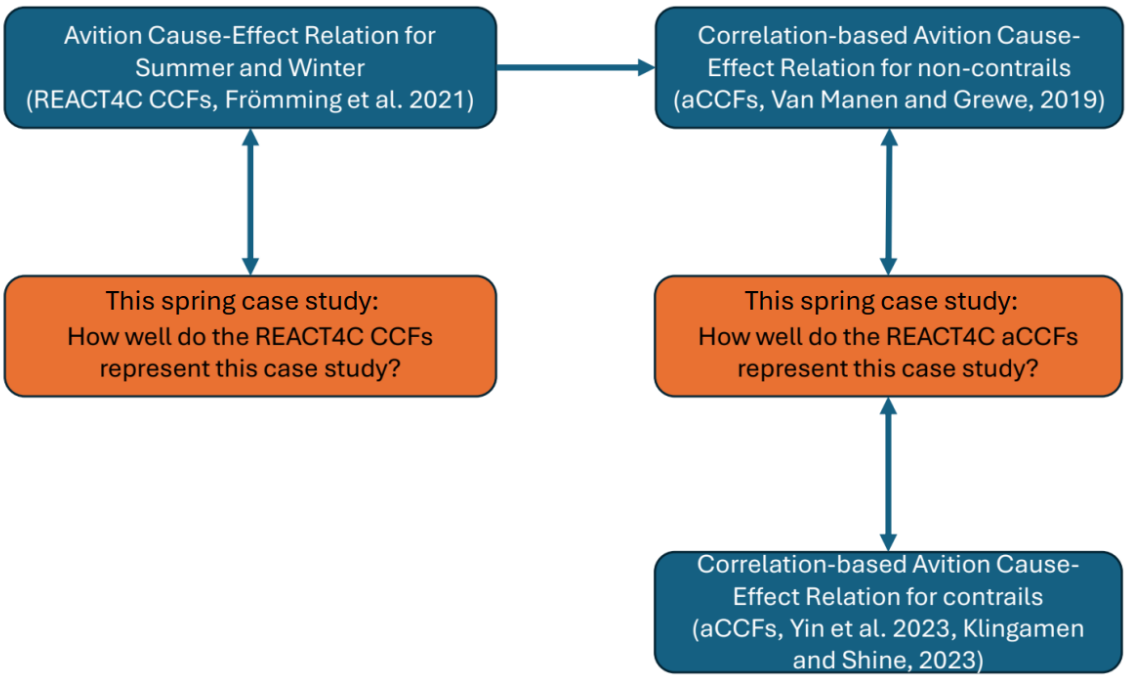
R1:

General comments

The authors present a limited evaluation of algorithmic climate change functions (aCCFs) as a method to evaluate the climate impact of specific flights, using revised calculations of climate change functions (CCFs) based on calculations with the ECHAM/MESSy EMAC climate model. The evaluation looks specifically at two days in March 2014, comparing revised CCFs in a limited spatial domain at low resolution (2.8x2.8 degrees) to algorithmic CCFs (aCCFs) derived from the original CCFs. The title promises "[e]xpanding spatial and temporal coverage" with "[a]ssessment and comparison with aCCFs", two different goals: with regards to the first, the experimental setup makes this rather a case study as it is not clear how to combine the new results with the original CCFs. On the second goal, the paper does not provide any external evaluation of (a)CCFs but rather just compares one set of CCFs to another. More concerningly, seemingly strongly negative results are presented positively in the abstract in a way which seems very misleading. I do not think the conclusions drawn are supported by the study's results, which instead show what appears to be a general failure of existing CCFs and aCCFs to represent variability in results from the same model.

