Reply to

RC2: 'Comment on jecats-2026-3' , Anonymous referee #2, 20 Mar 2026

We thank the Reviewer for the thoughtful and insightful evaluation of our manuscript. We highly appreciate the constructive comments, which have helped us to improve both the presentation and the scientific clarity of our work. We have carefully considered all remarks and revised the manuscript accordingly.
Our detailed point-by-point responses are provided below, providing review comments (**in black**) and our responses (**in blue**).

This manuscript by Matthes et al. presents an overview of the D-KULT project, which aims at integrating eco-efficiency into routine flight operations. The workflow for this integration is presented, as well as new developments for each of its bricks that were conducted within the D-KULT project. A simulation-based ATM-led contrail avoidance trial is also presented. Recommendations for additional developments and smooth integration in the routine workflow are formulated. This study is important given the current focus on contrail avoidance from multiple stakeholders. The presentation of all achievements of the D-KULT project is very interesting and deemed for publication.

However, I have a few concerns about the current form of the paper. First, the scientific conclusions of the D-KULT project are not clearly stated, and its added value, although significant, is not sufficiently highlighted. Second, the manuscript does not sufficiently relate the advancements made in D-KULT to the state of the art and the existing literature. Third, I find the paper and its overall structure quite difficult to read, going back and forth on multiple subjects.

Therefore, I recommend publication after major revisions. Below I provide additional details about the major concerns I have. I also provide comments throughout the manuscript, some of them being linked to my major comments, and some minor comments.

Major comments

1) It is not clear, at the moment, what the scientific objective of the paper is. This comment is mostly driven by the fact that a vast majority of the results are reported in other publications and to some extent in other manuscripts in preparation but cited in this manuscript. This said, I see value in two main objectives and conclusions, but they need to be more highlighted, and additional details could be provided.
Reply: We reformulated the abstract to make the objective of the paper clearer.

1.1) First, the recommendations derived from the experience of D-KULT could be more clearly stated, as well as the challenges lying ahead. Most sections where I expected recommendations are in fact pure descriptions of the tools and what was done in the project. Typically, section 5 aims at describing the challenges, as stated in section 5.1. However, most of the subsections are summaries of what was already described elsewhere in the manuscript, with the exception of subsections 5.8 and 5.9. Similarly, I would have expected some conclusions on which tools should be used (sections 4.2 or 4.5), which methods for implementing contrail avoidance may be the most appropriate between airline-based (tactical or pre-tactical) or ATM-based avoidance, etc.
Reply: Thank you for the suggestion. We have now a list of issues that need to be solved

before eco-efficient flying can become usual practice. There is discussion of these issues in the new section 6 and the end of the abstract lists these issues as well, however, in short and not-discussed form. Even if the word "recommendation" does not appear, it should be clear from the context that we recommend these issues to be tackled, since otherwise there will be no progress.

1.2) Second, I would expect additional discussion on the results of the DFS simulation trial. At the moment, many qualitative conclusions are drawn, mainly that the workload of controllers has increased. It seems that no separate manuscript is planned for this trial, so this would be a good place to provide additional details. E.g., quantify the additional workload; conclude on the feasibility of avoidance led by ATM; could it be done within a less dense airspace? How to link PPC=1 area avoidance with eco-efficient planning, as the results suggest that climate impact was in fact increased? Should we pursue PPC=1 area avoidance or try something else, such as the concept of avoiding highly-warming climate-sensitive regions presented in section 4.2.3?

Reply: Thank you for this comment. We take it as a hint that the real-time simulation and its outcomes should be described in more detail. As the current paper is already quite long, we prefer not to enlarge the sections about the real-time simulation. Instead, DFS together with DLR & DWD have decided to generate an additional paper related to real-time simulation and follow-up activities as climate impact assessments and evaluation about capacity reduction. This paper is already in preparation.

2) At multiple places, the manuscript presents the achievements of D-KULT but without referring to or comparing to other methods / results available in the literature. It is difficult to understand what D-KULT has achieved when there is no comparison with existing studies.

