We thank the reviewer for constructive and helpful input and provide our response in detail below.

Response to Reviewer Comments

We thank the reviewers for their careful comments, which improved the quality of the manuscript. Below, the reviewer's comments are repeated in the italic text. Our response follows in normal letters. Blue text is used if text from the revised manuscript is cited. When page and line numbers are specified, they refer to the clean version of the revised manuscript. Additionally, we highlighted the changes in the manuscript by track-change mode, and attached them to the reply.

REFEREE 1 (RC1)

In this study, the authors calibrate a previous version of the algorithmic climate change functions (aCCFs), obtained by fitting climate model simulations to a small selection of meteorological parameters, to the impulse response model AirClim. Both models aim at estimating the average temperature response to CO2 and non-CO2 emissions of flights, but follow different approaches and assumptions. The calibration produces an alternative set of aCCFs, with sizeable differences in the estimated average temperature response of flights compared to the original aCCF set.

The study is well written, and the figures and tables illustrate the discussion well. My comments below aim at making the description clearer and more accurate in places and should amount to minor revisions because new analyses are not required.

Main comment:

Section 3 and 5 need to be more upfront on two important points. First, it is important to point out in section 3 that the "calibration factors" (Table 2) that translate aCCFv1 into functions comparable to AirClim imply very large changes. Those are not small corrections. So the two models, aCCFv1 and AirClim represent two very different views of the climate impact of flights, for the reasons listed (only qualitatively unfortunately) in the conclusion. Second, Section 5 does not answer the basic question of why one would want to calibrate aCCFv1 to AirClim. As a measure of uncertainty? Not really because the two model are unlikely to cover the whole uncertainty range. As a way to choose between different philosophies (tagging/ perturbation and climatological/synoptic)? But what would be the rationale for such a choice?

- We have reconsidered using the term calibration, and we suggest to replace the term "calibration" with "alignment" in the whole manuscript.
- We are scaling absolute estimates of individual non-CO_2 effects (P-ATR20) estimated from aCCFs in a way that they become identical to absolute estimates from AirClim. First, the alignment factors show, which of the two model results in larger P-ATR20 values of the individual effects, exploring absolute estimate. Second, relative strength of individual non-CO_2 effects to the overall climate effect gets aligned to the relative strength of AirClim climate estimates, while maintaining with this alignment of aCCF, identical spatial and temporal 'relative distribution' of each individual non-CO_2-aCCFs-effect.
- [Lines 19-25] Individual **alignment** factors (i.e. AirClim **alignment**-factors) are provided for the respective non-CO_2 effects comprising contrail cirrus, water vapour and NO_x-induced climate effects on ozone and methane. **Aligned aCCF-V1.0A can be seen as one realization within the range of possible values.**

Other comments:

Line 12: "climate effects" Be specific: the aCCFs give the Average Temperature Response.

- We added the physical climate metric used, by way of example (see track changes in the abstract), as aCCFs can be used to estimate climate effects providing various climate physical climate metrics, e.g. GWP, GTP, ATR.

Lines 19-20: Clarify: "lower estimates" of the ATR?

- Thank you for that comment. We have updated the text accordingly, explaining that the aligned aCCFs are aligned in strength and relative strength of individual non-CO_2 effects of AirClim V2.0 results (see track changes of revised document).
- **[Lines 21-23] The aCCFV1.0A provide** ~~resulting~~ **a scaled version of** ~~uniformly in lower estimates in~~ aCCF-V1.**0**A **for climate effects measured as ATR** of all **non-CO_2** species **which are aligned for a European air traffic sample to the absolute and relative strength as estimated by AirClim V2.0**. ~~compared to the earlier version aCCF-V1.0.10.~~

Line 49: Not sure that "short-cut" is the right word, because there is also a loss of accuracy. "Approximation" would be a better word.

- Yes, you're right: Approximation describes this concept well We have updated the text accordingly (see track changes)
- [Line 52-53] Such response surfaces can be seen as **a direct** link and approximation ~~short-cut~~ between aviation emissions and associated climate effects.

Lines 55 and 62: Need to define what is meant by "perturbation approach" and "tagging contribution approach". It might be useful to add a paragraph before this one to clearly define those two approaches to attribution, and what they mean in terms of radiative forcing. As you found in Grewe et al. (2019) https://doi.org/10.1088/1748-9326/ab5dd7, the impact of the choice of methods is very large: a factor 6.7 according to Table 1 in that paper.

