer (Figure R 5).

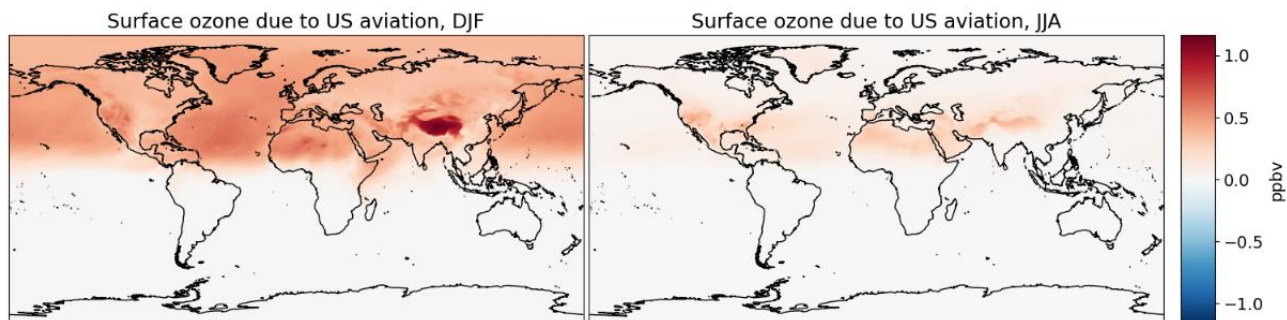


Figure R 5. Change in surface-level ozone worldwide due to US aviation in winter (DJF, left) and summer (JJA, right). Results are here from the core simulation at C180.

This effect is exceeded by the greater overall effect of aviation on surface level ozone during winter, when the impact is more diffuse. When all aviation is simulated, surface level ozone over the US is increased by 1.6 ppbv during winter compared to 0.64 ppbv during summer (a 60% reduction). This implies that US aviation is responsible for 21% of aviation-attributable exposure in the US during winter, and 34% during summer. However, since the overall summertime exposure is lower, the wintertime effect – where global aviation is responsible for 79% of the surface ozone change in the US - is dominant.

For $PM_{2.5}$, we find a similar outcome with a larger in-region signal. During summertime, US aviation is responsible for an additional 21 ng/m^3 of US $PM_{2.5}$, 55% of the total effect from global aviation. During wintertime, the absolute contribution of US emissions increases to 49 ng/m^3 , but the relative contribution falls to 31%. As in the case of aviation-attributable ozone, in-region emissions have a larger relative effect during summer, but the largest absolute contribution occurs during winter and is the result of out-of-region emissions.

This finding has significant implications for policy makers. Regardless of the season, out-of-region emissions are the dominant contributor to in-region impacts, even for a region with as much fuel burn and covering as much physical space as the US. Similarly, exposure to both $PM_{2.5}$ and ozone are greater during winter, when the contribution of in-region emissions is reduced. However, in-region emissions have a greater influence on $PM_{2.5}$ than on ozone, and (due to the fact that baseline summertime ozone exposures are typically greater) few studies have considered whether an increase in wintertime ozone might have less severe health consequences than an increase in summertime ozone. If indeed seasonal changes in ozone exposure are found to be important, then a focus on regional aviation emissions may still have disproportionate benefits for public health.

Second, as a modeling study the manuscript lacks a necessary evaluation section. Almost all discussions were made based on simulation results, so it is very important to demonstrate the reliability or remaining uncertainty of the simulation results. Without solid evaluation the rest of discussion regarding contributions of aviation to air pollution and associated mortalities would be difficult to interpret.

GEOS-Chem has been extensively evaluated for its ability to reproduce global observations of both atmospheric composition and air quality, and has also been widely used in air quality assessment. However, we recognize that these results still reflect only those of a single model. We therefore provide both citations of recent assessments and a statement to this effect on lines 143-150. We have also endeavored to improve and extend our comparison to include results from models other than GEOS-Chem (see Section 3.5).

There are a few other minor issues with the organization of the manuscript. For example, line #80 mentioned simulations were conducted for 400, 100, and 50km and section 3 demonstrated the differences between. Apparently finer resolution can better reproduce atmospheric chemistry especially for ozone which is sensitive to mixing of NO_x and VOCs. The manuscript indeed shows a large difference between fine and coarse simulations, but this seems more like a technical improvement other than a key innovative science finding.

We respectfully disagree with the assessment that this is purely technical, but recognize that the manuscript as it was previously organized did not sufficiently highlight the scientific findings. We hope that the changes made in response to the comments above help in this respect, but we have also worked to try and more prominently highlight the following findings:

- *First, that aviation's impacts on air quality are global in nature and not dominated by near-airport emissions. This is now more specifically illustrated through our analysis of the relative role of near-airport and regional-scale air quality (lines 283-292), and our quantification of the relative role of cruise-altitude NO_x emissions (as opposed to non-NO_x emissions (line 375) and LTO emissions (line 376)).*
- *Second, the finding that in-region emissions of NO_x from aircraft do not cause the majority of air quality impacts associated with aviation (Section 3.2), and that this is largely the result of seasonal differences in ozone chemistry and transport (Section 3.3).*
- *Third, that the effect of aviation emissions on global air quality is greater than previously estimated and dominated by ozone – making it almost unique in comparison to other sectors (opening of Section 3).*
- *Finally, the finding that moderate-resolution simulations (~100 km) are sufficient to capture impacts that are resolved at 50 km, but that low resolution (~400 km) simulations are likely to underestimate impacts (Section 3.1).*

In addition to the above points, we believe that this manuscript helps to address a standing point of confusion in the literature – that being whether very high resolution simulations would be expected to significantly reduce the simulated impact of aviation. This finding was reported by Vennam et al. (2017) and was a possible explanation for the results of Yim et al. (2015). Our work provides a plausible explanation for both outcomes, while also demonstrating that the hemisphere-scale increase in surface-level ozone and PM_{2.5} associated with cruise-altitude aviation is unlikely to disappear as model resolution continues to advance.

Thank you for considering our submission for *Atmospheric Chemistry and Physics*. We hope you agree that these comments have been addressed and look forward to your response.

Regards,

Sebastian Eastham

References

- Barrett, Steven R. H., Rex E. Britter, and Ian a. Waitz. 2010. "Global Mortality Attributable to Aircraft Cruise Emissions." *Environmental Science & Technology* 44 (19): 7736–42.
- Cameron, M. A., M. Z. Jacobson, S. R. H. Barrett, H. Bian, C. C. Chen, S. D. Eastham, A. Gettelman, et al. 2017. "An Intercomparative Study of the Effects of Aircraft Emissions on Surface Air Quality." *Journal of Geophysical Research, D: Atmospheres* 122 (15): 8325–44.
- Eastham, Sebastian D., and Steven R. H. Barrett. 2016. "Aviation-Attributable Ozone as a Driver for Changes in Mortality Related to Air Quality and Skin Cancer." *Atmospheric Environment*. <https://doi.org/10.1016/j.atmosenv.2016.08.040>.
- Eastham, Sebastian D., Michael S. Long, Christoph A. Keller, Elizabeth Lundgren, Robert M. Yantosca, Jiawei Zhuang, Chi Li, et al. 2018. "GEOS-Chem High Performance (GCHP V11-02c): A next-Generation Implementation of the GEOS-Chem Chemical Transport Model for Massively Parallel Applications." *Geoscientific Model Development*, no. 11: 2941–53.
- Fountoukis, C., and A. Nenes. 2007. "ISORROPIA II: A Computationally Efficient Thermodynamic Equilibrium Model for K^+ - Ca^{2+} - Mg^{2+} - NH_4^+ - Na^+ - SO_4^{2-} - NO_3^- - Cl^- - H_2O Aerosols." *Atmospheric Chemistry and ...*, 4639–59.
- Geels, Camilla, Morten Winther, Camilla Andersson, Jukka-Pekka Jalkanen, Jørgen Brandt, Lise M. Frohn, Ulas Im, Wing Leung, and Jesper H. Christensen. 2021. "Projections of Shipping Emissions and the Related Impact on Air Pollution and Human Health in the Nordic Region." *Atmospheric Chemistry and Physics* 21 (16): 12495–519.
- Holmes, Christopher D., Timothy H. Bertram, Kaitlyn L. Confer, Kelly A. Graham, Allison C. Ronan, Charles K. Wirks, and Viral Shah. 2019. "The Role of Clouds in the Tropospheric NO_x Cycle: A New Modeling Approach for Cloud Chemistry and Its Global Implications." *Geophysical Research Letters* 46 (9): 4980–90.
- Hu, Lu, Christoph A. Keller, Michael S. Long, Tomás Sherwen, Benjamin Auer, Arlindo Da Silva, Jon E. Nielsen, et al. 2018. "Global Simulation of Tropospheric Chemistry at 12.5 Km Resolution: Performance and Evaluation of the GEOS-Chem Chemical Module (V10-1) within the NASA GEOS Earth System Model (GEOS-5 ESM)." *Geoscientific Model Development* 11 (11): 4603–20.
- Jerrett, Michael, Richard T. Burnett, C. Arden Pope, Kazuhiko Ito, George Thurston, Daniel Krewski, Yuanli Shi, Eugenia Calle, and Michael Thun. 2009. "Long-Term Ozone Exposure and Mortality." *The New England Journal of Medicine* 360 (11): 1085–95.
- Lefohn, Allen S., Heini Wernli, Douglas Shadwick, Sebastian Limbach, Samuel J. Oltmans, and Melvyn Shapiro. 2011. "The Importance of Stratospheric–Tropospheric Transport in Affecting Surface Ozone Concentrations in the Western and Northern Tier of the United States." *Atmospheric Environment* 45 (28): 4845–57.
- Lin, Meiyun, Arlene M. Fiore, Owen R. Cooper, Larry W. Horowitz, Andrew O. Langford, Hiram Levy, Bryan J. Johnson, Vaishali Naik, Samuel J. Oltmans, and Christoph J. Senff. 2012. "Springtime High Surface Ozone Events over the Western United States: Quantifying the Role of Stratospheric Intrusions." *Journal of Geophysical Research, D: Atmospheres* 117 (D21). <https://doi.org/10.1029/2012JD018151>.
- Lin, Meiyun, Arlene M. Fiore, Larry W. Horowitz, Andrew O. Langford, Samuel J. Oltmans, David Tarasick, and Harald E. Rieder. 2015. "Climate Variability Modulates Western US Ozone Air Quality in Spring via Deep Stratospheric Intrusions." *Nature Communications* 6 (May): 7105.
- Malley, Christopher S., Daven K. Henze, Johan C. I. Kuylentierna, Harry W. Vallack, Yanko Davila, Susan C. Anenberg, Michelle C. Turner, and Mike R. Ashmore. 2017. "Updated Global Estimates of Respiratory Mortality in Adults ≥ 30 Years of Age Attributable to Long-Term Ozone Exposure." *Environmental Health Perspectives* 125 (8): 087021.
- Martin, Randall V., Sebastian D. Eastham, Liam Bindle, Elizabeth W. Lundgren, Thomas L. Clune, Christoph A. Keller, William Downs, et al. 2022. "Improved Advection, Resolution, Performance, and Community Access in the New Generation (Version 13) of the High Performance GEOS-Chem Global Atmospheric Chemistry Model (GCHP)." *Geoscientific Model Development*. <https://doi.org/10.5194/gmd-2022-42>.
- Moch, Jonathan M., Eleni Dovrou, Loretta J. Mickley, Frank N. Keutsch, Zirui Liu, Yuesi Wang, Tracy L. Dombek, et al. 2020. "Global Importance of Hydroxymethanesulfonate in Ambient Particulate Matter: Implications for Air Quality." *Journal of Geophysical Research, D: Atmospheres* 125 (18): e2020JD032706.
- Nault, Benjamin A., Duseong S. Jo, Brian C. McDonald, Pedro Campuzano-Jost, Douglas A. Day, Weiwei Hu, Jason C. Schroder, et al. 2021. "Secondary Organic Aerosols from Anthropogenic Volatile Organic Compounds Contribute Substantially to Air Pollution Mortality." *Atmospheric Chemistry and Physics* 21 (14): 11201–24.
- Prashanth, Prakash, Sebastian David Eastham, Raymond L. Speth, and Steven R. H. Barrett. 2022. "Aerosol Formation Pathways from Aviation Emissions." *Environmental Research Communications*, February. <https://doi.org/10.1088/2515-7620/ac5229>.

