Replies to reviewer's comments on "Impacts of aviation fuel sulfur content on climate and human health" by Z. Z. Kapadia et al.

Correspondence to: Z. Z. Kapadia (pm08zzk@leeds.ac.uk)

We would like to thank both of the anonymous reviewers for their helpful and constructive comments.

Reviewer	Comment number		Comment text	How the comment has been addressed/response
1	1	a	***Overall the work appears to be carefully executed. The paper would benefit from a better articulation of how this treatment differs from published previous work (including better explanation of different results).***	We have added a more detailed description of how our work differs from previous studies. See our responses to points 1b and 4 below (from reviewer 1), and point 2 from reviewer 2. We have explicitly mentioned the current broad range of literature estimates in the aerosol direct radiative and aerosol cloud albedo effects (section 1, pp.2, paragraph 3), and how our work is used to re-evaluate these radiative effects using a coupled tropospheric chemistry-aerosol microphysics (section 1, pp.3, paragraph 5). The investigation of the use of desulfurised fuel, ULSJ fuel and variations in FSC above the cruise phase of flight are now explicitly mentioned to identify how this work differs from previous work (section 1, pp.3, paragraph 5).
		b	***and more description/model evaluation for the chemistry results***	The model description has been extended (section 2.1). This has been done by splitting section 2.1 "Coupled chemistry-aerosol microphysics model" in to two parts: 2.1.1 Model Description and; 2.1.2 Model Evaluation. As suggested, we have added a section on model evaluation (section 2.1.2). The model has been extensively evaluated in previous studies and we now describe these previous evaluations in more detail (pp.4, Section 2.1.2). In this study, we focus our new evaluation on aspects of the model that are most pertinent to this study - namely the vertical profile of sulphate and nitrate aerosol. We evaluate simulated vertical profiles of speciated aerosol mass concentrations against aircraft observations from Heald et al., (2011) – section 2.1.2. We add a new figure (new Fig. 1 – shown at the end of this document) that summarises aerosol model-observation comparisons. Overall, we demonstrate that

				globally the model overestimates sulfate, and underestimates nitrate, ammonium and OA. After broad stratification of field campaigns in to polluted, biomass burning and remote regions (as per campaign classifications used by Heald et al., (2011)) we find that the model best performs over polluted regions, with lower skill over remote regions, and greatest underestimation over biomass burning regions. Additionally, we add another new figure (new Fig. 2 – shown at the end of this document) where we demonstrate the model's ability to simulate ozone in the UTLS. In comparison to observational ozonesonde profiles compiled by Tilmes et al., (2012) from 41 different global locations from between 1995 to 2011, we find that globally the model slightly overestimates ozone.
		c	***The mortality methodology needs better explanation and improvement, as it uses out-of-date concentration-response functions, and should include more discussion of the appropriateness of using such factors worldwide.***	We thank the referee for pointing out how better explanation of the mortality methodology could aid the paper. As such we have included an in-depth explanation of the mortality methodology– section 2.5. While the concentration-response function utilised in this study may not be the most recently published function, it allows for this study to provide estimates of aviation-induced mortality directly comparable with previous work from Barrett et al., (2012) and Yim et al., (2015) – please also refer to our response to comment 6. As stated by Butt et al., (2016), the CRFs employed in this study are based on the American Cancer Society Prevention cohort study. WHO recommend the use of a log-linear model (as used here), as linear models could result in unrealistically large RR values when high $PM_{2.5}$ concentrations are considered ($PM_{2.5}$ > 30 μg m^{-3}); as also stated the supplementary information for Barrett et al., (2012). Though the application of the same CRF factor is global (as per Ostro et al., (2004) and recommendation from the WHO), we use regional population data and regional baseline mortalities for both cardiopulmonary disease and lung cancer.
	2		***Specific comments: abstract, line 7, line 16: significant figures***	Taking in to account that the values reported are estimates we have rounded mortality