- We agree that it is useful to the reader to briefly introduce and discuss both approaches. However, we do not agree with the estimated factor of 6.7 when referring to Grewe et al. (2019), By clearly distinguishing the separate updates, table 1 indicates a factor of 1.6 for the climate effect of the ozone signal, and a factor of 2,1 for the combined net NOx-effect. We have added a section in the discussion section exploring the different approaches We included the following text to the revised document:
  [Line 486-473] Overall, ~~we would like to stress~~ when comparing **the resulting estimates of climate effects from** AirClim **alignment** ~~calibration~~ factors (Table 2), which align ~~calibrate~~ non-CO_2 effect estimates from EMAC/AirTraf/ACCF to the values quantified by AirClim, **we find t**hat **result lie** within the confidence intervals for effective radiative forcing of individual climate effects (when comparing relative variation) which agree with those suggested by assessment studies, e.g., Lee et al., 2021.
  [Line 474-477] Results show that no constant factor is identified, but variations in the order of 20% can be identified. Hence, we consider that changing from perturbation to tagging approach introduces an additional uncertainty in the order of 20%, but does not affect the overall validity of the aCCF concept. With the proposed additional sensitivity study (repeating the workflow for other flight levels), we derive separate set of alignment factors, while exploring vertical dependence of such differences.

Line 61: I do not understand "for eight specific days". Do you mean eight specific weather situations?

- Initial baseline simulations that were applied for development of aCCFs covered eight archetypical weather situations; with each being represented by one typical day, hence overall eight individual days. We have further clarified that in the text:
  [Line 67] …for the North Atlantic Flight Corridor (NAFC) for eight specific days **and their synoptic weather situations** that consider representative weather types in summer and winter (Frömming et al., 2021 […].
  [Lines 70-72] The overall concept of aCCFs development relies on statistical methods which link and correlate non-CO_2 climate effects as quantified by the **set of** CCFs. Base simulations were performed on explicitly calculated in comprehensive numerical chemistry climate model simulations) [..]

Line 68: "to local meteorological conditions" is too broad. It is in fact "to a selected set of variables that represent the local meteorological conditions".

- It is in fact "to a selected set of variables that represent the local physical state of the atmosphere and meteorological conditions".
  [Lines 72-75] to a selected set of variables that represent the physical state of the atmosphere and local meteorological conditions at location of emission. Therefore, the strength of **the prototype** aCCF is that their implementation solely requires the local state of the atmosphere as input data**, however this statistical method also introduces uncertainties in the estimated values and represents a simplification.**

Line 77: Need to clarify what is meant by "effectiveness" here.

- The term effectiveness here refers to the application of NOx-aCCFs in order to guide flights in such a way, that they avoid regions where high NOx-aCCFs gives large estimates, in order to generate an emission inventory which has an overall lower NOx-induced climate effect. We have added more explanation here for clarity.
- [Line 85-89] For an initial proof of concept **to estimate achievable mitigation gains**, Yin et al. (2018) and Rao et al. (2022) applied an atmospheric chemistry model chain within the EMAC framework to evaluate the effectiveness of NOx-O3 aCCFs using their climate effect information for trajectories optimizations, **hence evaluating if flight guidance which avoids regions with large climate effects as described by aCCF, result in lower overall NOx-induced climate effects.**

Line 83: The meaning of "the room for a calibration process is open" is only clarified in the next paragraph (line 88) by "one realization within the range of plausible values", so I would suggest merging the two paragraphs to clarify what you are doing. And it should be noted that such calibration does not replace the need for a way to properly account for uncertainty in both AirClim and aCCFs.

- We agree that we need to describe more in detail, what our study is doing. Hence, we have updated this paragraph accordingly. Additionally, we now suggest to use the term alignment for the factors that we have identified when comparing aCCFs estimates with AirClim estimates.
  The aCCF-V1.**0A** can be seen as **another** ~~one~~ realization within the range of plausible values (event horizon) considering the current level of scientific understanding of climate effects of aviation emissions and their associated uncertainty. The alignment process was performed ~~is was done~~ by evaluating and comparing the climate effects of aviation for a European ~~arbitrary~~ air traffic inventory ~~scenarios~~, resulting from aCCF-V1.0 **(Yin et al. 2023),** with an alternative numerical model; here we used the **advanced** state-of-the-art climate-response model AirClim. A detailed description, visualization, and application of aCCF-V1.0 are provided.

Line 109: "it is equivalent to selecting another value from this interval". Is that true? If there's a probability distribution, then some values are more probable than others, although I agree that the most probable value is not necessarily the mean.

- All values within a confidence interval are possible, however, it is correct, that a probability distribution provides additional information if some values are more probable than others, and the mean value is not necessarily the most probable one. We rephrased the text to be more comprehensive.