- Prashanth, Prakash, Raymond L. Speth, Sebastian D. Eastham, Jayant S. Sabnis, and Steven Barrett. 2020. "Post-Combustion Emissions Control in Aero-Gas Turbine Engines." *Energy & Environmental Science*, December. <https://doi.org/10.1039/D0EE02362K>.
- Punger, Elizabeth M., and J. Jason West. 2013. "The Effect of Grid Resolution on Estimates of the Burden of Ozone and Fine Particulate Matter on Premature Mortality in the United States." *Air Quality, Atmosphere and Health* 6 (3). <https://doi.org/10.1007/s11869-013-0197-8>.
- Quadros, Flávio D. A., Mirjam Snellen, and Irene C. Dedoussi. 2020. "Regional Sensitivities of Air Quality and Human Health Impacts to Aviation Emissions." *Environmental Research Letters: ERL [Web Site]* 15 (10): 105013.
- Seltzer, Karl M., Drew T. Shindell, Prasad Kasibhatla, and Christopher S. Malley. 2020. "Magnitude, Trends, and Impacts of Ambient Long-Term Ozone Exposure in the United States from 2000 to 2015." *Atmospheric Chemistry and Physics* 20 (3): 1757–75.
- Shah, V., D. J. Jacob, R. Dang, and L. N. Lamsal. 2023. "Nitrogen Oxides in the Free Troposphere: Implications for Tropospheric Oxidants and the Interpretation of Satellite NO₂ Measurements." *Atmospheric*. <https://acp.copernicus.org/articles/23/1227/2023/>.
- Sun, Youwen, Hao Yin, Xiao Lu, Justus Notholt, Mathias Palm, Cheng Liu, Yuan Tian, and Bo Zheng. 2021. "The Drivers and Health Risks of Unexpected Surface Ozone Enhancements over the Sichuan Basin, China, in 2020." *Atmospheric Chemistry and Physics* 21 (24): 18589–608.
- Tarín-Carrasco, Patricia, Ulas Im, Camilla Geels, Laura Palacios-Peña, and Pedro Jiménez-Guerrero. 2022. "Reducing Future Air-Pollution-Related Premature Mortality over Europe by Mitigating Emissions from the Energy Sector: Assessing an 80 % Renewable Energies Scenario." *Atmospheric Chemistry and Physics* 22 (6): 3945–65.
- Turner, Michelle C., Michael Jerrett, C. Arden Pope III, Daniel Krewski, Susan M. Gapstur, Ryan W. Diver, Bernardo S. Beckerman, et al. 2015. "Long-Term Ozone Exposure and Mortality in a Large Prospective Study." *American Journal of Respiratory and Critical Care Medicine* 64 (6): 897–905.
- Vennam, L. P., W. Vizuete, K. Talgo, M. Omary, F. S. Binkowski, J. Xing, R. Mathur, and S. Arunachalam. 2017. "Modeled Full-Flight Aircraft Emissions Impacts on Air Quality and Their Sensitivity to Grid Resolution." *Journal of Geophysical Research, D: Atmospheres* 122 (24): 13472–94.
- Wang, Peng, Ruhan Zhang, Shida Sun, Meng Gao, Bo Zheng, Dan Zhang, Yanli Zhang, Gregory R. Carmichael, and Hongliang Zhang. 2023. "Aggravated Air Pollution and Health Burden Due to Traffic Congestion in Urban China." *Atmospheric Chemistry and Physics* 23 (5): 2983–96.
- Wang, Xinyue, Yutian Wu, William Randel, and Simone Tilmes. 2020. "Stratospheric Contribution to the Summertime High Surface Ozone Events over the Western United States." *Environmental Research Letters: ERL [Web Site]* 15 (10): 1040a6.
- Whitt, D. B., M. Z. Jacobson, J. T. Wilkerson, a. D. Naiman, and S. K. Lele. 2011. "Vertical Mixing of Commercial Aviation Emissions from Cruise Altitude to the Surface." *Journal of Geophysical Research* 116 (D14): 1–13.
- Williams, Ryan S., Michaela I. Hegglin, Brian J. Kerridge, Patrick Jöckel, Barry G. Latter, and David A. Plummer. 2019. "Characterising the Seasonal and Geographical Variability in Tropospheric Ozone, Stratospheric Influence and Recent Changes." *Atmospheric Chemistry and Physics* 19 (6): 3589–3620.
- Yang, Junhua, Kun Wang, Mang Lin, Xiufeng Yin, and Shichang Kang. 2022. "Not Biomass Burning but Stratospheric Intrusion Dominating Tropospheric Ozone over the Tibetan Plateau." *Proceedings of the National Academy of Sciences of the United States of America*. National Acad Sciences.
- Yim, S. H. L., G. L. Lee, I. W. Lee, F. Allroggen, A. Ashok, F. Caiazzo, S. D. Eastham, R. Malina, and S. R. H. Barrett. 2015. "Global, Regional and Local Health Impacts of Civil Aviation Emissions." *Environmental Research Letters: ERL [Web Site]*.
- Yu, Karen, Christoph A. Keller, Daniel J. Jacob, Andrea M. Molod, Sebastian D. Eastham, and Michael S. Long. 2018. "Errors and Improvements in the Use of Archived Meteorological Data for Chemical Transport Modeling: An Analysis Using GEOS-Chem V11-01 Driven by GEOS-5 Meteorology." *Geoscientific Model Development* 11 (1): 305–19.
- Zhan, Chenchao, Min Xie, Hua Lu, Bojun Liu, Zheng Wu, Tijian Wang, Bingliang Zhuang, Mengmeng Li, and Shu Li. 2023. "Impacts of Urbanization on Air Quality and the Related Health Risks in a City with Complex Terrain." *Atmospheric Chemistry and Physics* 23 (1): 771–88.
- Zhang, Yuqiang, Drew Shindell, Karl Seltzer, Lu Shen, Jean-Francois Lamarque, Qiang Zhang, Bo Zheng, Jia Xing, Zhe Jiang, and Lei Zhang. 2021. "Impacts of Emission Changes in China from 2010 to 2017 on Domestic and Intercontinental Air Quality and Health Effect." *Atmospheric Chemistry and Physics* 21 (20): 16051–65.

Global impacts of aviation on air quality evaluated at high resolution

Sebastian D. Eastham*^{1,2}, Guillaume P Chossière¹, Raymond L. Speth^{1,2}, Daniel J. Jacob³, Steven R.H. Barrett^{1,2}

¹Laboratory for Aviation and the Environment, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³Atmospheric Chemistry Modeling Group, John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA

10 *Correspondence to:* Sebastian D. Eastham (seastham@mit.edu)

Abstract. Aviation emissions cause global changes in air quality which have been estimated to result in ~58,000 premature mortalities per year, but this number varies by an order of magnitude between studies. The causes of this uncertainty include differences in the assessment of ozone exposure impacts and in how air quality changes are simulated, and the possibility that low-resolution (~400 km) global models may overestimate impacts compared to finer-resolution (~50 km) regional models. We use the GEOS-Chem High Performance chemistry-transport model at a 50 km global resolution, an order of magnitude finer than recent assessments of the same scope, to quantify the air quality impacts of aviation with a single internally consistent, global approach. We find that aviation emissions in 2015 resulted in 21,200 (95% confidence interval due to health response uncertainty: 19,400 – 22,900) premature mortalities due to particulate matter exposure and 53,100 (36,000 – 69,900) due to ozone exposure. Compared to a prior estimate of 6,800 ozone-related premature mortalities for 2006 our central estimate is increased by 5.6 times due to the use of updated epidemiological data which includes the effects of ozone exposure during winter, and by 1.3 times due to increased aviation fuel burn. The use of fine (50 km) resolution increases the estimated impacts on both ozone and particulate matter-related mortality by a further 20% compared to coarse-resolution (400 km) global simulation, but an intermediate resolution (100 km) is sufficient to capture 98% of impacts. This is in part due to the role of aviation-attributable ozone, which is long-lived enough to mix through the Northern Hemisphere and exposure to which causes 2.5 times as much health impact as aviation-attributable PM_{2.5}. This work shows that the air quality impacts of civil aviation emissions are dominated by the hemisphere-scale response of tropospheric ozone to aviation NO_x rather than local changes, and that simulations at ~100 km resolution provide similar results to those at two times finer spatial scale. However the overall quantification of health impacts is sensitive to assumptions regarding the response of human health to exposure, and additional research is needed to reduce uncertainty in the physical response of the atmosphere to aviation emissions.

15
20
25
30

1 Introduction

Aviation is a unique source of climate and air quality impacts. For example, the contrails (ice clouds) which form in aircraft exhaust do not occur for any other major mode of transportation, but have been estimated to cause as much climate forcing as all of the carbon dioxide emitted during flight (Lee et al., 2020). Similarly, the nitrogen oxides (NO_x) and other species emitted during flight can have long-lasting chemical consequences which result in global-scale degradation of air quality (Eastham and Barrett, 2016). It has been estimated that emissions during the cruise portion of the flight specifically contribute around 80% of the 8,000 – 58,000 premature mortalities each year attributable to aviation emissions (Barrett et al., 2010a; Eastham and Barrett, 2016; Quadros et al., 2020). When monetized, these air quality impacts are similar in magnitude to the net climate costs of aviation, including CO₂ and contrails (Grobler et al., 2019).

However, there remain several key uncertainties regarding the impacts of aviation emissions on air quality due to the practical challenges associated with simulating a global influence on a local quantity. Global models are well suited to quantify the global change in oxidative capacity due to aviation emissions, but must simulate the entire atmosphere to do so. Models used in previous studies have split the atmosphere up into grid cells which are between 2 and 5° (latitude and longitude) along each side, or roughly 200 – 500 km (Barrett et al., 2010a; Eastham and Barrett, 2016). This means that the models artificially diffuse local, airport-scale emissions over the surrounding area, potentially failing to resolve the co-location of near-airport emissions and exposed populations and underestimating the relative contribution of non-cruise emissions to air quality. A study by Pungler and West (2013) found that, due to this co-location effect alone, coarse resolution (> 250 km) global models would likely be biased low by 30-40% when estimating US population exposure to PM_{2.5}. The use of large grid cells also means that models treat areas of up to 20,000 km² over a city as being a single, well-mixed air mass, and will not be able to resolve non-linear chemical processes which could increase or decrease the air quality response to aviation emissions. If the population mean exposure to air pollution varies strongly with resolution, air quality impact estimates from coarse models could therefore provide misleading results.

Studies with nested, regional models can address this question for limited areas but incur an inconsistency at the model boundary, either in resolution or in the model being used to quantify impacts. These studies calculate global atmospheric composition changes at a relatively coarse resolution, and then use those results to provide the boundary conditions for a finer-resolution simulation. Whereas some studies using nested regional models have found air quality impacts from cruise emissions of a similar magnitude to those from global studies (Yim et al., 2015; Quadros et al., 2020), Vennam et al. (2017) found that the finer-resolution nested model produced changes in surface ozone and fine particulate matter (PM_{2.5}) which were 70 and 13 times smaller respectively than in a global model.

A related question is whether the air quality impacts of aviation are dominated by local sources (e.g. landing and takeoff (LTO) operations and flight through local airspace) or are the result of larger atmospheric changes. If the former is true, then regional regulations and emissions standards applied for only near-surface operations may be sufficient to reduce impacts. However, if impacts are dominated by large scale atmospheric responses to global aviation, then emissions standards will

only be effective if applied both globally and to fuel burn beyond LTO – since LTO accounts for only 9.1% of the global total **from aviation** (Simone et al., 2013).

These questions urgently need to be resolved. New regulations for aviation NO_x emissions are being debated by the International Civil Aviation Organization. Since the establishment of the 1981 CAEE standard limiting NO_x emissions per unit of thrust for landing and take-off operations (LTO), subsequent CAEP regulations have continued to increase stringency, resulting in the current CAEP/8 standard which was set in 2010 (ICAO, 2017). Since the amount of NO_x produced during take-off and during cruise are closely related for most current combustor architectures, these regulations are also the relevant limiting factor for cruise NO_x emissions. Recent studies have suggested that the climate impacts of greater NO_x emissions are sufficiently small relative to the benefits of reduced CO₂ that fuel efficiency should be prioritized over further NO_x emissions reduction (Skowron et al., 2021). However, if air quality impacts are included in this analysis then such a prioritization could cause net environmental damages rather than improvement – dependent on accurate estimation of the air quality consequences (Miller et al., 2022).

This study quantifies the global air quality response to aviation in a single, consistent modeling framework, evaluating both the role of model grid resolution and the relative contribution of local and remote aviation emissions. We use three different global resolutions to quantify how grid resolution affects the simulated outcomes, varying from 400 km to 50 km globally. To isolate the role that in-domain (“local”) and out-of-domain (“remote”) emissions have on air quality in the contiguous United States we use a perturbation approach, performing an additional simulation at each resolution in which aviation emissions over the **United States** are set to zero. **These are supplemented by sensitivity simulations described in Section 2.3. All calculations are performed with the objective of quantifying potential benefits of rapidly eliminating aviation emissions, whether through policy or technological approaches such as post-combustion emissions control (Prashanth et al., 2020).**

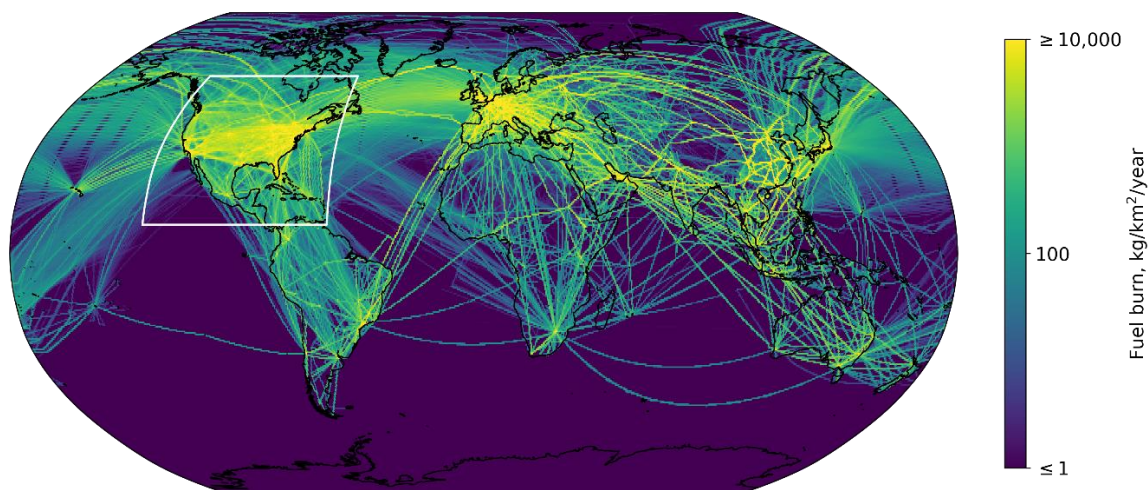
2 Method

We simulate aviation’s impacts on global air quality at three resolutions, ~400 km, ~100 km, and ~50 km, without the use of regional refinement or nesting. For each of the three resolutions, we perform three simulations. We first simulate global atmospheric composition using version 12.6.2 of the GEOS-Chem High Performance model (GCHP) (Eastham et al., 2018) with all aviation emissions enabled (AVGLOBAL) and with all aviation emissions disabled (AVOFF). The simulated differences in concentrations of ozone and particulate matter (PM_{2.5}) in the surface layer are then taken as the total air quality impacts of aviation, with health impacts calculated as discussed below.