Reply: We address comparison/evaluation of CCFs/aCCFs in Section 3.2.1. Further text was added regarding the verification of contrail detection algorithm (see below).

What struck me most is the absence of comparison in the observation and verification section. A new contrail detection algorithm based on the Mannstein (1999) algorithm is described. However, this algorithm relies on rather old concepts. How does this algorithm compare to more recent algorithms, e.g. Chevallier et al. (2023), Meijer (2024), Ng et al. (2024)? What are the advantages and drawbacks? Why only geostationary satellites were used, while other papers discussed the use of LEOs or ground cameras?

Reply: The reviewer addresses various points with his comment. We go through all of them in the following.

1) *First, the wording was probably misleading in the sense that we didn't aim at developing a fully new contrail detection. Instead, the goal of this work was to improve the Mannstein et al. (1999) method for contrail detection that had been implemented for LEO satellites, the AVHRR instrument, and directly translated to MSG. In fact, we adapted the overall Mannstein strategy for contrail detection to increase, in particular, the probability of detection. This is a central task when you want to evaluate contrail avoidance. This has been added to the text. In addition, the word "new" has been replaced by the word "improved".*
2) *The improved contrail detection has been compared to the initial implementations of the Mannstein method for MSG. This was the only comparable source of data available at the time of writing. Thus, a comparison is present in the validation section. For a comparison to the newer contrail detection methods, please see 4).*

3) *With respect to other algorithms for satellite-based contrail detection, we are well aware of the work by the colleagues you mentioned. In fact, for the current generation of American geostationary satellites GOES-R/S/T machine learning has been applied to the task of contrail detection with good success (Zhang et al., 2018; Kulik, 2019; Meijer et al., 2022; Chevallier et al., 2023; Ng et al., 2024). We haven't mentioned them because they are not directly applicable to MSG especially due to the better spatial resolution of the satellite observations. But of course, we are happy to cite them and have added a small paragraph.*
4) *With respect to the performance of the improved algorithm for MSG compared to these newer AI-based methods one must first point out that the comparison is challenging due to differences in sensor characteristics, spatial resolution, observational domain and time, input data, and evaluation criteria. Meijer et al. (2022) report a recall of 50.2 % and a precision of 52.5 %, while Ng et al. (2024) obtain a precision of 80 % for a 50% recall. These precision values are (substantially) higher than ours with a similar recall, thus suggesting that AI-based algorithms for the current generation of geostationary satellites of the GOES-R series generally achieve higher skill, but this may result not only from differences in methods, but also from differences in data, meteorological conditions and/or evaluation approaches. This has been added to the text.*
5) *The advantage of this traditional image processing technique is that it is easily extendable to other satellite sensors, as the application of Mannstein to MSG/SEVIRI shows, although it was originally developed for AVHRR. The disadvantages are that it does not exploit more modern techniques for image processing or ML. A short sentence has been added to the text.*

The reason why only geostationary satellites and no ground-based cameras or LEO satellites are used was given in part: "In principle contrails can be detected with ground-based instruments – cameras exploiting visible and/or thermal radiation. However, only satellites offer a global coverage and an unhindered view, independently of low-level clouds and illumination conditions, on contrails, which are formed in the upper troposphere and lower stratosphere (UTLS). In particular, geostationary platforms have a very large field of view and high repetition rates such that entire flight paths flown at any time of the day and the night can be investigated to identify contrails in their vicinity and are expected to prove whether an avoidance or an observation flight (one without avoidance to control the contrail prediction) was successful." The explanation has been expanded to include polar orbiting satellites.

I understand that the objective of the manuscript is not to provide an extensive comparison of each new development to the state-of-the-art. However, as one of the objectives is to describe a workflow for implementing contrail avoidance, I would argue that at least comparing the tools used in D-KULT to what already exists would be valuable.
Reply: See above. In addition, in the revised version we give more attention to the differences between aCCF and CoCiP and causes thereof. Further we added a short paragraph on advantages and drawbacks regarding aCCFs and CoCIP.