[Line 123-129] The mathematical formulation given in aCCF-V1.0 (Yin et al. 202**3**) provides a numerical value which represents ~~the statistical mean value of the probability distribution per~~ individual CO2 and non-CO2 climate effects~~; hence it relates to a confidence interval of possible values. From a statistical point of view, each value in this interval is possible. Instead of choosing the mean value to calculate the climate effect, it is equivalent to selecting another value from this interval~~. Consequently, for our **alignment** ~~calibration~~ process, we explore ~~in this study the underlying range of uncertainties~~, and align**ment factors which would bring** the quantitative estimates **from** ~~of~~ aCCFs to **estimates from an** ~~the~~ alternative modelling approach, i.e. ~~studies with~~ the climate-response model AirClim.

Lines 126-127, lines 150-151, line 154: Again, avoid using the imprecise "climate effect". Here, I assume that it is the ATR that is calculated.

- We intentionally use the term climate effect as we see an equivalence between different physical climate metrics, which can be transferred from one into the other. We agree, that it is helpful to specify, that our study provide estimates in order to scale values that relies on average temperature response (ATR).

Line 140: "as boundary conditions" – ambiguous. Do you mean that there is no nudging inside the model domain? By the way, is the model configured to simulate Europe only, or is it global?

- Yes, it is misleading to call our nudging field boundary conditions, hence we have deleted this explanation. We have 4-dimensional fields of atmospheric state variables, and we relax numerically EMAC model conditions towards these values. Thus, we deleted "boundary conditions" (see track changes).

Line 149: The aCCFs are made out of the CCFs for 8 weather patterns. If I understand well that synoptic information is lost in the aCCFs, so it is not possible to relate aCCFs to specific patterns in the North Atlantic. Is that correct? It would be good to clarify that here. And by the way, are the aCCFs fitted over the 8 weather patterns together? Or are they fitted over some mean of the 8 weather patterns?

- Yes, aCCFs do not relate to one individual synoptic pattern. Moreover, during their construction of the base simulations, particular attention was paid to cover the set of possible weather patterns by identifying eight archetypical - different - synoptic weather patterns, in order to span the range of possible conditions in the North Atlantic Flight Corridor. For the parameterizations, data from all eight weather-pattern were combined in order to identify a fitting algorithm, which can be applied independently from the prevailing weather situation. We mention the set of CCFs which are used to generate one formulation of aCCFs. This has been clarified in the introduction and further clarified in this methods sections.
  [Line 158-160] Specifically, they allow identifying regions of the atmosphere where aviation emissions induce a strong climate effect, e.g. via the formation of warming contrails or the production of radiatively active species like ozone **independently from the prevailing weather situation.**

Line 173: Why A330 flights only?

- For the construction of the emission inventory, we assume that the whole European sample of city pairs is flown by A330 aircraft, as within the EMAC submodel AirTraf corresponding aircraft performance data is integrated for this aircraft.

Lines 182-186: Again, no need to keep the suspense on the actual calculation by using the imprecise "climate effect". Say that you are calculating P-ATR20 from the start of the paragraph.

- We agree, that it is helpful to mention the used physical climate metric, P-ATR20. We added some explanation in the revised manuscript now.
  [Line 208-212] While ATR20 or ATR combined with any other time horizon is not addressed in any political framework, it shows advantages for representing aviation related non-CO2 effects (Megill et al. 2024). We note that the short time horizon chosen here, puts more weight on the short-term effects. However, the chosen metric may be customized by constant, species dependent factors to other metrics and other time horizon for both pulse or continuous future emission scenarios similar as in Yin et al. (2023, Tab. 1).

Caption of Table 1: The "without forcing efficacy" comes as a surprise. There should be an explanation of what that means in the main text.

- We agree, and we give a very short explanation, but refer to relevant literature in the text now.
- [Line 234] (provided in P-ATR20 in [K] and **including the forcing dependent efficacy values (e.g.; Ponater et al. 2006; Lee et al. 2021)**)

Line 227: Converting to F-ATR20 when Section 3 has been all about P-ATR20 seems an unnecessarily confusing step. Why do it? And here again, the reader would benefit from a reminder of the notion of efficacy, and how it might change the results compared to Section 3.

- Thanks for pointing that out. In the revised document we now apply aCCFs with P-ATR20 including efficacy for a day in winter (all figures are changed accordingly), with that we are in line with above alignment chain and moreover contrail aCCFs - which were developed only for typical winter weather patterns (see supplement of Yin et al. 2023) – are applied in the designed area and season.