These simulations are followed by a further simulation in which aviation emissions are included everywhere except **directly on or above the contiguous United States within a domain** covering 10-60°N and 60-130°W (**AVNOUS**). This domain has previously been used in nested model simulations of regional air quality change, and can therefore be considered to be representative of models which are focused on regional change only (Kim et al., 2015; Hu et al., 2018). The difference in air quality changes between the AVGLOBAL and **AVNOUS** is then taken as the effect of in-region emissions on regional air

quality. The remaining differences in surface air quality within the region, calculated as the difference in air quality between AVNOUS and AVOFF, then correspond to the effect of out-of-region emissions on surface air quality. Figure 1 shows global distribution of fuel burn and the region in which emissions are set to zero when estimating influence of “out-of-region” emissions on “in-region” air quality.

100 To quantify the effect of model resolution, we perform the above set of three simulations at three different global resolutions. The GCHP model uses a cubed-sphere model grid with finite volume advection (Putman and Lin, 2007). Model resolutions are therefore denoted as CN where N is the number of grid elements along each edge of the cube such that higher numbers correspond to finer grid resolution. We perform our simulations at global resolutions of C180, C90, and C24, in which the average side lengths of a grid cell is 51 km, 100 km, and 380 km respectively. As such, they are also approximately
105 equivalent to global resolutions of $0.5^\circ \times 0.625^\circ$, $1^\circ \times 1.25^\circ$, and $4^\circ \times 5^\circ$, but without distortion in grid cell size near the poles and equator. By comparing the air quality impacts calculated at each model resolution, we quantify the effect that increasing model resolution has on simulated air quality. This approach uses a single, global model to accomplish this, avoiding discrepancies due to the use of regional models or different chemical mechanisms (Yim et al., 2015; Vennam et al., 2017).



110 **Figure 1. Simulated global distribution of fuel burn in 2015. Data shown include fuel burned at all altitudes, at an approximate global resolution of $0.5^\circ \times 0.625^\circ$. Emissions over the contiguous United States within the white box are set to zero when estimating the net effect of “out-of-region” aviation emissions on aviation’s air quality impacts in the contiguous US.**

2.1 Atmospheric simulation

We use the global chemistry transport model (CTM) GEOS-Chem version 12.6.2 (<https://doi.org/10.5281/zenodo.3543702>)
115 to simulate all scenarios, as implemented in the GEOS-Chem High-Performance model (Eastham et al, 2018). GEOS-Chem includes unified tropospheric-stratospheric chemistry and has previously been used in a global, coarse-resolution configuration to estimate aviation’s impacts on surface air quality (Eastham and Barrett, 2016; Eastham et al., 2014). Multiple horizontal resolutions are used as described above, but all share a common vertical discretization using 72 non-

uniform layers extending from the surface to a maximum altitude of around 80 km. Each simulation is integrated forwards in time for a total of 13 months from July 1st 2014 through to July 31st 2015, of which the final 12 are used to calculate annual mean changes in surface ozone. This period is used to ensure that the northern hemisphere winter season (December through February, inclusive) is from a single, continuous period, as this is when the greatest air quality impacts from aviation are expected (Eastham and Barrett, 2016). All simulations use meteorological data from the NASA Global Modeling and Assimilation Office Modern Era Retrospective for Research and Analysis version 2 (MERRA-2) reanalysis product.

Emissions from civil aviation in 2015 are calculated using the Federal Aviation Administration (FAA) Aviation Environmental Design Tool (AEDT) (Wilkerson et al., 2010). This includes all emissions from civil airliners in that year during taxi, take-off, climb, cruise, descent, and landing operations. We do not account for military flights or business jets. Emissions are gridded at a resolution of 0.25°×0.25° globally, finer than the highest resolution model grid (C180). A fuel sulfur content of 600 ppm by mass is assumed for all scenarios, unless otherwise stated (Hileman et al., 2010). All nitrogen oxide (NO_x) and volatile organic compound (VOC) emissions are speciated as described in AEDT guidance documents (Barrett et al., 2010b). Table 1 shows the total emitted mass of each relevant compound. Black and organic carbon (BC and OC) emissions are estimated using a fixed emissions factor of 30 mgC per kg of fuel burned for each species. Emissions of a secondary organic aerosol precursor (SOAP) are calculated as 69 g of carbon per kilogram of CO emitted (Kim et al., 2015). GEOS-Chem uses a bulk aerosol parameterization and therefore does not require a size distribution or calculation of the number of particles emitted. The impacts of water vapor emissions and condensation trails on atmospheric composition are not included in this analysis.

Table 1. Total emitted mass for the fleet in 2015. The mass basis for the reported quantity is given separately for each species. All quantities are shown to two significant figures. The rightmost column shows the fraction of global emissions which occur in the simulated North American domain over the US. *US HC emissions data are not separately archived so this number is estimated.

	Global	US domain	% in US domain
NO_x as NO, NO₂, and HONO (TgN)	1.1	0.17	15%
Carbon monoxide (Gg)	590	130	22%
Hydrocarbons (Gg CH₄ mass equivalent*)	62	15	24%
Soot (black carbon, GgC)	7.2	1.2	17%
Organic carbon aerosol (GgC)	7.2	1.2	17%
SOA precursor (GgC)	41	9.2	22%
Sulfur dioxide (GgS)	140	23	17%
Sulfate aerosol (GgS)	2.9	0.48	17%
Total fuel burn (Tg)	240	40	17%

140

Non-aviation emissions are provided by a collection of standard inventories, described in full in the Supplemental Information.

145 GEOS-Chem has been extensively used for air quality assessments in the past, including for evaluations of the effects of aviation (Quadros et al., 2020; Barrett et al., 2010a; Yim et al., 2015), long-range air pollution (Vohra et al., 2021; Huang et al., 2017), and regional-scale changes (Potts et al., 2021; Vohra et al., 2021) on air quality. A multi-model intercomparison performed by Cameron et al. (2017) found that the effects calculated by GEOS-Chem of aviation on surface air quality were consistent with estimates from other widely-used global atmospheric models including the Community Atmosphere Model 5 and the GEOS-5 Earth System Model. GEOS-Chem has also been continuously evaluated against global observations of both atmospheric composition and air quality (Zhang et al., 2011; Christian et al., 2018; Dasadhikari et al., 2019; Quadros et al., 2020). Nonetheless the results of this study reflect only the estimates from a single model.

2.2 Health impact estimation

Air quality impacts are calculated based on the difference in concentrations of ozone and PM_{2.5} in the lowermost simulated atmospheric layer. The lowermost layer is approximately 120 meters thick for a surface pressure of 1013.25 hPa, and the average concentration within this surface layer is treated as the exposure-relevant value. Model predictions are not-bias corrected to observations, in part because the relatively small effect of aviation emissions on ozone and particulate matter has a different spatial pattern than background concentrations (Cameron et al., 2017). Bias correction may therefore impose a non-physical pattern on the changes in surface concentration due to aviation.

160 Impacts are calculated on the 30 arc-second (1/120th of a degree) global grid on which population density data is provided by the Gridded Population of the World version 4.11 (Center for International Earth Science Information Network - CIESIN - Columbia University, 2016). The concentrations of ozone and PM_{2.5} are taken from whichever simulation grid cell contains that 30 arc-second cell, with no interpolation. If the population cell straddles the boundary between two simulation grid cells, the area-weighted mean concentration is used. The age distribution and baseline mortality rates within each grid cell are supplied by the World Health Organization (WHO) through their 2016 Global Health Estimates (World Health Organization, 2018). For non-member countries of the WHO, we use the world region mean rate instead from the same source.

165 In each grid cell and for each age bracket, we calculate the relative risk of mortality due to chronic exposure to ozone and PM_{2.5} with and without aviation (RR_{BASE} and RR_{NOAV} respectively). The change in the annual mortality (ΔM) due to some disease for that age bracket is then calculated for each grid cell as

$$\Delta M = M_{BASE} \times \frac{RR_{NOAV} - RR_{BASE}}{RR_{BASE}} \quad (1)$$

170 where M_{BASE} is the number of mortalities due to that disease in 2016. The relative risk is calculated by comparing the simulated exposure-relevant concentration without aviation (χ_{NOAV}) to the concentration simulated when aviation is included

(χ_{BASE}) using an appropriate concentration response function. In this case ΔM is expected to be negative, implying that reducing aviation emissions to zero would reduce mortality rates. The mortality due to aviation reported in this paper is therefore $-1 \times \Delta M$. For both ozone and $\text{PM}_{2.5}$ exposure we first calculate the relevant daily quantity and then average over the year. Our evaluations focus on the link between long-term (chronic) exposure and increased mortality, rather than acute impacts associated with short-term increases in exposure.

For ozone, the increase in relative risk of mortality is calculated based on the association between exposure and mortality identified by Turner et al. (2015). Said study analyzed a cohort of 669,046 participants in the American Cancer Society Cancer Prevention Study II from 1982 to 2004, finding a 12% increase (95% confidence interval: 8.0-16%) in respiratory mortality per 10 ppb increase in annual mean, maximum daily 8-hour average (MDA8) ozone concentration. We use this data in a log-linear concentration response function (CRF), such that the relative risk due to ozone exposure is calculated as

$$\text{RR}_{\text{NOAV}} = \exp(\beta_{\text{LL}}[\chi_{\text{NOAV}} - \chi_{\text{BASE}}]) \quad (2)$$

where the central value of β_{LL} is calculated as $\ln(1.12)/10 = 0.011 \text{ ppb}^{-1}$. Only adults over the age of 30 are included when calculating the increase in mortality, as this was the cohort in which the relationship was observed. Uncertainty in the concentration response function is quantified by treating β_{LL} as a triangularly-distributed random variable. We use $\ln(1.12)/10$ as the mode and fit a triangular distribution such that $\ln(1.06)/10$ lies at 2.5% and $\ln(1.18)/10$ as 97.5% along the cumulative distribution function.

For particulate matter, relative risks are calculated using the Global Exposure Mortality Model (GEMM) (Burnett et al., 2018). The GEMM is a set of non-linear concentration response functions which estimate the increase in relative risk of mortality based on associations with the annual-average, 24-hour mean $\text{PM}_{2.5}$ mass concentration at standard temperature and pressure. We apply the age-specific GEMM CRF for combined non-communicable disease and lower respiratory infection (NCD+LRI). The parameters for the GEMM are constructed based on a meta-analysis of 41 cohort studies worldwide examining the relationship between exposure to fine particulate matter and non-accidental mortality, and are described in detail in Burnett et al. (2018).

Impacts are calculated by performing 1,000 random draws of the CRF parameters in a paired Monte-Carlo simulation. From this, the mean and 95% confidence interval of mortality due to individual sources and mortality due to ozone and $\text{PM}_{2.5}$ combined can be estimated.

2.3 Sensitivity simulations

The core simulations described above allow us to quantify the total health impacts which could be avoided by eliminating aviation emissions, the sensitivity of that estimates to model resolution, and the degree to which in-region emissions control in-region outcomes. In order to better understand the causes of these impacts and their limitations, we also perform a series of additional sensitivity simulations at C90 resolution. We first repeat the core simulations (AVGLOBAL, AVNOUS, and AVOFF) using a different version (14.2.2) of the GCHP model (see Supplemental Information). This allows us to assess the

205 degree to which changes in chemistry, background emissions, and process representation can affect the overall health impact associated with aviation. We also extract data in each of these simulations about chemical production of ozone and global chemical loss rates to support analysis of the chemistry and fate of aviation emissions (Section 3.3).

We then simulate two shorter sensitivity scenarios with the updated model in which aviation NO_x is eliminated globally (AVNONOX) and aviation emissions are simulated above 1 km altitude only (AVNOLTO). These simulations allow us to establish to what extent aviation NO_x (rather than other emissions) or LTO emissions (rather than those from cruise) are responsible for changes in population exposure to both ozone and $\text{PM}_{2.5}$. These simulations cover the period August 2014 to February 2015 (inclusive) in order to capture the wintertime response to aviation.

210 Finally, two simulations are performed at resolutions of C24 and C90 which evaluate how the rate of vertical mixing is affected by model resolution by estimating the concentration of radioactive tracer ^7Be throughout the atmosphere. This test has been previously used to assess atmospheric model accuracy in simulating vertical transport (Yu et al., 2018; Liu et al., 2001) and provides a benchmark for the degree to which high-altitude atmospheric composition can affect surface conditions.

