We are grateful to the reviewer for their interest and comments on the paper. These comments are very valuable and have helped improve the manuscript. Here we outline how we have addressed these comments in the revised manuscript. The newly added discussions and rephrased sentences have been highlighted in green in our replies below.

The paper reports and discusses the results of a modelling effort to attribute aerosol climate impacts to changes in energy consumption as well as emission control technology. The window for analyses is 1970-2010 based on the EDGAR emission inventory. In the study, the authors designed two sets of modelling simulations using the Community Earth System Model or CESM. The first set used prescribed sea surface temperature and sea ice for the purpose to derive the effective radiative effects or ERF of aerosols. The second set includes various equilibrium type of long integrations using coupled CESM with different aerosol configurations. Both methods are commonly used in climate studies. The result represents an interesting incremental progress by connecting the aerosol climate impact with sectional emissions of aerosol and aerosol precursors. The paper itself is well organized despite certain presentation issues (see later comments). Its content is adequate for the reader of ACP. There are, however, still a few issues in the manuscript to be resolved before the paper can be accepted for publication.

While revising the manuscript, we realised that the isolation of the effects of aerosol changes in the "best estimate" experiment was probably biased due somehow to the experimental set-up. To address this issue, we have carried out a new experiment, which is now used throughout the revised manuscript (including Figures and texts). Note that, while this has resulted in different estimates of the impacts of aerosol changes in the best estimate case, it does not have any bearing on the major findings of this study on the comparison of the two different retrospective emission scenarios.

1. A Clear issue in presentation, mostly appears in Section 3, is missing the term of "change" throughout discussions from aerosol burden to ERF, and beyond. This confuses the difference between two time slices with the absolute quantity of aerosol burden and radiative effects. A few examples, "BC emissions generate a global mean positive radiative forcing of +0.06 W/mˆ2" (Pg. 5, LN 29); and "The global mean ERF of sulphate aerosols…" (Pg.6, LN 1). Regarding the precipitation change, it is not clear why the insignificant changes were highlighted even by quantity in, e.g., Abstract, while the much more profound regional changes were not mentioned there at all.

We thank the reviewer for pointing these issues out. We have revised the entire manuscript accordingly.

Regarding the abstract, we feel it is important to mention the global mean changes in both temperature and precipitation. Yet, in response to the reviewer's comment, we added a sentence in the Abstract following L15 as "Despite the relatively small changes in global mean precipitation, these two emission drivers have profound impacts at regional scales, in particular over Asia and Europe".

2. An interesting while somewhat puzzling result of this study is the nonlinearity revealed in several aspects related to aerosol, from ERF to model equilibrium sensitivity to aerosol ERF. The reason behind the fact that aerosol-species-based ERFs do not add up might have something to do with the (uniformly) internal mixing nature of the aerosol model where the hygroscopicity of aerosol is largely decided by organic carbon content due to its dominance in volume (ERF is largely a reflection of aerosol-cloud interaction or indirect effect of aerosols). Additional discussions are needed. Regarding the model's equilibrium sensitivity, could the different integration times in various simulations be at least a part of the reason responsible for the "nonlinearity"? Note that although the TOA forcing residual might be minimized throughout the quasi-equilibrium stage, ocean status such as SST evolution might still differ from time to time. Note also that aerosol forcing is rather small comparing to many internal factors of the model. Therefore, comparing simulations at different stages could likely introduce an arbitrary discrepancy in derived mean values. The authors should experiment using the same time slice for equilibrium analysis.

We agree with the reviewer that the nonlinearity in ERFs may be related to the internal mixing state of the different aerosol species. We also agree with the reviewer regarding the potential role of different oceanic states on the inferred equilibrium of the various experiments.

In response to the comment, we expand P9 Ls3-7 into "This reflects partly the nonlinear effect associated with the mixing state of different aerosol species as well as the importance of background aerosol loadings. This is particularly important for BC whose effects depend also on the presence of sulphate and organic aerosols (Ramana et al., 2010). That is, given that aerosol species are internally-mixed in MAM3 (i.e. different chemical species are mixed within an aerosol particle), the hygroscopicity of aerosol particles is dominated by the volume of soluble species (organic compounds and sulphate). This means that the nonlinearity in the isolated aerosol ERF may be a reflection of the aerosol scheme in CESM1. More specifically, BC particles tend to be coated with other species during ageing, thereby enhancing the absorption effects and the subsequent impacts on cloud microphysics, as well as amplifying their radiative forcing (Haywood and Ramaswamy, 1998; Kim et al., 2008; Chung et al., 2012; Wu et al., 2016). …".