			on annual mortality? Are 3597 and 624 really the estimates?	estimates up to the nearest 10.
	3		***The introduction could better establish what is not known, and what this study contributes relative to the work that has been done before, especially Barrett et al. 2010, 2012, and Morita et al. 2014.***	To better establish what is known, and the differences in estimates of premature aviation-induced mortality from other studies, we have included a paragraph in the introduction (pp.2, section 1, last paragraph) – "Barrett et al. (2012) and Barrett et al. (2010) using the methodology outlined by Ostro (2004), estimated that aviation emissions are responsible for ~10,000 premature mortalities a^{-1} globally, due to increases in cases of cardiopulmonary disease and lung cancer. Yim et al. (2015) revised this estimated, using the same methodology, to 13,920 (95% CI: 7,220–20,880) mortalities a^{-1}. Morita et al. (2014) using the methodology to derive the relative risk (RR) from exposure to surface $PM_{2.5}$ from Burnett et al. (2014) estimate aviation results in 405 (95% CI: 182–648) mortalities a^{-1} due to increases in cases of lung cancer, stroke, ischemic heart disease, trachea, bronchus, and chronic obstructive pulmonary disease. Jacobson et al. (2013) using the methodology from Jacobson (2010) estimate 310 (95% CI: –400 to 4,300) mortalities a^{-1} due to cardiovascular effects. These studies demonstrate that the different methodologies employed and modes of mortality considered produce a wide range in estimated mortalities due to aviation emissions of between 310 – 16,000 mortalities a^{-1}". Additionally, to acknowledge what is not known we have highlighted the large uncertainty in estimates in the aviation-induced cloud albedo effect and how this paper investigates the cloud albedo effect for all FSC cases, through stating in the section 1 (pp.2, paragraph 3) "Few studies estimate the cloud albedo effect from aviation-sourced aerosol. Those that have estimated this effect, show large uncertainties: Righi et al. (2013) assessed the aCAE to be –15.4 mW m^{-2} ± 69%, while Gettelman and Chen (2013) report an uncertainty of ± 52% associated with their estimate of –21 mW m^{-2}". Along with stating in Section 1 (pp3, paragraph 4) "; thus evaluating the ozone and aerosol direct radiative effect, and cloud albedo effect".
	4		***Section 2.1. Some model evaluation would be useful here, including information on stratosphere-troposphere***	We thank the reviewer for pointing out the value of model evaluation here. We now include speciated aerosol model-observation comparisons from a global synthesis of

			exchange.	aerosol mass spectrometer measurements (see response to comment 1b above and our new Fig. 1). It is unclear which datasets could be used to evaluate stratosphere-troposphere exchange directly, however we now include comparisons of model-simulated with vertical profiles of ozone and aerosol in the upper troposphere with observations from ozonesonde and aircraft. The comparisons described show no evidence of stratosphere-troposphere exchange being a problem in the model. Examining the shape of vertical profile comparisons (new Fig. 2), we find no evidence of systematic model bias in the upper troposphere.
	5	a	***Section 2.2. While the authors explain differences in estimates from ranges esp. SO2, OC, CO which are outside previous ranges, a bit more information is warranted here. Specifically, why do the authors think that the fuel burn inventory, or OC emissions index, better reflects reality?***	Our intention was to develop a comprehensive aviation emissions dataset that includes an expansive set of pollution species emitted by aviation. Many previous aviation emission datasets only include a subset of emitted species. For example, the CMIP5 emissions dataset only includes aviation emissions of NOx and BC. We use emissions indices, which are shown to be in agreement and consistent with previous work to produce estimates of additional species emissions (Table 1). There are few previous estimates of global OC aircraft emissions. Wilkerson et al., (2010) using data from Wayson et al., (2009) assume a greater emission index for OC (of 0.015 g kg(fuel)$^{-1}$) but note that "this is intended for airport operations at ground level conditions rather than cruise-related operations" and as such may overestimate OC emissions. We use the relationship between BC and OC of 4:1 (Bond et al., 2004; Hopke 1985) to derive an OC emissions index of 0.00625 g kg(fuel)$^{-1}$. Without developing models for the complex relationship between operating conditions, fuel flow, aircraft type, ambient conditions, and actual flight paths, we see this as a viable and appropriate approach. We add a short statement to the paper (pp.8, section 2.2, paragraph 4): The lower EI_{OC} applied here (in comparison to Wilkerson et al. (2010)) is due to the phase of flight considered when deriving the AEDT emissions inventory. Wilkerson et al. (2010) derive EI_{OC} focusing on airport operations at ground level