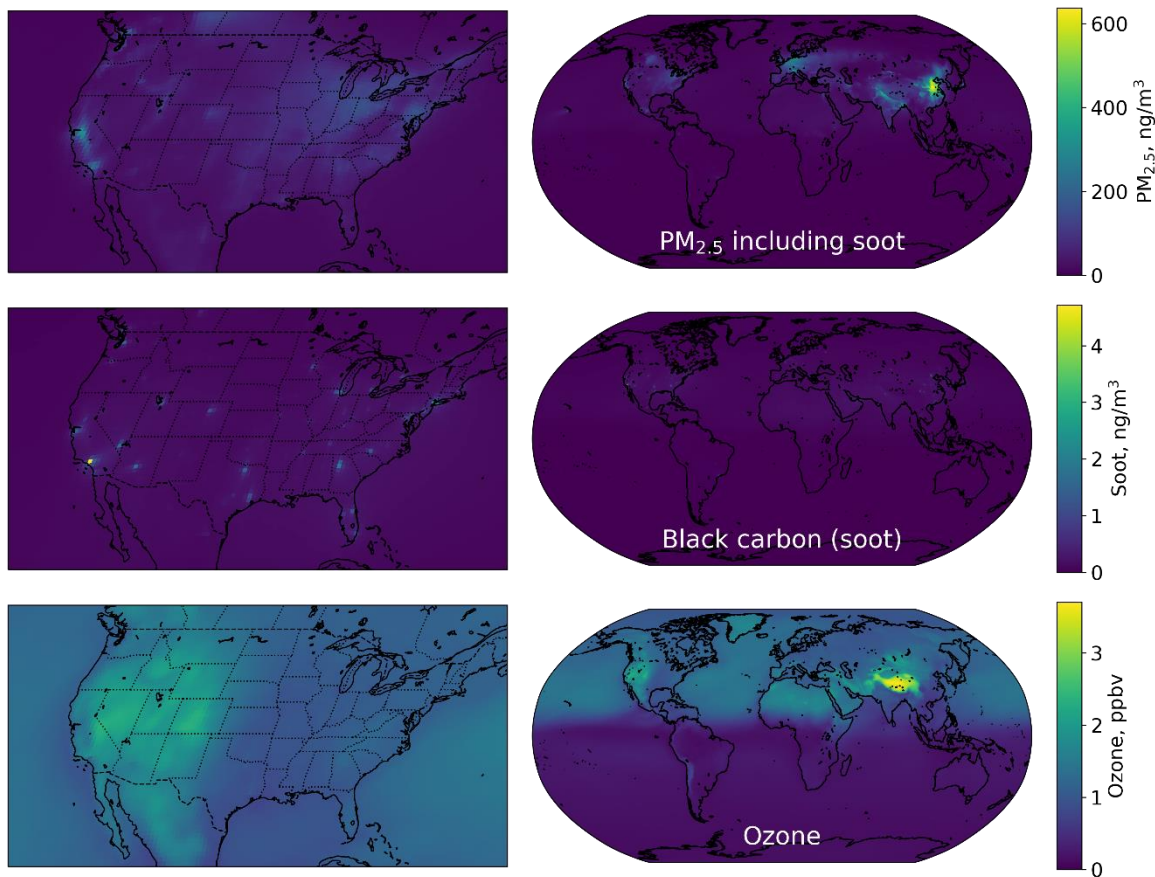
3 Results

Based on results generated at ~50 km resolution globally, we find that the increase in $\text{PM}_{2.5}$ exposure results in an additional 21,200 (95% confidence interval **due to uncertainties in health response**: 19,400 to 22,900) mortalities globally each year. 220 Aviation-attributable ozone exposure results in an additional 53,100 (36,000 to 69,900) mortalities. Combined, we estimate that aviation results in 74,300 mortalities (57,300 to 91,100), or 311 mortalities per teragram of fuel burned. **These results are analyzed in Sections 3.1 to 3.4 and compared to prior studies in Section 3.5.**

Figure 2 shows how the changes in surface concentrations are distributed, both in the US (left) and globally (right). Changes in ozone are diffuse throughout the Northern Hemisphere, reaching a maximum over the Himalayan plateau and the Western United States. Changes in total particulate matter are more heterogeneous, with the greatest increases occurring over western 225 China, northern India, and western Europe. This effect is likely driven by the effect of increased oxidant concentrations acting on non-aviation precursor gas emissions (Eastham and Barrett, 2016). Increases in soot (black carbon) are shown as soot is rapidly removed by precipitation, and therefore highlights near-airport regions as these are the locations most strongly affected by direct particulate matter emissions from aircraft. However, peak concentrations of aviation-attributable soot are 230 two orders of magnitude smaller than the more diffuse increases in total $\text{PM}_{2.5}$.

Our findings suggest that direct exposure to cruise-attributable ozone is the largest air quality concern for aviation, causing 2.5 times as many premature mortalities as $\text{PM}_{2.5}$. This means that the dominant air quality impact of aviation is not localized to airports, but is also not globally or regionally homogenous. Although aviation-attributable ozone is present throughout the Northern Hemisphere with a mean mixing ratio of 1.1 ppb, elevated regions can reach annual average mixing ratios in excess 235 of 3 ppb. In these regions air more rapidly reaches the surface from the cruise level, meaning that less of the cruise-altitude

ozone is destroyed before the air reaches the population. This is consistent with studies which have shown that surface ozone in the Western US can be strongly influenced by stratospheric intrusions (Lin et al., 2012) and is analyzed in Section 3.3. Although the greatest increases are on the Himalayan plateau, aviation-attributable ozone in the western US also exceeds 2 ppb. This is due to the descent of upper tropospheric air in the lee of mountain ranges, such that the largest impacts of aviation emissions on US air quality are just east of the Rocky Mountains.



245 **Figure 2. Changes in surface air quality due to global aviation. From top to bottom: changes in PM_{2.5} (including soot), changes in soot only, and changes in ozone. The left column is a zoomed-in version of the right plot, focusing on the contiguous United States. Soot is explicitly singled out due to its potential as an indicator of the direct influence of aviation exhaust without any chemical intermediaries.**

250 The distribution of aviation-attributable PM_{2.5} follows a different pattern. Aviation-attributable soot is concentrated near airports, but this peaks at less than 5 ng/m³ (see Section 3.1). The greatest changes in PM_{2.5} exceed 600 ng/m³ and are spread over larger areas in locations which already have elevated background pollution concentrations, such as western China, central Europe, and the US Northeast.

Figure 3 further explores the factors contributing to aviation-attributable increases in surface PM_{2.5}. The change in surface PM_{2.5} due to aviation is correlated with the concentration of background (non-aviation) PM_{2.5}, calculated as the concentration with all other sources included, with an R^2 of 0.495. However, aviation-attributable PM_{2.5} concentration is less strongly correlated (R^2 of 0.140) with the change in aviation-attributable black carbon, an indicator of nearby aviation activity. This is consistent with the hypothesis that increases in PM_{2.5} due to aviation for a location in the Northern Hemisphere occur because aviation-attributable ozone contributes to the formation of secondary particulate matter from existing precursors.

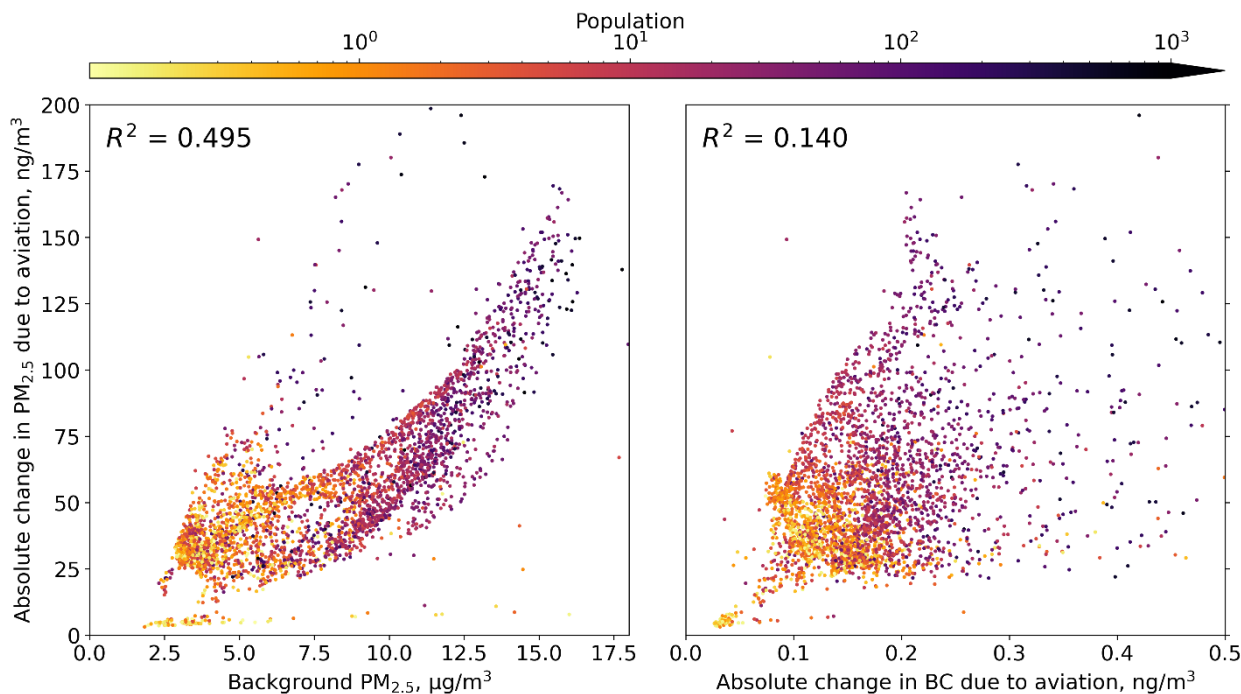
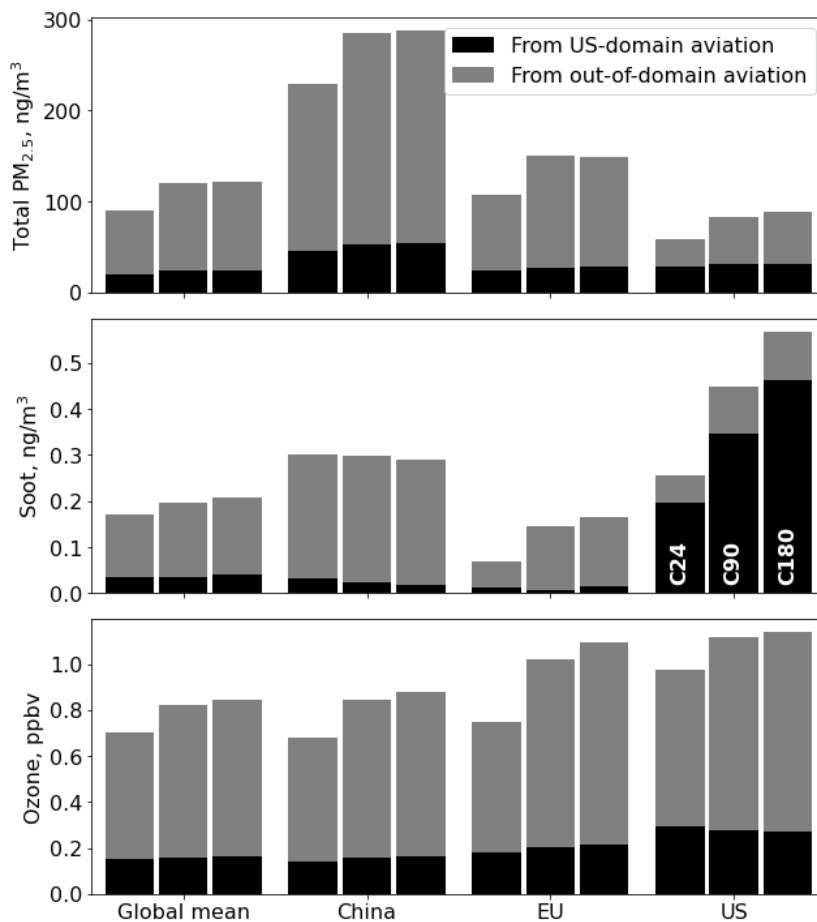


Figure 3. Scatter plots of the change in PM_{2.5} due to aviation as a function of other variables. Left: the absolute change in surface PM_{2.5} due to aviation plotted as a function of the background (non-aviation) concentration of PM_{2.5}, colored by the total population in the grid cell. Right: aviation-attributable PM_{2.5} plotted as a function of the absolute change in black carbon due to aviation. Each point corresponds to a single grid cell at C180 resolution. Only grid cells in the Northern Hemisphere with a population of at least 0.1 people per square kilometer are shown. R^2 values are based on a linear least-squares fit.

3.1 The effect of resolution on the simulated air quality impacts of aviation emissions

Figure 4 shows the population-mean exposure to PM_{2.5} (including soot), soot, and ozone for the global mean, China, the 27 European Union member states (“EU”), and the US when simulations are carried out at three different resolutions. All differences are given relative to the aviation-attributable change calculated at C180 unless otherwise stated. We find that simulations performed at the coarsest, C24 (~400 km) resolution result in a global mean ozone exposure which is 17% lower than when calculated at the C180 (~50 km) resolution, compared to 3.0% lower at C90. For PM_{2.5} the exposure is 26 and 1.2% lower respectively, whereas for soot it is 18 and 6% lower. In total, the net mortalities calculated when performing simulations at a resolution of ~50 km are 2.5% higher than at ~100 km, and 25% higher than at ~400 km resolution.

275 The greater resolution has two effects: physical phenomena are more finely resolved, and changes in surface air quality which occur local to population centers can be more accurately collocated. These effects can be separated by performing exposure calculations using the same high-resolution (~50 km) output data but downgrading it to low resolution (~400 km) before calculating exposure. Doing so increases the calculated ozone exposure by 3.7% but decreases calculated PM_{2.5} exposure by 14%. This suggests that the higher-resolution simulation of atmospheric phenomena is more important than population collocation for ozone but that the collocation effect is significant for PM_{2.5}, consistent with prior work focused on non-aviation exposure (Punger and West, 2013).