Line 245: I would expect variability in the contrail aCCF, but I would also expect some correlation between time steps, since ISSRs are not advected randomly. But then the aCCFs do not have the concept of ISSR, so what do the patterns shown in Figure 2c tell us? It is also surprising to have pockets of cooling in the middle of warming zones – that does not look like the day/night contrast alluded to in the text.

- Firstly, please note that Fig. 2 in the revised document now displays aCCF patterns for a specific winter condition, as contrails aCCFs were developed for winter weather patterns only.
- aCCFs do integrate the concept of ISSR: we calculate ISSR region every time step using a for ERA5 derived threshold for the relative humidity over ice and a temperature criterion. If ambient air is supersaturated we apply aCCF formulas. We added one sentence in the description of the contrail aCCFs, see Appendix A.
  [Line 920] Note moreover, that contrail aCCFs are only calculated at location and time where persistent contrails are forming, and accordingly regions without persistent contrails are set to zero.

- The intension of Fig. 2c is to shows how ISSR (and associated climate estimate) are varying. Ice supersaturation is a temporarily highly variable field that induces changes within 6h patterns (here pressure level 250 hPa). In some cases, there are ISSR patterns that are quite stationary (see Fig. 2c, 2018-12-02, 0, 6, 12 UTC), here we assume that uplifting of airmasses and connected adiabatic cooling in a stationary system (e.g. connected to a front) induces these ISSR.
- No, this is not day/night contrast, the day/night contrast only can be seen for sunrise and sunset. The current prototype contrail aCCF formulation is a function of the outgoing longwave radiation (see Annex A). Thus, according to the equation, the RF for the daytime contrails can take positive and negative values, depending on the OLR. As the OLR mainly depends on surface temperature and cloud cover, aCCF formulas produce a sharp change in the contrail climate effect estimate in regions where conditions change from clear-sky to cloudy-sky. This explains these pockets of cooling in the middle of warming ones, although we are not sure if these estimates (with strong gradients) are realistic. As mentioned aCCFs are prototypes and we like to stress the limitation mainly of the the day-contrail formulation. In order to address this issue, we added text in the discussion section of the revised manuscript.
- [Line 481] E.g., Yin et al. (2023) discussed several open issues to be addressed further, e.g., the temporal and spatial applicability of the current prototype aCCFs and the uncertainties related to contrail aCCFs. We would like to stress a specific limitation when estimating radiative effects with the help of any parametric response function which relies on numerical weather forecasts. The current algorithm relies on OLR as the basic variable for the estimation of net contrail effects, however no exact information on the existence of clouds in the vertical column is available, e.g. at what altitudes, which is expected to have a non-neglectable influence on the potential climate effect of contrails. However, these effects are in general neglected when providing quantitative estimates of the contrail climate effect response functions.

Line 307: What is the implication of that narrower distribution for climate-optimised routing?

- The narrower distribution of the merged non-CO_2 aCCF indicates that NO_x effects and contrail effects are getting closer, thus the tradeoff of contrail and NO_x effects will change. Moreover, also the trade-off with CO_2 will change significantly (not shown here), as non-CO_2 effects are scaled down for aCCF-V1.0A (see Table 2).
  [Line 345] Additionally, the distribution of the merged non-CO_2 aCCF is narrower for aCCF-V1.0A, indicating that NO_x effects and contrail effects are getting closer and thus leading to a slightly different trade-off between contrail and NO_x effects. Moreover, also the trade-off with CO_2 will change significantly (not shown here), as all non-CO_2 effects (unlike for CO_2) are scaled down for aCCF-V1.0A (see Table 2).

Lines 366-367: What do those confidence intervals look like? How do you go from your calibration factors to confidence intervals?

- Thanks for pointing that out. We know explain more about it in the discussion section. See subsection 5.3: "Uncertainties in prototype aCCFs and in alignment factors" of the revised document.

Technical comments:

Line 26: Suggest rewriting to "long-term background ozone is reduced by the NOx-induced methane decrease".

- Done

Line 43: Presenting the three types of studies as bullet point would make that paragraph easier to read.

- Even though, this is a valid point, we do not like to build points in the introduction, thus we did not follow your advice here.

Line 88: I have never seen "event horizon" used in this context. Is that a correct use of the term?

- This term has been used in scenario analysis, in order to describe the "limits" of the possible results and realization. This might not be a familiar expression; hence, we suggest to delete it. [Line 99] *Term "event horizon" has been deleted.*

Line 115: Typo aligned

- Done

Line 117: "is corresponding" -> corresponds

- Done

Figure 1: Typo relevant (and it would be good to disable the spelling checker to avoid those words underlined in red)

- Done

Line 164: green function -> Green's function

- Done