				conditions while we consider aircraft operations after ground idle conditions, acknowledging the risk of overestimating aviation OC emissions.
		b	***Given a fast-growing sector (especially in highly-populated areas such as Asia), why is the year 2000 still relevant? Some comment on the effect this choice has on results would be warranted.***	While we appreciate the fast-growing nature of the sector, investigations on the impacts of aviation for year 2000 are still pertinent given current literature which use year 2000 as a base for comparison for the future impact of aviation e.g. Righi et al., (2015). This choice allows for estimates of the radiative effect from aviation emissions to be compared against current literature. We add a statement to our paper to further acknowledge that aircraft emissions are growing rapidly and that emissions in 2015 are greater than 2001 (pp.2, section 1, paragraph 1), "while demonstrating a 85% rate of growth in Revenue Passenger Kilometres (RPK's) between 2001–2015 (Airbus, 2015)".
	6		***Section 2.5. The health effects calculation uses older concentration-response functions that are not state-of-the-science. The authors should revisit their choice here. They should consider using a concentration-response function consistent with previous work to enable comparisons, even if this is just as a sensitivity study.*** ***The functional form should also be given here, and its uncertainty discussed.***	While we acknowledge that the CRF function employed here is older, use of this CRF allows our estimates to be compared to previous studies using the same CRF (Barret et al., (2010), Barret et al., (2012), Yim et al., (2015)). We have added a more detailed description of the function (pp.9, section 2.5). We add the following (pp.20): "Future work needs to estimate the health impacts of aviation using newly available concentration response functions (Burnett et al., 2014)."
	7		***Section 3.1. It is unclear why there are 'increases' under the NORM scenario? Relative to what?***	These are increases relative to a scenario with no aviation emissions, i.e. the NOAVI simulation. We have edited the text to make this clear.
	8		***Figures 1-3 and 5: axes and text a bit too small to read.***	Text size within figures has been increased to enhance readability.
	9		***p 18932 I'm surprised by the strong linearity (R2=1?) of PM2.5 to sulfur content. While I'm not surprised that this is roughly linear, an R2=1 suggests to me that important potentially nonlinear parameters might not have been included in the model. Can the authors comment on this?***	R^2 values of 1 were arrived after rounding to 2 d.p. – these values have been amended in the text to $R^2 > 0.999$ to make clear that these are not precisely unity. The model includes many non-linear processes, including chemistry and aerosol microphysics. We agree with the referee that the linear response is may seem surprising. The near-linear response is likely due to the small emission perturbations that we have applied relative to global aerosol emissions.

				We add the following statement (pp.12, section 3.1, paragraph 5): "Larger emission perturbations would likely lead to a non-linear response in atmospheric aerosol".
	10	a	***p 18932 line 20: Why is the estimate of sulfate attributable to NOx so different from Barrett et al. 2010?***	Thank you for pointing this out to us. As far as we can tell, the values are actually rather similar, although the Barrett et al study does not give an exact figure (we estimate a value of 36.2% and Barrett et al., (2010) estimate a value of "more-than half"). We have rephrased the text to: "Barrett et al., (2010) using GEOS-Chem previously found for a standard FSC (FSC = 600 ppm) that more than half of aviation-attributable sulfates formed at the surface are associated with aviation NO_X emissions and not aviation SO_2 emissions. Here using GMV4-nitrate we find that for a standard FSC aviation is responsible for 36.2% of the aviation-induced sulfates at the surface. Differences in estimates of surface sulphate produced due to aviation-emitted NOx can be attributed to differences model chemistry and microphysics, and different inventories of aviation NOx emissions." – (Section 3.1, pp.14, paragraph 2).
		b	***What differences are there between the models? Is it more likely to be chemistry or transport parameters?***	There are likely to be a variety of differences between the models: the emissions, transport and chemical schemes, and that GLOMAP-mode is a size resolved model. Without conducting a model comparison experiment that includes sensitivities to different parameters, it is difficult to determine whether differences in transport or chemistry are most important.
	11	a	***3.2. A comparison of how differences in premature mortalities are affected by the choice and assumed slope of CRFs is needed here.***	Thank you highlighting this point. The 95% confidence interval range highlights how the slope of the CRF employed effects estimates in premature mortality. Additionally, we have included a description of the main uncertainties involved and captured within the confidence interval reported evaluating the long-term health effects of exposure to $PM_{2.5}$, referring to Ostro (2004) (pp.15, section 3.2, paragraph 2).