280 **Figure 4. Changes in air quality resulting from aviation emissions, calculated for three different resolutions (C24, C90, and C180 from left to right). Black bars show the contribution of emissions over the US only, while grey bars show the contribution of non-US domain emissions.**

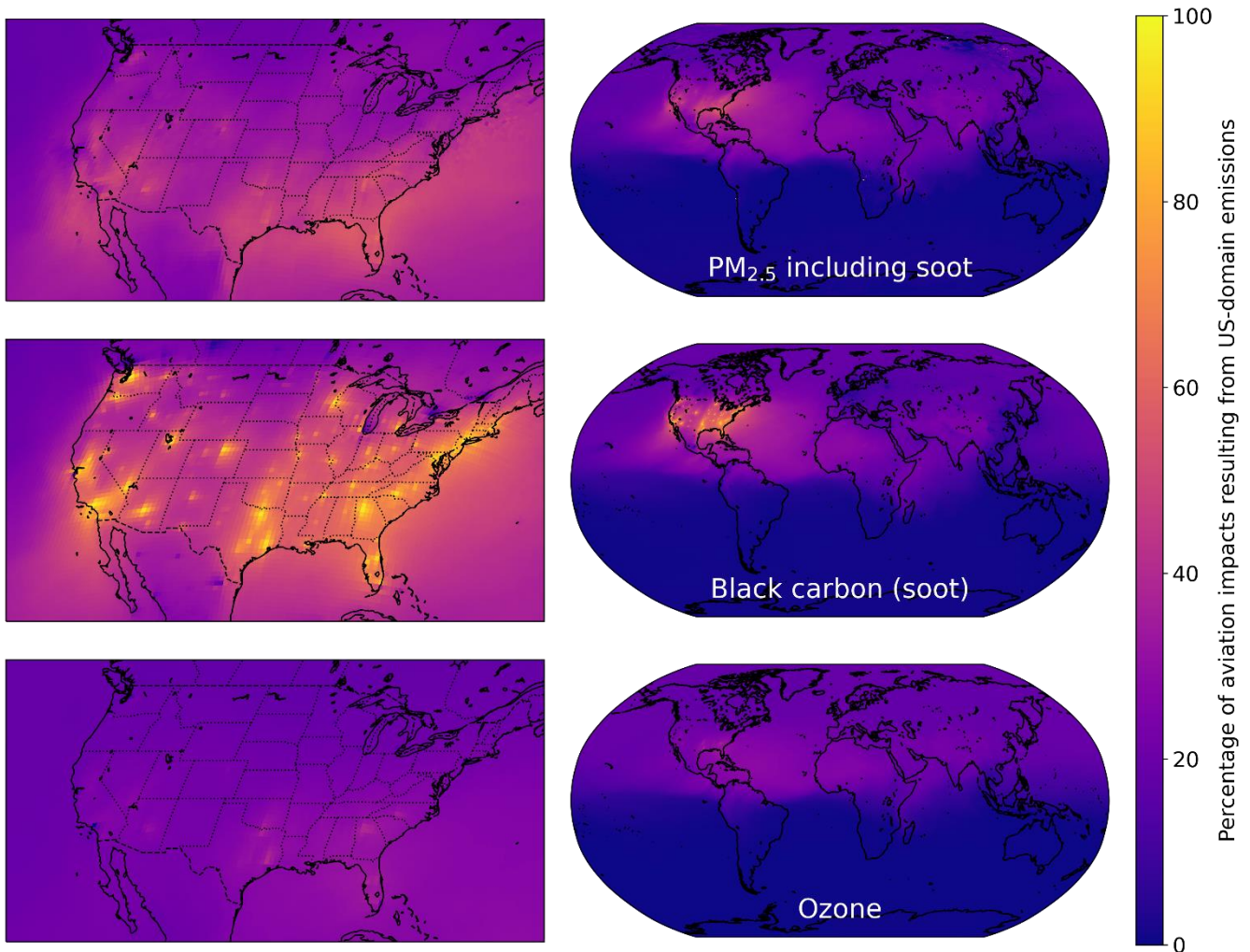
285 This effect is further illustrated by the different simulated effect of aviation on regional and near-airport air quality at different resolutions. At both C90 and C180 resolution, the pattern of change in surface-level PM_{2.5} and ozone is qualitatively similar, and the greatest changes are not necessarily in the immediate vicinity of airports. Figure S2 shows that,

over the US West Coast, $PM_{2.5}$ is increased most in the California central valley and ozone is increased most over the Plains States, with the same effect visible at both C180 and C90. However, at C24 these patterns are not well resolved. Exposure to black carbon (soot) is greater in the vicinity of major airports and more accurately captured at C180, but even in a hotspot around Los Angeles International and Santa Monica airports where soot concentrations are increased by 4.7 ng/m^3 , only 2.7% of the aviation-attributable increase in $PM_{2.5}$ is due to soot. The rest is due to a more diffuse increase in $PM_{2.5}$ extending throughout the California central valley which is captured by simulations at both C90 and C180, but which at C24 is diffused over most of California. We explore the potential causes of differences in estimated ozone and $PM_{2.5}$ changes in more detail in Section 3.3.

3.2 In-region versus out-of-region emissions

Concerning the role of long-range transport compared to in-region emissions, of the 21,200 mortalities due to aviation-attributable exposure to $PM_{2.5}$, 1,610 (1,470 to 1,740) occur in the United States, compared to 3,940 (2,670 to 5,180) of the 53,100 mortalities due to aviation-attributable ozone exposure. This yields a combined estimate of 5,550 (4,290 to 6,790) premature mortalities in the US due to global, year-2015 aviation emissions. Fuel burn in the US domain therefore results in an additional 39 US mortalities per Tg of fuel burn, whereas fuel burn outside of the domain results in 20 US mortalities per Tg of fuel burn. However, aviation emissions outside the US domain cause 72% of aviation-attributable health impacts within the US domain since 83% of global fuel burn occurs outside the US.

The relative contribution of US aviation emissions to aviation-attributable surface air quality degradation in each region is shown in Figure 4 as the black segment of each bar. This is again calculated as the expected change in surface air quality if aviation emissions could be eliminated over the US. Aviation emissions from within the US domain contribute 37% and 24% of aviation-attributable exposure to $PM_{2.5}$ and ozone respectively in the US. At C24 (~400 km) resolution these contributions are higher - 49% and 30% respectively - despite total exposure being 26% and 17% lower relative to that calculated at C180 (~50 km).



310 **Figure 5. Relative contribution of aviation in the US domain to all air quality impacts of aviation. Light colors indicate that emissions over the US domain are the dominant contributor, while dark colors indicate that non-US emissions are dominant. Data is shown for calculations performed at a global resolution of C180 (~50 km).**

The same US emissions contribute to a lower proportion (19%) of aviation-attributable exposure to PM_{2.5}, but a greater absolute value of 54 ng/m³ in China compared to 32 ng/m³ in the US. This is consistent with the hypothesis that aviation-attributable emissions are promoting the formation of PM_{2.5} from non-aviation sources, rather than direct emissions of aviation PM_{2.5} being responsible.

315

The exception to this is soot. 82% of US exposure to aviation-attributable soot is the result of aviation emissions from within the US domain, whereas 6.2% of aviation-attributable exposure in China is attributable to US-domain aviation emissions. However, aviation-attributable soot exposure in the US at C180 makes up 1.4% of total aviation-attributable PM_{2.5} exposure due to the rapid wet scavenging of soot from the atmosphere (Wang et al., 2014). The reason for the greater concentration and resolution-sensitivity of aviation-attributable soot in the US is unclear.

320

This result is shown globally in Figure 5. As in prior figures, we show the annual average value of PM_{2.5} and ozone. The relative contribution of US-domain aviation emissions to US air quality is greatest for black carbon, where it reaches nearly 100% around airports. For ozone, impacts across most of the Northern hemisphere are close to the fractional contribution of US emissions to global fuel burn, 17%. For PM_{2.5} the relative contribution is more heterogeneous than for ozone but less than for black carbon. This is due to the contribution of secondary particulate matter, such as ammonium nitrate and acidic sulfate aerosols, which form in response to aviation-attributable ozone.

3.3 Chemistry and fate of aviation emissions by season and the role of model resolution

In order to understand the results discussed above, we also investigate the mechanism by which surface-level ozone and PM_{2.5} are changed in response to aviation. Whitt et al. (2011) showed that PM_{2.5} emitted or formed at cruise altitudes is unlikely to survive to the surface, and previous work has shown that the change in surface-level PM_{2.5} is likely to be the result of aviation NO_x-attributable ozone reaching the surface and promoting formation of PM_{2.5} locally from existing (non-aviation) precursor species (Eastham and Barrett, 2016). Aviation's effects on surface air quality have also consistently been found to be maximized during winter (Eastham and Barrett, 2016; Lee et al., 2013; Cameron et al., 2017; Phoenix et al., 2019). We therefore focus our analysis on seasonal changes in aviation-attributable ozone, using results from the sensitivity simulations performed with GCHP 14.2.2 (see Section 2.3).

Figure 6 shows how the vertical profile of ozone averaged from 15° to 45°N, in addition to odd oxygen production rates, is affected by aviation emissions during each season. Aviation-attributable surface-level ozone is maximized during winter and minimized during summer. This is true at all altitudes below 10 km, whereas background ozone (rightmost panel) is greatest during summer over the same range of altitudes.

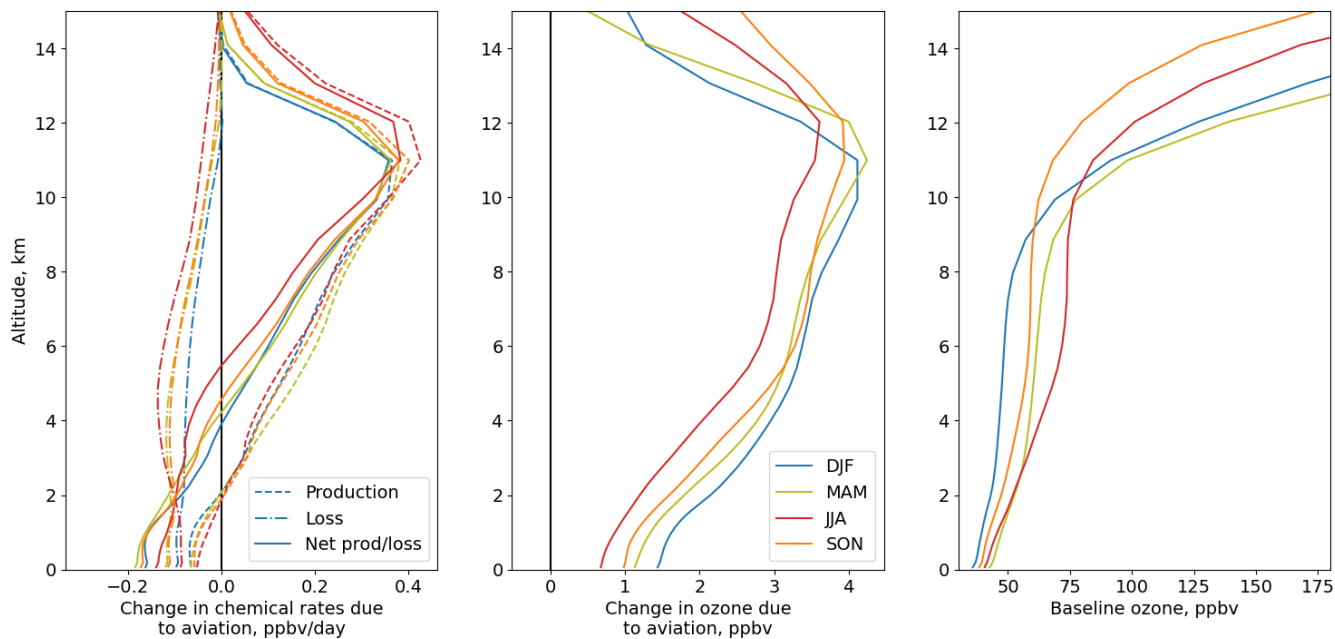


Figure 6. Changes in production and loss rates of odd oxygen (left) and the net effect on ozone mixing ratios (centered) along with baseline ozone mixing ratios (right), separated by season. Values are for the average from 15° to 45°N at C90 resolution.

345 We attribute the difference to two causes. First, the change in odd oxygen production rates in the free troposphere is consistent throughout the year, while odd oxygen loss rates are reduced during the winter. This results in longer lifetimes for aviation-attributable ozone without a decrease in production.

The second factor is transport. Figure 7 overlays the aviation-attributable change in ozone with the mean pressure (vertical) velocity in Northern Hemispheric winter (DJF) and summer (JJA), excluding sub-grid convection. As discussed by e.g. Williams et al. (2019), the contribution of stratospheric air to surface ozone is controlled by both the degree to which upper tropospheric air is enriched by stratospheric ozone and the rate of vertical mixing from the upper troposphere to the surface. 350 The former factor is not relevant to aviation's influence, which as discussed above shows relatively little variability between seasons in this latitude range although aviation-related increases in ozone are greater in JJA north of 45°N. The latter factor of increased transport rates, however, does appear to be significant. Downwelling motions in Northern midlatitudes during winter months allow ozone to descend from cruise altitudes to reach the surface. When combined with reduced ozone loss 355 rates, the result is an increase in the amount of aviation-attributable ozone reaching the surface during winter.