3) I find the structure of the paper difficult to follow. The authors chose to describe all existing tools in section 2, all new developments in section 3, and the application or evaluation of these new developments in section 4. However, this induces multiple repetitions between the

sections, references to Figures presented well above, and upon reading, some questions the reader have are only answered much later. Typically, when reading section 2.4, I'd like to know which of these tools are used and how, how they are interconnected, etc. But this information comes much later and in two batches, in section 3.3 and 4.2.
I would recommend to adapt the structure of the paper and present in different sections the bricks of the workflow. E.g. in section 2, current state and developments for met services would be described, along with the evaluation. In section 3, the observation system and its developments made in D-KULT could be described, etc. I understand that this requires a major revamp of the manuscript, but I want to stress again that the current manuscript is quite difficult to follow and could be much improved, in my opinion, if its structure was revised.
Reply: We acknowledge the difficulties and have restructured the whole manuscript. Accordingly, repetitions were removed. We thank the reviewer for this suggestion.

Specific comments

P1L32 The wording "strategical" and "tactical" for respectively pre-departure planning and pre-take-off / in-flight planning got me lost. I am more used to the wording "strategical", "pre-tactical" and "tactical", for respectively pre-departure (at least H-12 before departure), pre-take-off and in-flight. The terminology may not be completely settled but I note that you use this latter wording P7L220. You should harmonize the wording throughout the manuscript, and I suggest you use the "pre-tactical" wording for pre-take-off planning.
Reply: Indeed, the term "pre-tactical" is more appropriate than the term "strategic". If we are thinking on eco-efficient routing, we always need some kind of strategy, but this is something quite general (for instance it could mean that one only reroutes if the probability of success is higher than x%). For planning single flights, as done in D-KULT, we take the actual weather forecast into account instead of the climatological properties of the flown region. Thus, this is more special than a general strategy and thus we agree with the suggestion to use "pre-tactial" instead of "strategic".

P2L46 You could explicitly mention the Paris Agreement here. Moreover, the objective is closer to "well below 2°C, 1.5°C if possible".
Reply: Done.

P2L48 A citation is needed to support this statement.
Reply: Citation is Grewe et al., (2021), which has been already cited. We cite it here again in the revised version. Additionally, we cite a recent report of ICCT (2026).

P3L85 "Basic methods" is rather unclear. You could replace it by "Current workflow of flight operations" for example.
Reply: Resolved during restructuring of the manuscript.

P3L90 This sentence is a repetition of the previous one.
Reply: Thanks for checking, in fact that was by mistake. This sentence was deleted.

P3L94 Typo "focussing on i.e." -> "focusing on e.g."
Reply: Corrected.

p4L123 I would expect in this section a small description of the advantages and drawbacks of aCCFs and CoCiP. While the advantages are well described, the drawbacks are not.
Reply: We added a short paragraph on advantages and drawbacks regarding aCCFs and CoCIP.

P5L154 The two parts of this sentence are not related. The sentence should be split in two.
Reply: Indeed, the sentence as it stands makes no sense. We change it into these two: "Strong sensitivity of climate response to humidity data is reported. The latest version of CoCiP is available as open-source software in the Python package pycontrails (Shapiro et al., 2023)." Then the first of these sentences is linked to the descriptions of tests in the sentence before.

P6L166 The pyTOM model is not described in this section, contrarily to the other two tools.
Reply: pyTOM was originally intended to provide comparable solutions for large flight scenarios simulated with Lido Flight 4D in order to quantify the efficiency loss associated with legal trajectories compared to unconstrained climate-optimal trajectories. However, due to developments related to the 100-flights trial, a full implementation of this large-scale comparative assessment was no longer feasible. Nevertheless, pyTOM was successfully applied to a number of smaller scenarios and played a supporting role in the development and validation of the multi-criteria optimization capabilities of Lido Flight 4D

P6L178 I don't understand the added value of this paragraph. It comes out of the blues and does not provide any valuable information to the reader. I would either remove it or rephrase it significantly.
Reply: Deleted and no longer present in the revised version.