		b	***Concentration is not the only difference from previous work. Also, what about comparing to the results of Morita et al. (2014) in their present day scenario? Why are the USLJ changes different from Barrett et al. 2012?***	It is true that concentrations will not be the only difference and factors driving differences in aviation-induced mortality estimates. Differences in cause-specific coefficients (β) will also play a role. To acknowledge this, the following statement has been added (pp.14, section 3.2): Additionally, differences in mortality assessed here in comparison to estimates from Barrett et al. (2012) can be attributed to differences in the disease specific cause-specific coefficients (β), where this study uses β values recommended by Ostro. (2004), while Barrett et al. (2012) derive a β values for cardiopulmonary disease based on a relationship between the β values for lung cancer and the "All Cause" mortality function. As such the functional forms used here and by Barrett et al. (2012) are inherently the same, while differences in β values which drive these functions will partly explain differences in aviation-induced mortality estimates. We cannot confirm the effect of the β values used by Barrett et al., (2012) as these are not provided in their paper or supplementary information. A comparison between estimates in aviation-induced premature mortalities from lung cancer evaluated by this study (390 [95% CI: 150–640] mortalities a^{-1}) and Morita et al., (2014) (41 [95%: 7–67] mortalities a^{-1}) is now discussed. The main reason for the differences in estimates of aviation-induced premature mortality is due to the different CRFs used, with Morita et al., (2014) using the IER (integrated exposure response) methodology outlined by Burnett et al., (2014). The IER methodology is described in order to identify how the two methodologies differ, thus resulting in different estimates in aviation-induced premature mortality – (section 3.2, pp.15, paragraph 3). While we appreciate that through the application of a ULSJ fuel strategy Barrett et al., (2012) estimate that premature 2,300 mortalities could be avoided in comparison to our estimate of 620 mortalities avoided (estimates which differ by a factor of 3.7), we see similar rates of reduction in mortalities avoided when relative values are considered: we estimate a reduction in mortalities of 17.3%, while Barrett et al., (2012) estimates a reduction in aviation-induced mortalities of

				23%. Differences in estimates of mortalities avoided between this study and Barrett et al., (2012) for the ULSJ will again be a function of different β values used and differences in simulated surface-layer $PM_{2.5}$ concentrations, in both the base case and ULSJ.
	12		***Figure 8 is perhaps the most unique part of this work and deserves a bit more discussion.***	Additional attention has been given to this figure (now called Fig. 10), linking in to reply to comment 15 from reviewer 2.
2	1		***The main concern with this study is the use of an off-line model to study climate impacts of aircraft emissions. The model description says that that the GLOMAP-mode is embedded within the 3-D off-line Eulerian CTM to make it a coupled chemistry-aerosol microphysics model. It seems that the meteorological and chemical processes are not coupled. A discussion on the justification of an offline CTM to study climate impacts will greatly strengthen this paper.***	When referring to a "coupled chemistry-aerosol" model we mean that the aerosol microphysics and gas-phase chemistry are coupled. The referee is correct that our model is a chemical transport model, using offline meteorology and that there is no coupling between chemistry and meteorology. We now explain this more clearly in the paper (see Section 2.1.1, pp.3, paragraph 6). The advantage is that we can compare our short 1-year simulations with each other directly, since the meteorology in each run is identical, meaning we are only looking at chemical changes due to changes in emissions. We can then calculate the climate effects of our changes offline using the radiative transfer model.