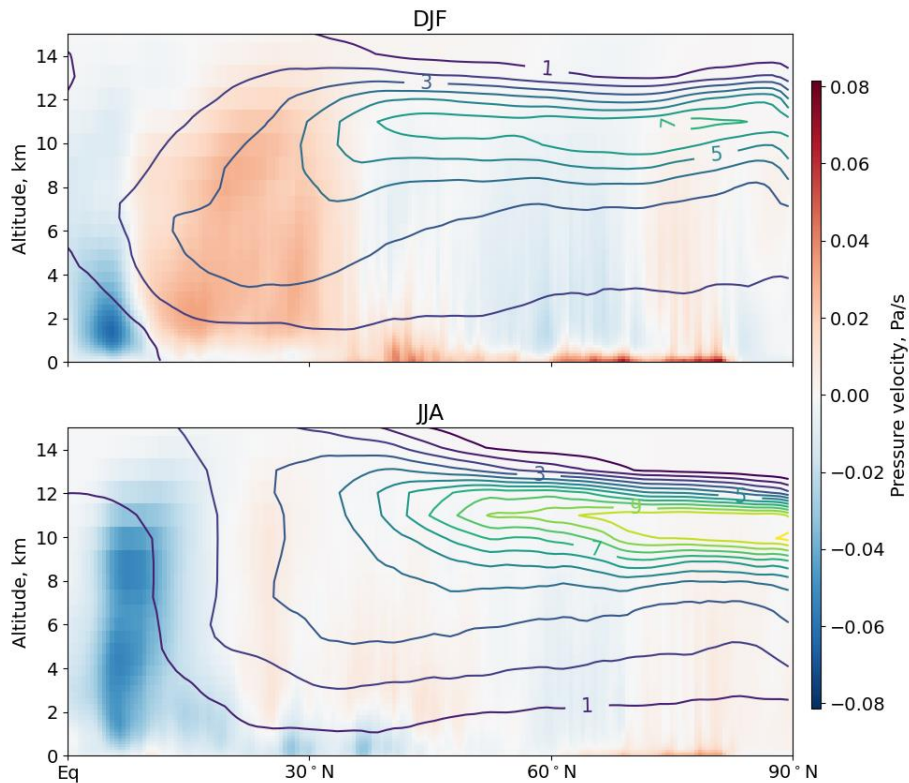


Figure 7. Seasonal mean pressure velocities (shading, Pa/s) and aviation-attributable changes in ozone (contours, ppbv). Blue colors indicate an upwelling motion, while red colors indicate a downwelling motion. Each contour corresponds to a change of 1 ppbv ozone. Pressure velocities are extracted from the MERRA-2 meteorological data used for all simulations.

360 An additional factor in the relationship between cruise altitude emissions and surface air quality is the relative timescale of vertical versus zonal transport. Based on data from the MERRA-2 reanalysis the annual mean zonal wind speed between 15 and 45°N, from 6 to 10 km altitude, is between 9.4 and 18 m/s. At these speeds an aircraft emission or its products would traverse the North American domain shown in Figure 1 at the equator within 4.9 to 9.6 days, on average. Given that aircraft cruise at a typical pressure of 300 – 200 hPa (i.e. 700 – 800 hPa from the surface) and that average pressure velocities are
 365 less than 0.1 Pa/s (or ~90 hPa/day) as shown in Figure 7, this implies that cruise-altitude emissions and their products over the US would typically leave the North American domain before they could reach the surface.

This explains why the global pattern of impacts from emissions over North America is not substantially different from the pattern of impacts from emissions over the rest of the globe (Section 3.2). During summertime, slower zonal mixing and shorter lifetimes result in US aviation emissions causing 3.9 times more ozone (population-weighted mean exposure) over the US than over the rest of the world. This ratio falls to 1.3 during winter. Since wintertime changes in ozone dominate
 370 aviation-attributable changes in surface air quality, this lower wintertime ratio dominates the overall signal such that out-of-region emissions contribute 79% of the aviation-attributable annual mean increase in US exposure to ozone.

375 The changes we calculate in surface air quality are predominantly due to aviation NO_x emitted at cruise altitude. In a sensitivity simulation where emissions of aviation NO_x are excluded, the wintertime (DJF) aviation-attributable increase in population-weighted mean exposure to ozone and PM_{2.5} is decreased by 108 and 103% respectively. In a second sensitivity simulation in which aviation emissions below 1 km altitude are eliminated (i.e. those associated with landing and take-off), the decrease is 0.029 and 5.9% respectively. This is consistent with the finding by Barrett et al. (2010a), Lee et al. (2013), and Prashanth et al. (2022) that emissions of cruise-altitude NO_x are the dominant contributor to aviation-attributable changes in atmospheric ozone and aerosol concentrations.

380 This leaves the question of how differences in the representation of the same physical phenomena at different resolutions might cause discrepancies in surface-level air quality changes attributable to aviation emissions. Although the emissions data and underlying meteorological fields are identical, simulated physical processes including transport, deposition, and chemistry are affected by the change in resolution.

385 With regards to vertical transport and scavenging Yu et al. (2018) found that, when using the same model with meteorological data from the same source as in this study, degrading the grid resolution from 0.25°×0.3125° (~C360) to 2°×2.5° (~C48) resulted in a reduction in the quantity of soluble material (⁷Be) descending from the tropopause to the surface. This effect was strongest in the subsiding subtropics, with a maximum effect of “up to 40%” as measured by changes in the surface-level concentration of ⁷Be. In a sensitivity simulation performed at both C90 and C24 resolution, we find that wintertime surface-level concentrations of ⁷Be between 15 and 45°N are 7.1% greater when simulated at C90 than 390 when simulated at C24. At 8 km altitude this is reversed, with concentrations 10% greater in the C24 simulation than in the C90 simulation. This suggests that vertical mixing across the tropopause is overestimated at C24 resolution but that vertical transport to the surface may be underestimated, and is consistent with the finding that aviation-attributable surface-level air quality impacts are underestimated when simulated at C24 resolution.

395 With regards to chemistry, we find that the changes in the model resolution cause differences in cruise-altitude concentrations of key precursor species. At C24 resolution, baseline mixing ratios of ozone during wintertime at 10 km altitude are 5.8% greater (from 15 to 45°N) than at C180, while concentrations of NO_x and NO_y are both 7.8% lower. The opposing signs are consistent with previous studies which found an increase in short-lived ozone from aviation emissions when simulated non-aviation NO_x emissions are reduced (Holmes et al., 2011). Since we also find that surface-level ozone is lower at C24 than at C180, this implies that differences in the representation of chemistry and background composition at 400 low resolution partially compensate for the reduction in vertical transport diagnosed above. At C90, these differences are reduced to less than 1%.

3.4 Additional sources of uncertainty in the atmospheric response to aviation emissions

405 Due to the diffuse nature of aviation emissions impacts, direct empirical constraints on aviation-induced changes in air quality are more challenging than might be the case for surface-based emissions sources. As the focus of this work is on understanding the role of model resolution and the relative contribution of in-region versus out-of-region emissions we do

not perform an exhaustive quantification of other potential sources of uncertainty. Nonetheless there are known sources of uncertainty which are likely to affect the simulated impacts of aviation on air quality, and we perform a brief analysis of our results in order to identify potential additional sources of uncertainty.

410 An evaluation of the climate impacts of aviation NO_x by Holmes et al. (2011) highlighted four areas of particular concern for understanding aviation-attributable ozone changes: chemical kinetics, non-aviation anthropogenic emissions, emissions of lightning NO_x, and modifications to the underlying model. In addition, work by Barrett et al. (2010a) and Quadros et al. (2020) highlighted the potential role that uncertainty in aerosol emissions and aerosol modeling might have with regards to estimates of aviation-attributable changes in surface-level PM_{2.5}.

415 With these factors in mind, we compare the results calculated at a single resolution (C90) between two versions of the same model, GCHP 12.6.2 and 14.2.2 (see Section 2.3). We find that, due to changes in physical modeling of atmospheric chemistry and physics, the increase in population mortality attributable to aviation decreases by 15% between the two model versions. This is similar to a 12% change in short-lived ozone radiative forcing estimated for two different versions of the same model in Holmes et al. (2011) but smaller than the range of uncertainty reported in a multi-model intercomparison by Cameron et al. (2017). These uncertainties are in addition to uncertainties in the health response already quantified.

420 With regards to aerosol modeling, we find that 76% of the total increase in surface-level fine particulate matter (as calculated in GCHP 14.2.2) is nitrate by mass, compared to 22% ammonium and less than 2% sulfate or carbonaceous aerosol. The baseline composition of aerosol (excluding sea salt and dust) is 25% nitrate, 19% ammonium, and 31% sulfate. Although GEOS-Chem includes a sophisticated treatment of inorganic aerosol, uncertainty in nitrate aerosol production and concentrations remains high with several recent updates to GEOS-Chem directly affecting inorganic aerosol (Shah et al., 425 2023; Moch et al., 2020). This means that, alongside other physical uncertainties, the above assessment is sensitive to ongoing work to improve the representation of inorganic aerosol, and in particulate nitrate, in global models.

3.5 Comparison to previous work

These impacts of aviation emissions on global air quality and mortality are consistent with a recent analysis which found that year-2005 aviation emissions resulted in 58,000 mortalities globally, of which 38,000 were due to ozone exposure (Quadros et al., 2020). Our 2015 emissions inventory includes 33% more fuel burn by mass than was used for their study, and our 430 estimate of the net mortality impact is also 33% greater, **although we find that 71% of mortalities are due to ozone whereas Quadros et al. (2020) find 66%**. The larger contribution of ozone in this work and Quadros et al. (2020) relative to earlier assessments is due to the use of more recent epidemiological data which includes wintertime ozone exposure (Turner et al., 2015). This is significant because, **as shown in Section 3.3**, aviation-attributable ozone is maximized during winter and 435 minimized during summer (Eastham and Barrett, 2016). If we instead use epidemiological data from an earlier study by Jerrett et al. (2009) which considers only summertime ozone, ozone-related health impacts of aviation emissions are estimated to be 68% lower and net impacts 49% lower. **Given that there are relatively few studies of the health impacts of chronic exposure to ozone compared to exposure to PM_{2.5}, this uncertainty is a key area of future research.** A more detailed

assessment of the sensitivity of our conclusions to the health impact assessment method is provided in the Supplemental
440 Information.

These results are also consistent with studies of aviation's air quality impacts as calculated using models other than GEOS-Chem. A multi-model intercomparison by Cameron et al. (2017) found that global surface ozone in 2006 was increased by 0.21 to 0.65 ppbv per teragram of emitted NO_x on a nitrogen mass basis (TgN). Although GEOS-Chem was included in that study, it was neither the lower nor upper bound of the ozone response. We find a value of 0.59 ppbv per TgN, within the
445 range reported by Cameron et al. (2017). Similarly, Vennam et al. (2017) used a hemispheric-scale version of the CMAQ model to investigate aviation's air quality impacts, finding an increase of 0.53 ppbv ozone per TgN. Lee et al. (2013) and Phoenix et al. (2019) used the Community Atmosphere Model (CAM) to investigate the same question and reported an increase of "several" and 1.78 ppbv per TgN respectively in wintertime Northern Hemispheric ozone. Similarly, the Cameron et al. (2017) evaluation estimated increases in global surface PM_{2.5} of between -210 and +96 ng/m³/TgN, compared
450 to our estimate of +12 ng/m³/TgN.

Nevertheless, our results constitute a significant increase in the estimate air quality impacts of aviation relative to studies such as Eastham and Barrett (2016) and Barrett et al. (2010a). The former reported 6,800 premature mortalities per year due to aviation-attributable ozone exposure and 9,200 due to aviation-attributable PM_{2.5} exposure, compared to 53,100 (95% CI: 36,000 – 69,900) and 21,200 (19,400 – 22,900) respectively in this work. Assuming linear or near-linear relationships the
455 factor of 7.8 increase in ozone-related impacts is mostly attributable to the updated epidemiological data as described above, which results in a 3.2 times increase in ozone-related mortality. This is accompanied by a factor 1.6 increase due to the Turner et al. data being applicable to all respiratory diseases and not just chronic obstructive pulmonary disease and asthma, yielding a net factor of 5.0 relative to the approach used in Eastham and Barrett (2016). An additional factor of 1.3 is due to increased fuel burn, since our work examines 2015 whereas these previous studies examine aviation in 2006. We also
460 estimate that the use of finer resolution of our analysis (~50 km) compared to the ~500 km resolution of the previous studies results in an increase in the estimated ozone-related impacts of aviation emissions by a factor of 1.2. These three factors combined imply a 7.6 times increase in ozone-related impacts, consistent with the factor of 7.8 which is observed.

For PM_{2.5} we estimate 2.3 times as many mortalities as were estimated by Eastham and Barrett (2016). This smaller factor is only partially explained by the factors described above. While the same factor of 1.3 applies for fuel burn, the concentration
465 response function used here results in 14% fewer mortalities being attributed to aviation emissions than if the approach used in the prior study is applied (Hoek et al., 2013). The greater resolution of our work increases the estimated PM_{2.5}-related mortality by a factor of 1.35, but combined these result in a factor of 1.5 compared to the observed factor 2.3 increase.