Regarding the use of different simulation periods in the analysis, we feel necessary to provide further motivations for this choice. As stated at P5 Ls 7-10, each fully-coupled model simulation was considered in equilibrium when the TOA net radiation was devoid of any significant trends (less than 5% relative to the mean values stabilizing at ~0.3 W $m^{-2}$) during a 30-50 period following previous works (*Samset et al.*, 2016; *Samset et al.*, 2018). This period was then used in the analysis. Note that our simulations are longer (~150 vs. 100 years) and with a smaller TOA radiation imbalance (~0.3 vs. above 0.5 W $m^{-2}$) compared to those analysed

in *Samset et al.* (2018), meaning that our integrations are even closer to equilibrium. While it is certainly possible to analyse the same time period in all experiments, we argue that this would lead to examining potentially different equilibrium states (as defined above), and thus to potential misinterpretations of the differences among them. Therefore, we believe it is more appropriate to compare the same equilibrium states rather than the same time period. In fact, this reflects a very important issue regarding the time scale of different forcing agents in climate models, a topic which deserves new and deeper analysis in the future.

Minor comments.

3. Pg. 2: LN 20, "forcing type", could be elaborated. LN 24, "ongoing debates" on what topic? LN 28-29, "developing region", perhaps "developing countries" is better.

In L20, by "Forcing type" we mean different aerosol species (i.e., BC, $SO_4$ and organic compounds). To make this clearer, we reworded "forcing type" into "aerosol species".

In L24, the "ongoing debates" refers to the sentences in L25-27. Namely, ongoing debates on whether aerosol forcing has larger impacts on mean climate and climate extremes compared to GHG. We have now expanded the sentence "despite ongoing debates" into "despite ongoing debates as to whether aerosol forcing has larger impacts on mean climate and climate extremes compared to GHG.".

"Developing region" has been reworded into "developing countries".

4. Pg. 4: Section 2.1, the relative changes of BC, OC, and sulphate from 1970 to 2010 referring to either 1970 or 2010 alongside relative ERF change should be provided. LN 23, "their number concentration", please note that for a modal aerosol model, the number concentration is defined for each mode rather than different aerosol compositions. LN 26, "to be coated", this could only be an assumption in performing certain calculation (e.g., optics) and is actually not necessarily consistent with the model's configuration.

We thank the reviewer for all the suggestions.

Since the 1970-2010 changes in BC, OC and $SO_2$ emissions have been thoroughly presented in a published paper by one of our co-authors (*Crippa et al.* (2016)), we have added Figure S1 in the supplement to show emission changes, and decided to direct the readers to *Crippa et al.* (2016) by for more details. As such, we added the sentence " For the 1970-2010 changes in emissions of each individual aerosol/precursor species, please refer to Figure S1 in the supplementary file and *Crippa et al.* (2016). " Following P4 L7,

P4 L 23: "Several aerosol species (sulphate, organic carbon (OC), black carbon (BC), sea-salt, and dust) are simulated and their number concentrations and mass are prognostically calculated" has been slightly rephrased into "Several aerosol species (sulphate, organic carbon

(OC), black carbon (BC), sea-salt, and dust) are simulated, and their number concentration and mass are prognostically calculated for each aerosol mode".

To avoid confusion, we deleted the statement on BC at P4 L 25-26.

5. Pg. 5, LN 16, the configuration of paired Fsst simulations should be listed in either Table 1 or a separate table.

We thank the reviewer for the suggestions. We have updated Table 1 to include the Fsst runs. Also note as mentioned above, we have modified the table to account for the new experiment used to isolate the aerosol effects for the best-estimate case scenario.

**Table 1** Overview of the fully-coupled (Fcpd) and the paired stimulation (Fsst) where sea surface temperature and sea ice are fixed. They are: the baseline 2010 (B10) simulation, fixing aerosol-related emissions in 1970 levels (SAA), stagnation of anthropogenic aerosol-related emissions from energy use in 1970 levels (SEN), and stagnation of aerosol-related emissions related to technology and abatement measures in 1970 levels (STC). All Fcpd simulations are run into equilibrium (numbers in brackets denote the lengths of model integrations in years), while all Fsst runs are integrated for 40 years. Only the last 30 years of each Fcpd and Fsst run are used for analysis. Note the difference in the integration lengths of Fcpd simulations, which is determined on the criterion that the top-of-the-atmosphere radiation imbalance no longer shows significant trends (stabilizing at around ~0.3 W $m^{-2}$ in this case) during the last few decades of each run (see the main text). The response to the best estimate of 1970-2010 anthropogenic aerosol-related emissions: best estimate = B10-SAA. Similarly, energy use growth = B10–SEN; technology advances = B10–STC.

| **Experiment (length of Fcpd/Fsst)** | **Greenhouse gases** | **Ozone** | **Natural aerosols** | **Anthropogenic aerosols** |
|---|---|---|---|---|
| B10 (150/40) | 2010 | 2010 | 2010 | 2010 best estimate |
| SAA (120/40) | 2010 | 2010 | 2010 | 1970 best estimate |
| SEN (220/40) | 2010 | 2010 | 2010 | 2010 STAG_ENE |
| STC (170/40) | 2010 | 2010 | 2010 | 2010 STAG_TECH |

6. Pg. 6, Section 3.2, according to Figure 4, it seems that majority of statistically significant changes in temperature appear over oceans rather than land.