	2		***There is no discussion of evaluation of the model – either for meteorological variables or for air pollutant concentrations. While this is not at the core of this study, model evaluation is an essential prerequisite for any application study like this. It is suggested that the authors include results from the evaluation, and to specifically focus on the model's ability to predict both PM2.5 mass and speciated components in different parts of the world for the year studied.***	We thank the reviewer for highlighting the added value model evaluation can bring to the manuscript. We have added a model evaluation section (section 2.1.2) which details previous model evaluation work conducted on TOMCAT-GLOMAP-mode, and aerosol (sulfate, nitrate, ammonium and organic aerosol) and ozone specific to the nitrate-extended version of the TOMCAT-GLOMAP-mode coupled model used in this study. The model's ability to simulate sulfate, nitrate, ammonium and organic aerosols is evaluated against observations of aerosol mass from aircraft field campaigns compiled by Heald et al., (2011) (see new Fig. 1 and associated text)., Model-simulated ozone profiles are evaluated against ozonesonde profiles compiled by Tilmes et al., (2012). Please also see our response to referee #1 (comments 1b and 4).
	3		***There is no discussion of the emissions inventories used for non-aircraft sources. While documenting the source of these for this study, putting those in context with the other key studies referred in this study is important. Aircraft emissions react with background emissions from other sources such as NH3 to form aviation-attributable PM2.5, specifically inorganic PM2.5 which is at the core of this study. So, a discussion of NH3 emissions used in this study is critical but lacking.***	The model description section has been reordered and extended (section 2.1.1, pp.3-4) in order to: 1. Provide a description of the hybrid solver employed to simulate the dissolution of of semi-volatile inorganic gases (such as H2O, HNO3, HCl and NH3) in the aerosol-liquid-phase. 2. Identify where information on cloud fraction and cloud top pressure fields are taken from (ISCCP-D2), and, 3. Highlight the differences between the TOMCAT CTM and p-TOMCAT CTM. 4. Provide a description of sources of non-aviation emissions, 5. Sources of anthropogenic and natural emission sources are now listed in the model description section – stating that NH3 emissions are from the EDGAR inventory (Bouwmann et al., 1997).
	4		***Since Barrett et al (2012) used 3 different models, two of which were applied globally, when comparison are made to Barrett et al (2012), it is helpful to know which of the two models are being referred to in this study.***	We thank the reviewer for making this point. For clarity within the manuscript, and where it seems appropriate we have stated that results from this study are being compared to Barret et al.,'s simulations using GEOS-Chem.
	5		***Suggest including findings from two recent studies, and put these results into context. The first one is by Morita et al, ES&T 2014 which was published last***	We thank the reviewer for this suggestion. We integrated findings from Morita et al., (2014) and Brasseur et al., (2015) in to the paper. Estimates in the ozone and aerosol direct

			year and is relevant from the health risk aspects of aircraft emissions using the NASA GISS ModelE2, and the other is more recent one by Brasseur et al, BAMS 2015, which is relevant from the climate impact aspects of aircraft emissions using multiple global-scale models.	radiative, and cloud albedo effects from Brasseur et al., (2015) have been included in a new paragraph added (pp.2, para 3, section 1) in order to help put in to context and convey current estimates for aviation. Morita et al., (2014) has been mentioned, reporting their estimates in mortality from standard aviation, while making reference to the methodology used as this will impact how directly comparable the values estimated by Morita et al., (2014) are with the mortality estimates derived here (pp.15, section 3.2, paragraph 3).
	6		***Section 1 Pg 18926 Line 10: "A coupled tropospheric chemistry-aerosol microphysics model including nitrate aerosol…" Why the emphasis on nitrate aerosol, and not inorganic PM in general? Can the authors clarify this?***	The explicit mention of nitrate aerosols was included as other versions of the TOMCAT-GLOMAP-mode coupled model do not include the formation of nitrate aerosols. This emphasis has been removed now.
	7		***Section 2.1 Pg 18926 This study has used TOMCAT, and one of the models used by Barrett et al (2012) against which several comparisons are made in this study was p-TOMCAT. Since the names are so close to each other, a brief discussion of how these two models are different will be relevant for the sake of the comparisons presented.***	Both TOMCAT and p-TOMCAT started from the same model version several years ago, but their development has now diverged. The main differences are: our version has a different chemistry scheme, it has coupled online aerosols, a different photolysis scheme, a different dry deposition scheme and different emissions. The 'p' in p-TOMCAT stands for parallel, as it was the first version of the existing TOMCAT to use Message Passing Interface (MPI) to make the model suitable for massively parallel machines. Those modifications are now incorporated into the main TOMCAT model used here. The version of the TOMCAT CTM used in this study can be considered as an updated version of TOMCAT, on which p-TOMCAT was based.
	8		***Section 2.2 Pg 18927-18928 While discussing the aircraft emissions inventories in Table 1, providing comparisons against previous global inventories as ranges is helpful. However, given the subsequent multiple comparisons of air quality and health risk estimates from this study with Barrett et al (2012), which used the Wilkerson et al (2010) inventories, having an additional column with***	As suggested we have used speciation information from Wilkerson et al., (2010) to provide information on the annual emissions of speciated HCs to allow CMIP5-extended annual emissions to be compared to annual emissions from Wilkerson et al., (2010). We have taken the suggestion to have an additional column presenting Wilkerson et al., (2010)'s numbers. Though a separate column has not been added to Table 1 presenting their data (Wilkerson et al., (2010)) specifically, global annual VOC emissions from