The remaining factor 1.6 difference in PM_{2.5} mortality between the two studies may be in part due to growth in non-aviation emissions between 2006 and 2015. For this work we use emissions from 2014 in the Community Emissions Data System
470 (CEDS) (Hoesly et al., 2018) as a proxy for anthropogenic, non-aviation emissions. Figure 2 shows that the highest aviation-attributable concentrations of PM_{2.5} are found in Asia, and analysis of CEDS data suggests that emissions of NO and SO₂ from non-aviation sources in Asia (defined as a region bounded by 60 to 150°E and 10°S to 55°N) increased by 20% and

0.16% respectively between 2006 and 2014. In the same period, NO and SO₂ emissions fell by 5.3% and 16% respectively outside of Asia. This change was accompanied by a 9.4% increase in ammonia emissions in Asia, and a 7.3% increase outside. Such changes would increase ambient concentrations of PM_{2.5} precursors, and therefore increase the amount of PM_{2.5} formed as a result of aviation-attributable ozone descending to the surface. Other possible contributors to the increase in PM_{2.5} exposure-related mortality attributable to aviation include changes in baseline mortality rates and increases in the exposed population. For example, using the same geographical boundaries as were applied to calculate changes in Asia's emissions from CEDS, the total population in Asia increased by 11% from 2005 to 2015. A simple linear combination of increases – 20% in NO, 9.4% in ammonia, and 11% in population – would imply a factor 1.45 increase in exposure. However a detailed analysis would be required to fully understand the sources of the unexplained factor 1.6 difference. An additional potential cause of difference is the meteorological year or data used for each study. Barrett et al. (2010a), Eastham and Barrett (2016), and Quadros et al. (2020) each investigated the degree to which a change in meteorology could affect the impacts of aviation on air quality, finding differences of between 5 and 21% in estimated population exposure and health impacts.

4 Discussion

This work finds that aviation's air quality impacts are greater than has been previously estimated. Since prior assessments have shown aviation's monetized impacts on air quality to be similar in magnitude to its impacts on the climate (Grobler et al., 2019), this work suggests that impact mitigation options which do not address NO_x emissions-attributable air quality impacts will therefore not address one of the largest environmental impacts of aviation. This may change the balance of cost-effectiveness when considering trade-offs between CO₂ emissions and NO_x emissions.

These results also imply that coarse-resolution global simulations may underestimate the impacts of aviation on surface air quality. However, moderate-resolution studies at C90 (around 100 km or 1° resolution) resolve 98% of the impacts calculated by a C180 simulation at one quarter of the computational cost (or one eighth if a smaller time step is required). This includes 99% of PM_{2.5} impacts and 97% of ozone impacts. This work does not address the degree to which different parameterizations of physical processes might affect estimates of aviation impacts, as might be relevant when comparing results from regional or local-scale air quality models to those from global-scale models.

This work explains differences in the literature regarding aviation's air quality impacts. Yim et al. (2015) found that the use of finer resolution nested modeling increased estimated ozone exposure by 12% but decreased exposure to PM_{2.5} by 29%, whereas Vennam et al. (2017) found reductions of more than 90% in both. We hypothesize that the differences found between simulations at different resolutions in the former study are due to their use of two inconsistent models to represent the global and nested regions, as the difference in results is within the range reported in a model intercomparison of aviation's air quality effects (Cameron et al., 2017). Vennam et al. (2017) instead used boundary conditions from a single global simulation which included aviation emissions for both their aviation and non-aviation regional simulations of North

505 America, inherently removing the effect of aviation NO_x on hemispheric-scale tropospheric ozone and therefore neglecting the influence of a larger atmospheric response. Our work shows that this global response drives the majority of the change in surface air quality. **Not only do we find that each kilogram of fuel burned outside of US air space still cause about half as many mortalities per kilogram as fuel burned over the US, but we show that the vertical mixing timescales needed for cruise altitude emissions to affect surface concentrations are greater than the zonal mixing timescales (Section 3.3). This implies**
510 **that a model with fixed boundary conditions over a single region cannot capture the influence of cruise altitude emissions – even within the same region – on surface air quality.**

This work describes the potential benefit of a rapid reduction in aviation emissions, and the degree to which different modeling approaches can accurately capture the expected outcome of such a reduction while accounting for non-linearity in the response. Approaches which are intended to perform an attribution of current-day air quality or mortality impacts
515 **between different sectors, regions, or species using methods such as tagging (Emmons et al., 2012; Butler et al., 2018) may find different results. As discussed by Clappier et al. (2017) and Thunis et al. (2021a, b), the relevance of these results to planned policy will therefore depend on the context and objectives of the policy.**

A remaining important uncertainty regards the response of human health to ozone exposure. Although several recent epidemiological analyses have findings consistent with Jerrett et al. (2009) and Turner et al. (2015) that ozone exposure
520 increases morbidity and mortality (Zhao et al., 2021; Lim et al., 2019; Rhee et al., 2019), including at low concentrations (Yazdi et al., 2021), other studies have suggested that ozone exposure may not be significantly associated with mortality (Atkinson et al., 2016; Brunekreef et al., 2021; Huangfu and Atkinson, 2020). This would not affect the finding that over 20,000 mortalities per year due to $\text{PM}_{2.5}$ exposure are attributable to aviation.

There has also been research suggesting that exposure to certain specific components of $\text{PM}_{2.5}$ may be more harmful than
525 other constituents of $\text{PM}_{2.5}$, which may indicate greater impacts related to soot, ultrafine non-volatiles, or organic carbon (Verma et al., 2015). If so, our finding of a greater relative sensitivity of soot exposure to model resolution suggests that the more localized impacts of aviation soot emissions will require higher-resolution simulations or localized modeling approaches such as those in Yim et al. (2015) to quantify.

5 Conclusions

530 Our findings show that 74,300 **(95% confidence interval due to uncertainty in health response: 57,300-91,100)** premature mortalities each year are attributable to aviation emissions, based on 2016 data. This is 4.6 times greater than a previous assessment for 2005 finding 16,000 premature mortalities each year. A factor of 1.29 is due to our use of a global model with greater spatial resolution; a factor 1.84 is due to the availability of an ozone concentration response function which includes wintertime ozone; a factor 1.24 is due to the inclusion of a broader set of diseases in those impacted by ozone exposure; and
535 a factor 1.28 is due to increases in aviation fuel burn. Accounting for the different effects each of these factors has on ozone

and PM_{2.5}-attributable mortality, these alone explain a factor 4.1 increase in the estimate, 87% of the total. We hypothesize that the remaining 13% discrepancy may be the result of regional increases in non-aviation emissions.

We find no evidence of a decrease in air quality impact with increasing model resolution. Instead we find that the simulated impacts of aviation emissions on surface air quality increase by 24% when using a 50 km model compared to a 400 km resolution model. Impacts simulated at a resolution of 100 km globally are within 2% of those at 50 km, suggesting that moderate-resolution simulations are capable of accurately simulating aviation's air quality impacts.

Finally, we show that the impacts of aviation on air quality are global in nature. Emissions from aircraft flying over the US cause 39 mortalities per Tg of fuel burn in the US, compared to 20 mortalities in the US per Tg of fuel burned outside this region. We also find that PM_{2.5} concentrations in China are more strongly affected in absolute terms by aviation emissions within the US domain than concentrations of PM_{2.5} in the US, and that concentrations of aviation-attributable PM_{2.5} are more strongly correlated with the concentration of non-aviation PM_{2.5} than with a marker of local aviation emissions.

Previous work has shown that the monetized air quality impacts of a unit of aviation fuel burn are similar in magnitude to aviation's monetized climate impacts, including effects of contrails. The further increase in estimated air quality impacts we find suggest that aviation-attributable air quality degradation is a significant contributor to aviation's environmental impacts, and that mitigation of full-flight NO_x emissions should be considered alongside ongoing efforts to reduce the effects of aviation on the climate.

Acknowledgements and funding

The MERRA-2 data used in this study have been provided by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center. The core computations in this paper were run on the FASRC Odyssey cluster supported by the FAS Division of Science Research Computing Group at Harvard University. Sensitivity simulations were performed using the Svante cluster at the Massachusetts Institute of Technology. This work was partially supported by the U.S. Federal Aviation Administration (FAA) Office of Environment and Energy through ASCENT, the FAA Center of Excellence for Alternative Jet Fuels and the Environment, project 58 through FAA Award Number 13-C-AJFE-MIT under the supervision of S. Daniel Jacob and Jeetendra Upadhyay. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the FAA.

Code availability statement

The core of this work was completed using version 12.6.2 of the GEOS-Chem High Performance (GCHP) code. The specific codebase used can be found at https://github.com/geoschem/gchp_legacy/releases/tag/12.6.2. Additional sensitivity simulations were performed using version 14.2.2 of the same code as described in Section 2.3.

565 **Author contributions**

SDE, RLS, and SB conceived and designed the study. GC generated the aircraft emissions data. SDE performed all model simulations and analysis. DJJ provided computational resources for high-fidelity simulation. All authors contributed to the manuscript writing and editing.

Competing interests

570 The authors declare no competing interests.

References

- 575 Atkinson, R. W., Butland, B. K., Dimitroulopoulou, C., Heal, M. R., Stedman, J. R., Carslaw, N., Jarvis, D., Heaviside, C., Vardoulakis, S., Walton, H., and Anderson, H. R.: Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies, *BMJ Open*, 6, e009493, 2016.
- Barrett, S. R. H., Britter, R. E., and Waitz, I. a.: Global mortality attributable to aircraft cruise emissions, *Environ. Sci. Technol.*, 44, 7736–7742, 2010a.
- 580 Barrett, S. R. H., Prather, M., Penner, J., Selkirk, H., Doppelheuer, A., Fleming, G., Gupta, M., Halthore, R., Hileman, J., Jacobson, M., Kuhn, S., Miake-lye, R., Petzold, A., Roof, C., Schumann, U., Waitz, I., and Wayson, R.: Guidance on the use of AEDT Gridded Aircraft Emissions in Atmospheric Models, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2010b.
- Brunekreef, B., Strak, M., Chen, J., Andersen, Z. J., Atkinson, R., Bauwelinck, M., Bellander, T., Boutron-Ruault, M.-C., Brandt, J., Carey, I., and Others: Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM_{2.5}, BC, NO₂, and O₃: An Analysis of European Cohorts in the ELAPSE Project, 2021.
- 585 Burnett, R., Chen, H., Szyszkwicz, M., Fann, N., Hubbell, B., Pope, C. A., 3rd, Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., Lim, C. C., Turner, M. C., Jerrett, M., Krewski, D., Gapstur, S. M., Diver, W. R., Ostro, B., Goldberg, D., Crouse, D. L., Martin, R. V., Peters, P., Pinault, L., Tjepkema, M., van Donkelaar, A., Villeneuve, P. J., Miller, A. B., Yin, P., Zhou, M., Wang, L., Janssen, N. A. H., Marra, M., Atkinson, R. W., Tsang, H., Quoc Thach, T., Cannon, J. B., Allen, R. T., Hart, J. E.,
- 590 Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin, H., and Spadaro, J. V.: Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter, *Proc. Natl. Acad. Sci. U. S. A.*, 115, 9592–9597, 2018.
- Butler, T., Lupascu, A., Coates, J., and Zhu, S.: TOAST 1.0: Tropospheric ozone attribution of sources with tagging for CESM 1.2.2, *Geosci. Model Dev.*, 11, 2825–2840, 2018.
- 595 Cameron, M. A., Jacobson, M. Z., Barrett, S. R. H., Bian, H., Chen, C. C., Eastham, S. D., Gettelman, A., Khodayari, A., Liang, Q., Selkirk, H. B., Unger, N., Wuebbles, D. J., and Yue, X.: An intercomparative study of the effects of aircraft emissions on surface air quality, *J. Geophys. Res. D: Atmos.*, 122, 8325–8344, 2017.
- Center for International Earth Science Information Network - CIESIN - Columbia University: Gridded Population of the World, Version 4 (GPWv4): Population Count, <https://doi.org/10.7927/H4X63JVC>, 2016.
- 600 Christian, K. E., Brune, W. H., Mao, J., and Ren, X.: Global sensitivity analysis of GEOS-Chem modeled ozone and hydrogen oxides during the INTEX campaigns, *Atmos. Chem. Phys.*, 18, 2443–2460, 2018.
- Clappier, A., Belis, C. A., Pernigotti, D., and Thunis, P.: Source apportionment and sensitivity analysis: two methodologies with two different purposes, *Geosci. Model Dev.*, 10, 4245–4256, 2017.
- 605 Dasadhikari, K., Eastham, S. D., Allroggen, F., Speth, R. L., and Barrett, S. R. H.: Evolution of sectoral emissions and contributions to mortality from particulate matter exposure in the Asia-Pacific region between 2010 and 2015, *Atmos. Environ.*, 116916, 2019.
- Eastham, S. D. and Barrett, S. R. H.: Aviation-attributable ozone as a driver for changes in mortality related to air quality and skin cancer, *Atmos. Environ.*, <https://doi.org/10.1016/j.atmosenv.2016.08.040>, 2016.