The reviewer is correct in pointing out that the majority of statistically significant temperature changes are over the oceans rather than land. This is especially the case for the energy use experiment, and may be explained by the fact that the equilibrium climate responses are mainly related to the ocean.

We add a sentence to comment on this following P6 L20 as "It can be seen that the majority of statistically significant temperature changes in response to aerosol changes are over the ocean rather than the land. This is particularly true for the energy use experiment, and may reflect the fact that the equilibrium climate response is dominated by the slow response of the ocean."

7. Pg. 9, LN 2, "the residual (0.14 W/mˆ2)", could the authors elaborate on how to derive this residual?

First of all, we apologise that the number on the top right of Figure 2d was incorrectly reported as 0.10 W $m^{-2}$. The correct value is -0.11 W $m^{-2}$ (see Figure 3b). We feel this may have confused the reviewer, and have corrected this in the revised version.

The ERF due to changes in BC (0.06 W $m^{-2}$), OC (-0.04 w $m^{-2}$), and $SO_4$ (0.01 W $m^{-2}$) add up to 0.03 W $m^{-2}$. The difference between the sum of individual aerosol ERFs in Figure 1 (0.03 W $m^{-2}$) and the total in Figure 2b (-0.11 W $m^{-2}$) therefore produces a residual of -0.14 W $m^{-2}$.

8. Pg. 10, LN 1 temperature responses do not necessarily follow the ERF...", why? The scale of temperature response to ERF or the "equilibrium sensitivity" could differ from case to case, but for the same forcing agent in the same model, it should be the same, in other words, the temperature response should be always proportional to (or follow) ERF.

We understand the reviewer's point. However, as also demonstrated by other recent works (Persad and Caldeira, 2018; Lewinschal et al., 2019), the link between ERF and temperature response may hold particularly for long-lived and spatially homogeneous forcing factors, such as greenhouse gases, but may be weaker for short-lived aerosols which are, by nature, highly heterogeneous and involve complex atmospheric circulation adjustments (*Shindell and Faluvegi*, 2009; *Shindell et al.*, 2010).

9. Pg. 10, LN 18-22, "...it is also likely that aerosol emissions will increase...", this discussion actually raises an interesting issue that recent increase of aerosol emissions could occur not only in developing but also developed countries. Note that the EDGAR estimate used here is up to 2010, a year before Fukushima Daiichi nuclear disaster in March 2011. Due to the closure of nuclear facilities from Europe to East Asia following that event, it is likely that in recent years coal burning has already come back in many of these regions because of obvious shortage in energy supply otherwise from renewables alone.

We thank the reviewer for providing very interesting insights.

We expand L18-23 into "On the other hand, it is also likely that aerosol emissions will increase, especially over some developing regions, under scenarios where high inequality exists between and within countries. For example, in SSP3, expanding industrial sectors over Southeast Asia may continue to rely on fossil energy sources such as coal. Also, it is possible that the world may continue to rely on fossil energy sources more strongly than expected over the coming years, given the concerns about nuclear energy after the Fukushima Daiichi nuclear disaster in March 2011. As a consequence, aerosol emissions from energy use in some regions may increase and therefore offset aerosol reductions elsewhere.".

**References**

Crippa, M., G. Janssens-Maenhout, F. Dentener, D. Guizzardi, K. Sindelarova, M. Muntean, R. Van Dingenen, and C. Granier (2016), Forty years of improvements in European air quality:

regional policy-industry interactions with global impacts, Atmospheric Chemistry and Physics, 16(6), 3825-3841.
Samset, B., M. Sand, C. Smith, S. Bauer, P. Forster, J. Fuglestvedt, S. Osprey, and C. F. Schleussner (2018), Climate impacts from a removal of anthropogenic aerosol emissions, Geophysical Research Letters, 45(2), 1020-1029.
Samset, B., G. Myhre, P. Forster, Ø. Hodnebrog, T. Andrews, G. Faluvegi, D. Flaeschner, M. Kasoar, V. Kharin, and A. Kirkevåg (2016), Fast and slow precipitation responses to individual climate forcers: A PDRMIP multimodel study, Geophysical Research Letters, 43(6), 2782-2791.
Shindell, D., and G. Faluvegi (2009), Climate response to regional radiative forcing during the twentieth century, Nature Geoscience, 2(4), 294.
Shindell, D., M. Schulz, Y. Ming, T. Takemura, G. Faluvegi, and V. Ramaswamy (2010), Spatial scales of climate response to inhomogeneous radiative forcing, Journal of Geophysical Research: Atmospheres, 115(D19).