			Wilkerson et al numbers is suggested. Further, the EIs are listed for 6 explicit VOC species – formaldehyde, ethane, propane, methanol, acetaldehyde and acetone, and in the last column, no comparison is shown and N/A is stated. Two suggestions are offered to improve this table. Wilkerson et al report total hydrocarbons in their Table 4. The authors could compare their estimate of total HCs against that of Wilkerson, or use the speciation information of TOG in Wilkerson's Table 9, and compare explicitly for each of these 6 HCs. Since both Barrett et al (2012), and Morita et al (2014) use the Wilkerson et al inventories, providing this comparison upfront for all key species including HCs is of special relevance.	Wilkerson et al has been incorporated in to the last column of Table 1.
	9		***Section 2.5 Pg 18930 Some justification of why they chose a somewhat outdated C-R function for PM2.5 is helpful. The literature has evolved, and more recent functions including those used in the Global Burden of Disease, 2010 are available now.***	While we acknowledge that this methodology is not the most recent, we used methodology based on Ostro (2004) to allow us to be directly comparable with previous literature (Barrett et al., (2010) and Barrett et al., (2012)) that used the same CRF (please see our response to comment 1b and 4 from reviewer 1). We add a statement to the paper (pp.11, section 2.5, paragraph 6) to explain our choice and to acknowledge that newer functions are now available "We acknowledge that the CRF outlined by Ostro (2004) is not the most recent CRF available to evaluate mortality due to long term exposure to $PM_{2.5}$ (Burnett et al., 2014), this log-linear function from Ostro (2004) allows for aviation-induced mortality evaluated here to be compared against previous work (Barrett et al., 2012; Yim et al., 2015)".
	10		***Section 3.1 Pg 18931-18932 The authors acknowledge that the response of modelled inorganic PM2.5 is very non-linear and do a nice job illustrating examples where even "when aviation emissions contain no sulfur, aviation-induced sulfate is formed through aviation NOx-induced increases in OH***	This comment has brought to our attention that through the inclusion of the word "increased" within the associated explanation the wrong impression was given. We intended for that sentence to put across that even when aviation emissions contain no SO_2, sulfates are still formed through aviation-NO_X induced increases in OH concentrations. To clear this up we have removed the use of the word "increased", so as not imply there are

			concentrations, resulting in the increased oxidation of SO2 from non-aviation sources". However, this does not align with the fairly linear response of aviation-attributable PM2.5 to changes in FSC, as presented in Figure 2. A reconciliation of the non-linear response discussed above with the linear response in Figure 2 warrants additional explanation.	any changes in the rates of aviation-induced sulfates, irrespective of whether the source SO_2 emissions are from aviation or other sources. This hopefully clears any confusion about our message and helps clarify that a fairly linear response is still seen in Fig. 4 (previously Fig. 2), with our desulfurised case (FSC = 600 ppm) creating a "baseline" level of aviation-induced sulfates.
	11	a	***Section 3.2 Pg 18933 The comparison with Barrett et al (2012) can be improved here, and provide more insights to the reader on the differences being seen, especially if Barrett et al estimates are higher by factors of 5 and 2.5 in different parts of the world.***	Further analysis cannot be provided here as we are unable to compare changes surface nitrate and ammonium concentrations, as these are not shown by Barrett et al., (2012). A statement to this effect has been added (section 3.2, pp.15, paragraph 2): "Further differences in mortality assessed here in comparison to estimates from Barrett et al. (2012) can be attributed to differences in the disease specific cause-specific coefficients (β) utilised in both studies, where this study uses β values recommended by (Ostro, 2004), while Barrett et al. (2012) derive a β values for cardiopulmonary disease based on a relationship between the β values for lung cancer and the "All Cause" mortality function. In doing though the functional forms used here and by Barrett et al. (2012) are inherently the same, differences in β values which drive these functions will partly explain differences in aviation-induced mortality estimates. Additionally, different population datasets used – we use the GPWv3 population dataset while Barrett et al. (2012) use the GRUMPv1 dataset from Center for International Earth Science Information Network (CIESIN)".
		b	***In lines 20-22, when they attribute some of these differences to "other aerosol components", a quantitative comparison for each of these other components along with some explanation would be helpful.***	We acknowledge that this comparison would be helpful and aid further understanding the differences between these two pieces of work, but mean aviation-induced $PM_{2.5}$ changes from normal aviation are not reported by Barrett et al., (2012). Plots of aviation-attributable ground-level $PM_{2.5}$ concentrations for standard aviation using standard aviation fuel are presented by Barrett et al., (2012)'s supplementary information, but values are not reported in the text. A breakdown of changes in 'other' aerosol species is provided in section 3.1. when