- 610 Eastham, S. D., Weisenstein, D. K., and Barrett, S. R. H.: Development and evaluation of the unified tropospheric–stratospheric chemistry extension (UCX) for the global chemistry-transport model GEOS-Chem, *Atmos. Environ.*, 89, 52–63, 2014.
- 615 Eastham, S. D., Long, M. S., Keller, C. A., Lundgren, E., Yantosca, R. M., Zhuang, J., Li, C., Lee, C. J., Yannetti, M., Auer, B. M., Clune, T. L., Kouatchou, J., Putman, W. M., Thompson, M. A., Trayanov, A. L., Molod, A. M., Martin, R. V., and Jacob, D. J.: GEOS-Chem High Performance (GCHP v11-02c): A next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications, *Geosci. Model Dev.*, 2941–2953, 2018.
- Emmons, L. K., Hess, P. G., Lamarque, J.-F., and Pfister, G. G.: Tagged ozone mechanism for MOZART-4, CAM-chem and other chemical transport models, *Geoscientific Model Development*, 5, 1531–1542, 2012.
- Grobler, C., Wolfe, P. J., Dasadhikari, K., Dedoussi, I. C., Allroggen, F., Speth, R. L., Eastham, S. D., Agarwal, A., Staples, Sabnis, J., and Barrett, S. R. H.: Marginal Climate and Air Quality Costs of Aviation Emissions, *Environ. Res. Lett.*, 2019.
- 620 Hileman, J. I., Stratton, R. W., and Donohoo, P. E.: Energy Content and Alternative Jet Fuel Viability, *J. Propul. Power*, 26, 1184–1196, 2010.
- Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., and Kaufman, J. D.: Long-term air pollution exposure and cardio- respiratory mortality: a review, *Environ. Health*, 12, 43, 2013.
- 625 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O’Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geoscientific Model Development*, 11, 369–408, 2018.
- Holmes, C. D., Tang, Q., and Prather, M. J.: Uncertainties in climate assessment for the case of aviation NO, *Proc. Natl. Acad. Sci. U. S. A.*, 108, 10997–11002, 2011.
- 630 Hu, L., Keller, C. A., Long, M. S., Sherwen, T., Auer, B., Da Silva, A., Nielsen, J. E., Pawson, S., Thompson, M. A., Trayanov, A. L., Travis, K. R., Grange, S. K., Evans, M. J., and Jacob, D. J.: Global simulation of tropospheric chemistry at 12.5 km resolution: performance and evaluation of the GEOS-Chem chemical module (v10-1) within the NASA GEOS Earth system model (GEOS-5 ESM), *Geosci. Model Dev.*, 11, 4603–4620, 2018.
- 635 Huang, M., Carmichael, G. R., Pierce, R. B., Jo, D. S., Park, R. J., Flemming, J., Emmons, L. K., Bowman, K. W., Henze, D. K., Davila, Y., Sudo, K., Jonson, J. E., Lund, M. T., Janssens-Maenhout, G., Dentener, F. J., Keating, T. J., Oetjen, H., and Payne, V. H.: Impact of intercontinental pollution transport on North American ozone air pollution: an HTAP phase 2 multi-model study, *Atmos. Chem. Phys.*, 17, 5721–5750, 2017.
- Huangfu, P. and Atkinson, R.: Long-term exposure to NO₂ and O₃ and all-cause and respiratory mortality: A systematic review and meta-analysis, *Environ. Int.*, 144, 105998, 2020.
- 640 ICAO: Annex 16: Volume II, in: *The Convention on International Civil Aviation*, vol. II, 2017.
- Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., and Thun, M.: Long-term ozone exposure and mortality, *N. Engl. J. Med.*, 360, 1085–1095, 2009.
- 645 Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Jimenez, J. L., Campuzano-Jost, P., Froyd, K. D., Liao, J., Hair, J. W., Fenn, M. A., Butler, C. F., Wagner, N. L., Gordon, T. D., Welti, A., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Teng, A. P., Millet, D. B., Schwarz, J. P., Markovic, M. Z., and Perring, A. E.: Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model, *Atmos. Chem. Phys.*, 15, 10411–10433, 2015.

- 650 Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., Doherty, S. J., Freeman, S., Forster, P. M., Fuglestedt, J., Gettelman, A., De León, R. R., Lim, L. L., Lund, M. T., Millar, R. J., Owen, B., Penner, J. E., Pitari, G., Prather, M. J., Sausen, R., and Wilcox, L. J.: The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018, *Atmos. Environ.*, 117834, 2020.
- Lee, H., Olsen, S. C., Wuebbles, D. J., and Youn, D.: Impacts of aircraft emissions on the air quality near the ground, *Atmos. Chem. Phys.*, 13, 5505–5522, 2013.
- 655 Lim, C. C., Hayes, R. B., Ahn, J., Shao, Y., Silverman, D. T., Jones, R. R., Garcia, C., Bell, M. L., and Thurston, G. D.: Long-Term Exposure to Ozone and Cause-Specific Mortality Risk in the United States, *Am. J. Respir. Crit. Care Med.*, 200, 1022–1031, 2019.
- Lin, M., Fiore, A. M., Cooper, O. R., Horowitz, L. W., Langford, A. O., Levy, H., Johnson, B. J., Naik, V., Oltmans, S. J., and Senff, C. J.: Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, *J. Geophys. Res. D: Atmos.*, 117, <https://doi.org/10.1029/2012JD018151>, 2012.
- 660 Liu, H., Jacob, D. J., Bey, I., and Yantosca, R. M.: Constraints from ^{210}Pb and ^7Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields, *J. Geophys. Res.*, 106, 12109–12128, 2001.
- Miller, C. J., Prashanth, P., Allroggen, F., Grobler, C., Sabnis, J. S., Speth, R. L., and Barrett, S. R. H.: An Environmental Cost Basis for Regulating Aviation NO_x Emissions, *Environmental Research Communications*, 2022.
- 665 Moch, J. M., Dovrou, E., Mickley, L. J., Keutsch, F. N., Liu, Z., Wang, Y., Dombek, T. L., Kuwata, M., Budisulistiorini, S. H., Yang, L., Decesari, S., Paglione, M., Alexander, B., Shao, J., Munger, J. W., and Jacob, D. J.: Global Importance of Hydroxymethanesulfonate in Ambient Particulate Matter: Implications for Air Quality, *J. Geophys. Res. D: Atmos.*, 125, e2020JD032706, 2020.
- 670 Phoenix, D., Khodayari, A., Wuebbles, D., and Stewart, K.: Aviation impact on air quality present day and mid-century simulated in the Community Atmosphere Model (CAM), *Atmos. Environ.*, 196, 125–132, 2019.
- Potts, D. A., Marais, E. A., Boesch, H., Pope, R. J., Lee, J., Drysdale, W., Chipperfield, M. P., Kerridge, B., Siddans, R., Moore, D. P., and Remedios, J.: Diagnosing air quality changes in the UK during the COVID-19 lockdown using TROPOMI and GEOS-Chem, *Environ. Res. Lett.*, 16, 054031, 2021.
- 675 Prashanth, P., Speth, R. L., Eastham, S. D., Sabnis, J. S., and Barrett, S.: Post-Combustion Emissions Control in Aero-Gas Turbine Engines, *Energy Environ. Sci.*, <https://doi.org/10.1039/D0EE02362K>, 2020.
- Prashanth, P., Eastham, S. D., Speth, R. L., and Barrett, S. R. H.: Aerosol formation pathways from aviation emissions, *Environ. Res. Commun.*, <https://doi.org/10.1088/2515-7620/ac5229>, 2022.
- Punger, E. M. and West, J. J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the United States, *Air Qual. Atmos. Health*, 6, <https://doi.org/10.1007/s11869-013-0197-8>, 2013.
- 680 Putman, W. M. and Lin, S. J.: Finite-volume transport on various cubed-sphere grids, *J. Comput. Phys.*, 227, 55–78, 2007.
- Quadros, F. D. A., Snellen, M., and Dedoussi, I. C.: Regional sensitivities of air quality and human health impacts to aviation emissions, *Environ. Res. Lett.*, 15, 105013, 2020.
- Rhee, J., Dominici, F., Zanobetti, A., Schwartz, J., Wang, Y., Di, Q., Balmes, J., and Christiani, D. C.: Impact of Long-Term Exposures to Ambient $\text{PM}_{2.5}$ and Ozone on ARDS Risk for Older Adults in the United States, *Chest*, <https://doi.org/10.1016/j.chest.2019.03.017>, 2019.
- 685

- Shah, V., Jacob, D. J., Dang, R., Lamsal, L. N., Strode, S. A., Steenrod, S. D., Boersma, K. F., Eastham, S. D., Fritz, T. M., Thompson, C., Peischl, J., Bourgeois, I., Pollack, I. B., Nault, B. A., Cohen, R. C., Campuzano-Jost, P., Jimenez, J. L., Andersen, S. T., Carpenter, L. J., Sherwen, T., and Evans, M. J.: Nitrogen oxides in the free troposphere: implications for tropospheric oxidants and the interpretation of satellite NO₂ measurements, *Atmos. Chem. Phys.*, 23, 1227–1257, 2023.
- 690 Simone, N. W., Stettler, M. E. J., and Barrett, S. R. H.: Rapid estimation of global civil aviation emissions with uncertainty quantification, *Transp. Res. Part D: Trans. Environ.*, 25, 33–41, 2013.
- Skowron, A., Lee, D. S., De León, R. R., Lim, L. L., and Owen, B.: Greater fuel efficiency is potentially preferable to reducing NO_x emissions for aviation’s climate impacts, *Nat. Commun.*, 12, 564, 2021.
- Thunis, P., Clappier, A., Beekmann, M., Putaud, J. P., Cuvelier, C., Madrazo, J., and de Meij, A.: Non-linear response of PM_{2.5} to changes in NO_x and NH₃ emissions in the Po basin (Italy): consequences for air quality plans, *Atmos. Chem. Phys.*, 21, 9309–9327, 2021a.
- Thunis, P., Clappier, A., de Meij, A., Pisoni, E., Bessagnet, B., and Tarrason, L.: Why is the city’s responsibility for its air pollution often underestimated? A focus on PM_{2.5}, *Atmos. Chem. Phys.*, 21, 18195–18212, 2021b.
- 700 Turner, M. C., Jerrett, M., Pope, C. A., III, Krewski, D., Gapstur, S. M., Diver, R. W., Beckerman, B. S., Marshall, J. D., Su, J. G., Crouse, D. L., and Burnett, R. T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, *Am. J. Respir. Crit. Care Med.*, 64, 897–905, 2015.
- Vennam, L. P., Vizuete, W., Talgo, K., Omary, M., Binkowski, F. S., Xing, J., Mathur, R., and Arunachalam, S.: Modeled Full-Flight Aircraft Emissions Impacts on Air Quality and Their Sensitivity to Grid Resolution, *J. Geophys. Res. D: Atmos.*, 122, 13472–13494, 2017.
- 705 Verma, V., Fang, T., Xu, L., Peltier, R. E., Russell, A. G., Ng, N. L., and Weber, R. J.: Organic aerosols associated with the generation of reactive oxygen species (ROS) by water-soluble PM_{2.5}, *Environ. Sci. Technol.*, 49, 4646–4656, 2015.
- Vohra, K., Vodonos, A., Schwartz, J., Marais, E. A., Sulprizio, M. P., and Mickley, L. J.: Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem, *Environ. Res.*, 195, 110754, 2021.
- 710 Wang, Q., Jacob, D. J., Spackman, J. R., Perring, A. E., Schwarz, J. P., Moteki, N., Marais, E. A., Ge, C., Wang, J., and Barrett, S. R. H.: Global budget and radiative forcing of black carbon aerosol: Constraints from pole-to-pole (HIPPO) observations across the Pacific, *J. Geophys. Res.*, 119, 195–206, 2014.
- Whitt, D. B., Jacobson, M. Z., Wilkerson, J. T., Naiman, a. D., and Lele, S. K.: Vertical mixing of commercial aviation emissions from cruise altitude to the surface, *J. Geophys. Res.*, 116, 1–13, 2011.
- 715 Wilkerson, J. T., Jacobson, M. Z., Malwitz, A., Balasubramanian, S., Wayson, R., Fleming, G., Naiman, A. D., and Lele, S. K.: Analysis of emission data from global commercial aviation: 2004 and 2006, *Atmos. Chem. Phys.*, 10, 6391–6408, 2010.
- Williams, R. S., Hegglin, M. I., Kerridge, B. J., Jöckel, P., Latter, B. G., and Plummer, D. A.: Characterising the seasonal and geographical variability in tropospheric ozone, stratospheric influence and recent changes, *Atmos. Chem. Phys.*, 19, 3589–3620, 2019.
- World Health Organization: WHO methods and data sources for country-level causes of death 2000-2016, WHO, 2018.
- 720 Yazdi, M. D., Wang, Y., Di, Q., Requia, W. J., Wei, Y., Shi, L., Sabath, M. B., Dominici, F., Coull, B., Evans, J. S., Koutrakis, P., and Schwartz, J. D.: Long-term effect of exposure to lower concentrations of air pollution on mortality among US Medicare participants and vulnerable subgroups: a doubly-robust approach, *Lancet Planet Health*, 5, e689–e697, 2021.

Yim, S. H. L., Lee, G. L., Lee, I. W., Allroggen, F., Ashok, A., Caiazzo, F., Eastham, S. D., Malina, R., and Barrett, S. R. H.: Global, regional and local health impacts of civil aviation emissions, *Environ. Res. Lett.*, 2015.

725 Yu, K., Keller, C. A., Jacob, D. J., Molod, A. M., Eastham, S. D., and Long, M. S.: Errors and improvements in the use of archived meteorological data for chemical transport modeling: an analysis using GEOS-Chem v11-01 driven by GEOS-5 meteorology, *Geoscientific Model Development*, 11, 305–319, 2018.

Zhang, L., Jacob, D. J., Downey, N. V., Wood, D. A., Blewitt, D., Carouge, C. C., van Donkelaar, A., Jones, D. B. A., Murray, L. T., and Wang, Y.: Improved estimate of the policy-relevant background ozone in the United States using the GEOS-Chem global model with $1/2^\circ \times 2/3^\circ$ horizontal resolution over North America, *Atmos. Environ.*, 45, 6769–6776, 2011.

730 Zhao, N., Pinault, L., Toyib, O., Vanos, J., Tjepkema, M., and Cakmak, S.: Long-term ozone exposure and mortality from neurological diseases in Canada, *Environ. Int.*, 157, 106817, 2021.