- *Thank you for this comment. Indeed, the CCFs presented in this study are not combined with the original CCFs to expand them. We reformulated the title and parts of the abstract and text accordingly to avoid misinterpretations.*
- *An external evaluation of (a)CCFs is beyond the scope of this study. Furthermore, such an evaluation depends on the availability of weather-dependent, episodical impact studies from other models. Apart from contrails we are not aware of such studies. Nevertheless, we elaborate more on altitudinal or latitudinal climatological studies to assess our current estimates. Regarding contrail aCCFs a detailed comparison study of contrail climate responses from aCCFs with CoCiP has been performed within the EU project CICONIA, the manuscript of this comparison will be submitted shortly (Dietmüller et al., 2026).*
- *We reformulated the abstract to neutrally summarize the results and point out agreement and disagreement.*
- *We outline more clearly the two objectives we are following in the manuscript by using the spring case study CCFs:*
    1. *How well do statistics for emission-effect relations - expressed by CCFs of this case study (orange left box) - agree with previously calculated CCFs that cover winter and summer (blue left box), only.*
    2. *How well do aCCFs that were developed on the basis of the REACT4C CCFs (right blue box) represent the spring case study (orange right box)*



Objectives of this study (orange) in relation to previous work on CCFs (blue left box) and aCCFs (blue right box). At the end of the introduction, we clarified the outline and objectives of the manuscript more clearly. The two objectives are directly related to Sections 5 and 6 in the manuscript.

Despite this, I do feel that the work contains some excellent analysis and could potentially provide a useful case study of shortcomings in the general concept of CCFs and aCCFs. Ideally it could present solutions to some of these shortcomings, and provide a set of clear metrics for how future CCFs might be evaluated to show that they are beginning to accurately represent the original model (or ideally to show that they represent broad model consensus). However, the manuscript currently seems to be split between multiple conflicting objectives, seeking on the one hand to extend and support (a)CCF usage while also performing a case study (presented as an extension) which yields overall unfavorable results. As such I cannot recommend the manuscript for publication without either substantially reframing the work to make it clear what the results imply about the (in)accuracy of CCFs and aCCFs, or performing a substantial expansion of scope to address these errors and develop the promised expanded-scale CCFs.

- *Thank you for your comment. Indeed, multiple objectives are addressed in this manuscript. It reveals shortcomings of aCCFs beyond their original domain and season which limit their applicability, and points towards necessary developments for broader usage of aCCFs. We do not see the results "overall unfavorable", particularly in view of considerable uncertainties of non-CO_2 climate impact in general (Lee et al., 2021), but we do agree, that the results need to be reframed and presented neutrally to avoid misinterpretation. We reformulated the title, large parts of the abstract and text accordingly, to give an impartial view on the performance and point out both more clearly the limitations of aCCFs. We also indicate our objectives more clearly and the relation to the sections where those objectives are addressed.*
- *We largely revised the discussion section to address the reviewers' comments. We added a brief summary of the findings, added a section on the CCFs and aCCFs reliability, what further steps should be considered and how and where aCCFs can be applied.*

Specific comments

The work is not an extension, but rather a case study: if the exact same model had been used as for the original CCF work, then the argument that this study presents an extension (albeit a very limited one) would be reasonable. Alternatively, the study could have recomputed CCFs for the times and locations of the original study. Instead, we have to consider that the original CCFs are for 36 lat/lon pairs and 4 altitudes, computed using one model, over only the North Atlantic and for 8 different days which are meant to represent specific weather patterns. In this study we have CCFs for 56 lat/lon pairs (none of them matching the original pairs) and 6 altitudes (three of the four from the original study and three news ones) for two seemingly arbitrary days (specifically, three times in one day and two times in the other – none of which match those in the original study – on 2014, 13-14 years later than the original), now simulated using a different model version, at the same coarse resolution, and with nudging rather than free running (with other changes beyond these). This makes it extremely challenging to consider how the two could be used together, as the degree of (dis)agreement between the new and old approaches cannot be directly evaluated. Instead, extrapolation of the old CCFs to compare against the new - as indeed is done here – shows that the sets are apparently incompatible. I would recommend that, if the authors wish to "expand" the CCFs, they instead recompute the CCFs with their updated climate model and perturbation structure, and perform a structured evaluation which covers all the target time points (i.e. including the original weather patterns), not just these two days.