				discussing the use of ULSJ fuel (pp.12, paragraph 2), while when discussing the impact of ULSJ on premature mortality only the total change in 'other' aerosol species mass is provided (section 3.2, pp.15, paragraph 2).
	12		***Section 3.2 Pg 18934 Lines 1-8: ULSJ reduces global mean PM2.5 concentrations by 1.41 ng/m3 and 0.89 ng/m3 in this study and Barrett et al (2012). For inorganic PM2.5 components, this study estimates 1.61 ng/m3. How does this compare with Barrett et al (2012)?***	Unfortunately, we cannot make this comparison as Barrett et al. (2012) do not report all the necessary aerosol components.
	13		***Again, when using ULSJ, if the authors see a net reduction in surface PM2.5 of 1.41 ng/m3, what is causing an increase in other aerosol species of +0.20 ng/m3? Showing the aviation-attributable speciated PM2.5 will be helpful, perhaps as global average, and again for each of the major regions studied – Europe, North America, and Asia for key ULS scenarios.***	In section 3.1 (pp.12, paragraph 2) a breakdown of the global average changes in speciated aerosol from using ULSJ fuel is provided; providing mass and relative changes for the largest changes and relative changes for the species which see the smaller changes in mass.
	14		***Lines 25-28: This study shows a 17.4% reduction in global premature mortality, while Barrett et al (2012) show a 23% reduction. The authors attribute this to larger changes in PM2.5 in populated regions of the world. Can the authors comment on potential differences in the population datasets used in the two studies?***	Both Barrett et al., (2012) and this study use population data from the Center for International Earth Science Information Network (CIESIN). We use the "Gridded Population of the World, Version Three (GPWv3)", while Barrett et al., (2012) use the Global Rural-Urban Mapping Project, Version One (GRUMPv1) version. It is acknowledged that the GRUMPv1 dataset that Barrett et al., (2012) *"which provides a higher- resolution data product that moves populations out of thinly settled large administrative units into settlements"*, thus giving a better resolution at an urban scale. The differences between the population dataset has been acknowledged in section 3.2 (pp.15, paragraph 2 and pp.16, paragraph 3)
	15	a	***Section 3.5 Pg 18937 Figure 8 presents an interesting relationship between changes in mortality versus net radiative effect for the low, mid and high ranges of mortality sensitivities for various FSC scenarios. What***	The difference in the three slopes are due to the different cause-specific coefficients used (beta coefficients) which aim to highlight the uncertainties that are present when trying to evaluate premature mortality from long-term exposure to PM2.5 – linked to comment 11 from reviewer 1.

			would explain the differing slopes for the 3 ranges?	While addressing this comment and how the choice of cause-specific coefficient affect the mortality line and its associated relationship with net radiative effect, fig 8 (now fig 10) has also been paid more attention.
		b	***While it is appreciated that the authors have performed this analysis, additional discussion here would be helpful to understand the implications of the fairly stiff response for mortality at low range versus almost linear change at high range.***	The implications of the fairly stiff responses seen at the low range in comparison of the almost linear response seen at the high range are discussed in the response to reviewer 1's comment 11a, i.e. the range created by the low and high CRFs (driven by different β-values) are employed to try and account for uncertainties which are difficult to capture in long-term studies, such as mortality displacement of a few days and disease-relevant times, durations and intensities of exposure (Ostro, 2004). The following has been added to the manuscript on (section 3.2, pp.15, paragraph 1): "The use of low-, mid- and high-range cause-specific coefficients are employed to try and account for uncertainties which are difficult to capture in long-term studies, such as mortality displacement of a few days and disease-relevant times, durations and intensities of exposure (Ostro, 2004)".