- *The way we used the term "extension/extended" obviously led to some misunderstanding that was unintended and needs to be explained more clearly. We meant that the spatial domain of previously available CCF calculations was expanded in the spring case study. Our intention was not to imply, that we include these CCF results into a new expanded version of aCCFs. We reformulated the title and text accordingly to avoid misunderstanding.*
- *At the time, the study started, the actual version of EMAC (ECHAM/MESSy Atmospheric Chemistry) was employed. While the base atmospheric model ECHAM5 remained as in the original CCF studies, the version of MESSy, which couples submodels to the atmospheric model was updated. The updates are mainly of technical manner and do not imply substantial differences. Further, no systematic changes are expected because of a different emission "grid" (the results are subsequently interpolated to T42 model resolution anyway) or through the nudging of dynamic variables, as the formation of contrails and chemical mechanisms remain unaffected by these changes in general. But impacts of non-CO_2 aviation emissions are expected to change with altitude, with latitude, with different chemical background concentrations, with weather situation and with season. Nevertheless, CCF results of the present study show largely comparable magnitudes and patterns with those of WP3 of REACT4C within the shared domain (e.g. Figure 8,9,17,18 for O_3), despite the different model setup, emission grid and season.*
- *Obviously, the comparison of the different CCFs were not clarified properly, which lead to a misunderstanding. We do not "extrapolate" old CCFs to compare with new CCFs, but we compare old (winter weather situation W3) and new CCF results (spring case study) in their shared horizontal and vertical domains (despite the different season). Nevertheless, old CCF results (from 5 typical weather situations in winter and 3 in summer) were correlated with meteorological variables at the time of emission (in case of O_3 CCFs with e.g. temperature and geopotential) and this correlation was used to derive algorithmic Climate Change Functions (aCCFs) (van Manen & Grewe, 2019). These aCCF-formulations are employed (using only temperature and geopotential at the*

new emission date and time of the present study: 26.3.2014) to calculate the aCCFs for the corresponding date. The resulting aCCFs (correlation-based estimate) are then compared to the new spring case study CCFs (detailed simulation) of the same date (26.3.2014) which were computed with detailed chemistry-climate model simulations with high computational effort (CCF vs aCCFs). This comparison reveals, that the actual form of the algorithms cannot sufficiently reproduce all characteristics of the CCFs for the spring date. At main flight levels (200-300 hPa), the overall magnitude and general altitudinal and latitudinal gradients were generally reproduced, however with systematic overestimation in case of O_3. At other altitudes (O_3 below 300 hPa and above 200 hPa, and H_2O above 200 hPa), the algorithms have larger general problems and are not recommended to be used there. CH_4 CCFs have general problems to reproduce gradients in the spring case correctly and should be revisited. Contrail CCFs and aCCFs underestimate the climate impact systematically because of low optical depths and need to be revised, which is currently pursued (Dietmüller et al., 2026, Peter et al., 2026, both in preparation).

The "novel season" argument is not well supported: the authors claim that this is an expansion to a "novel season", but they simulate only two days in one year. The justification then states that this is close to a winter pattern (lines 310-311), and states that a different paper (specifically Peter et al., 2025) benefited from this simulation – rather than justifying this choice because of, say, its ability to meaningfully represent a novel season. As above, I would recommend that – if the authors wish to argue that a novel season is represented – a dedicated evaluation of the need for a novel season to be represented should be performed along the lines of the original argument for 8 weather patterns to represent two seasons.

- *Thank you for this comment. Of course, one single date was not meant to represent the whole spring season, but is only a case study of a spring date. We changed the title and text throughout the manuscript to clarify this. In addition, we adapted the reasoning of the analysis with the two objectives of testing the "old" CCFs against a case study for April and similarly for the aCCFs.*
- *Indeed, the circulation pattern of 26./27.3.2014 is similar to that of WP3, however, atmospheric properties (e.g. temperature, humidity, etc.) and background chemistry differ and may cause significant changes regarding contrails and chemical effects in response to non-CO_2 aviation emissions in contrast to a date in winter. We revised the text to clarify this.*

The evaluation of (a)CCF accuracy is very limited: it appears that CCFs have three primary shortcomings which have then propagated to aCCFs. First, they cover only a limited geographical scope yet are extrapolated to other regions. Second, they are derived from a single model. Third, they are derived on the basis that eight weather patterns represent the majority of variability – but this has only been shown for the original domain (van Manen and Grewe, 2019). This study does not communicate the implications of its findings that the original CCFs/aCCFs did not accurately predict the new results (the first recommended change), nor does it meaningfully address any of these three shortcomings (the alternative recommended change). A thorough evaluation of aCCFs, or indeed of CCFs, would require more than just comparing results from one model with itself. I would strongly recommend that the authors consider bringing in different models, ideally as independent as possible, to give a meaningful evaluation of the degree to which CCFs or aCCFs represent any sort of consensus. Without this they represent just one model's estimate of the atmospheric response under a very limited set of circumstances.

- *Indeed, CCFs and aCCFs cover only a limited geographical and temporal scope. The usage outside the design region ("off-design") is not intended nor recommended (Dietmüller et al., 2022), yet it is done deliberately in the present study and results are accordingly assessed.*
- *So far episodical aviation climate impact studies regarding weather dependent, altitudinal, latitudinal and temporal sensitivity are sparse, hence evaluation of CCFs and aCCF is challenging. Grewe et al. (2014) found ranges of CCF results to compare well with the literature, and Yin et al. (2022) found general agreement of climatological aCCFs with other climatological studies.*

The abstract and key conclusions of the paper are not supported by the results: the authors argue that the new CCFs "compare well with earlier CCFs" (line 565), and that "aCCFs successfully reproduce the overall magnitudes and most gradients" of CCFs. I do not agree that this accurately represents the rest of the paper, and am confused as to why these conclusions are drawn. Elsewhere the authors state that the aCCFs for NO_x "overestimate the climate response by 25-30%" even within the target altitude range, and that the functions should not be used outside that range at all (line 578). The authors note that the estimated contrail RF is low compared to other studies, is likely biased due to a factor of 5 to 15 underestimation of optical depth (line 592), and shows a sign of daytime radiative forcing which disagrees with the scientific consensus (line 596). They also note that their specific example demonstrates how the CCFs and aCCFs derived from winter and summer are inadequate to describe a spring scenario (lines 605-608), meaning that not only are aCCFs not able to represent CCFs, but CCFs are not particularly generalizable. More generally the analysis seems confused; results such as those shown in Figure 6 show a lack of consistency between the old CCFs, the new CCFs (which again cannot be used outside of the two days simulated), and the aCCFs, with neither pattern nor magnitude showing agreement. The text however seems to oscillate between acknowledging this ("correlation between contrail CCFs and aCCFs is weak (r=0.2) and statistically not significant", line 470) and then framing the results as positive ("we find comparable magnitudes of CCFs and aCCFs"). My takeaway from this study is that aCCFs are not yet fit for use in trajectory planning, being unable to represent the variability even within their source CCFs and with little evidence provided that the CCFs are extensible or in agreement with other models. Phrasing such as "aCCFs offer a fast alternative for trajectory planning" (line 17) imply that it would be appropriate to do so, but based on the results from this work I would expect the abstract to instead strongly discourage the use of aCCFs or CCFs.

- *Thank you for mentioning this concern. Indeed, there are species and altitude dependent over-and underestimations. However, for route optimization not the exact value of climate response is key, but the orientation of horizontal and vertical gradients as these provide meaningful incentives for re-routing. Nevertheless, we report the deviations and limitations of aCCFs more precisely in the abstract and conclusion section and also highlight the above criterion for optimization. We also think, splitting the "comparison with earlier results-section" into two sections, following two different objectives, helps to understanding our aims and avoids mixing comparison with "old" CCFs and comparison with aCCFs for the spring date. The two objectives are:*
  a) to analyze how different CCFs for one spring case study are relative to an exemplary REACT4C weather situation and
  b) to analyze whether aCCFs can represent CCFs for spring, outside their seasonal, geographical and vertical design region.