Fig. 1: Comparison of observed (Obs) and simulated (Mod) (a) sulfate; (b) nitrate; (c) ammonium, and; (d) organic aerosol mass concentrations. Observations are from airborne field campaigns compiled by Heald et al. (2011). Mean values are represented by black dots, median values as shown by horizontal lines, while boxes denote the 25th and 75th percentiles, and whiskers denote the 5th and 95th percentile values.

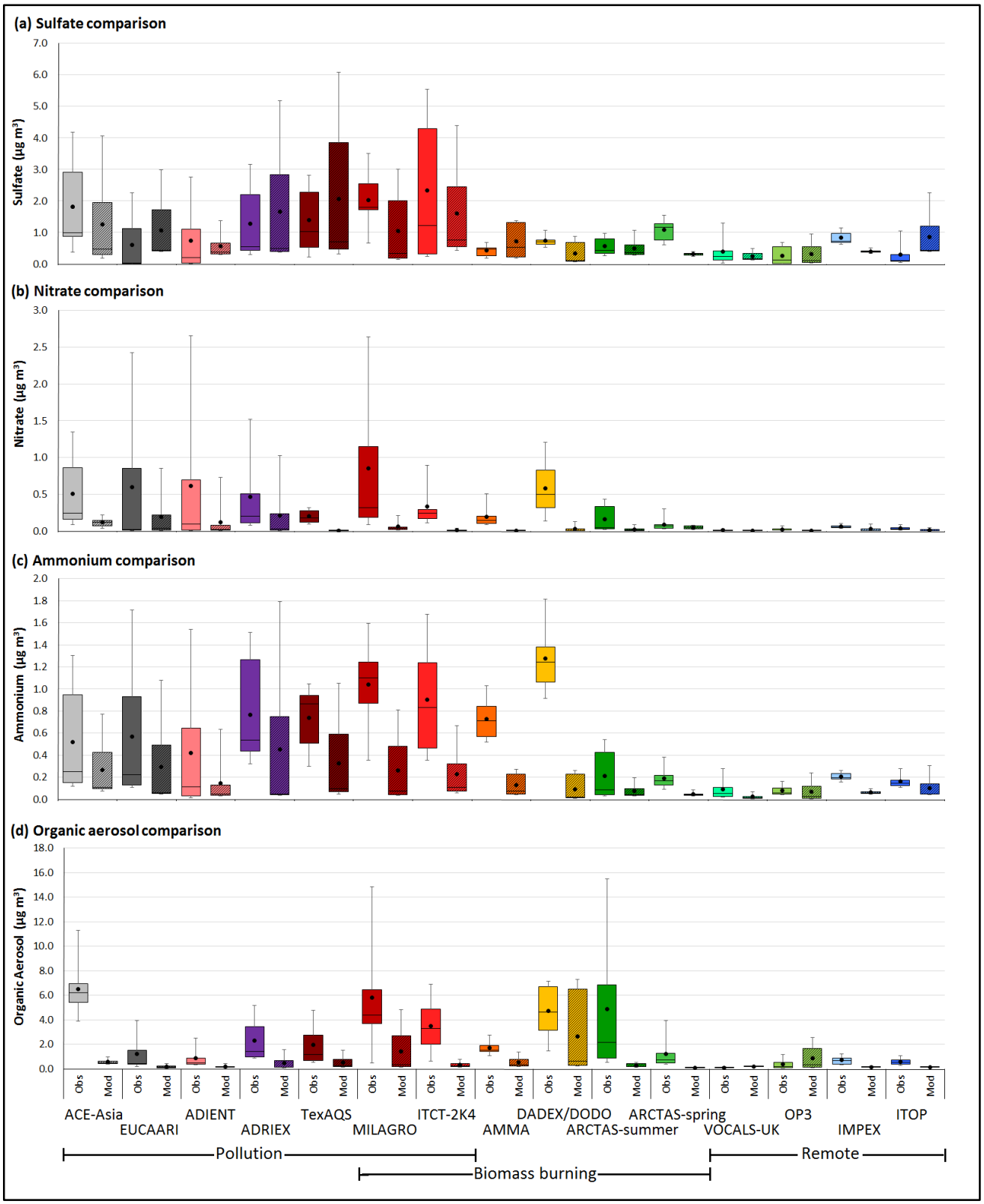

Fig. 2: Comparison of observed (solid lines) and simulated (dashed lines) ozone profiles. Observations are taken from ozonesonde observations, and arranged by launch location regions as arranged by Tilmes et al. (2012).