P7L203 "relatively high false alarm rates and low detection efficiencies" Could you provide a number?
Reply: We added this information in the text.

P7L207 This paragraph is difficult to read. I do not understand the difference between the first sentence ("project [...] investigate the feasibility of integrating…") and the last one ("the objectives of the ongoing projects are to investigate the feasibility…"). Please clarify this statement.
Reply: Reformulated the paragraph.

P7L207 Which projects? Or add a citation.

Reply: We orient the reader to a publication from the first project in this series which is REACT4C with was Matthes et al., 2012. For an overview on research and development projects we orient the reader to the publication from Simorgh et al. 2022.

P7L212 Which projects? Or add a citation.

Reply: This sentence mentioning projects has been deleted.

P8L227 Typo "were" -> "where"

Reply: Corrected.

P11L266 Why was this value chosen? A reference would be appreciated.
Reply: DFS has determined this value for flights in its airspace. There is no reference. This is now indicated in the paper.

P11L269 Can you provide a reference to support this statement?
Reply: Yes. Sperber and Gierens (2023) and Gierens et al. (2022) will suffice.

P11L271 How does the choice of a lower threshold compare to other published methods of humidity correction? Why other more advanced corrections were not considered. These corrections should at least be mentioned, as it is the subject of numerous other works (e.g., Teoh et al., 2024; Platt et al., 2024; Wang et al., 2025; Wolf et al., 2025).
Reply: Sorry, but this is irrelevant. In our projects we did not correct the RHi-field, we only chose a threshold value from which on we considered the water vapour as supersaturated with respect to ice. For diagnosing contrail persistence, this suffices. A correction of the humidity field may be necessary for simulating the evolution of a contrail in this field, but this was not done with the Clima-1 data. No change in the paper!

P11L283 For the Clima-1 dataset, could the average of the 40 members have been used rather than the deterministic run, to ensure consistency with the experimental Clima-1s? More generally, is taking solely the deterministic run more or less "valid" than taking the average of the members?
Reply: Of course, from a technical point of view this could be done. The starting point from the conventional workflow was usage of the deterministic run, hence Clima-1. However, intentionally the experimental operational simulation was foreseen as an ensemble, in order to explore atmospheric variability. For this purpose, the Clima-1 run was used as initial condition. It is correct that comparing either two single members, two ensemble mean states would represent a comparison between two dataset with identical variability. Furthermore, one reason for exploring the deterministic run was that it provides physically consistent states, while the average of the ensemble does not.

P11L289 "Contrail avoidance should only be performed if the contrail would have a measurable effect on the outgoing radiation" Shouldn't this be accounted for with the climate response functions? If these functions do not inform on the potential radiative effect of contrails, they have no use in my opinion. How does the TRM and the climate response function should be articulated during operational avoidance? Note that my comment focuses on the second purpose of TRM - I agree that this quantity is of great importance to assess whether a contrail can be observed, in particular for avoidance experiments.
Reply: Yes, indeed. We think you made a valid point here. The text is now rephrased accordingly.

(B7) P12L295 This sentence is a repetition of the paragraph starting at P12L301. I would simply delete it.
Reply: Several repetitions were removed during the restructuring of the manuscript.

P12L296 The F in front of F-ATR100 is not explained. Why was this metric chosen rather than other ones? In particular considering the metrics used in the MRV.
Reply: We added this information and address the choice of metric.

P12L301 How are the outputs of CoCiP converted in F-ATR100? Which value is used? How are the meters of contrails formed converted in kg of fuel? What is the assumed aircraft type? The underlying performance model?
Reply: We added this information to the text. A detailed overview of factors used for the calculation of climate response functions is given in Appendix B.

P12L305 The Clima-4s dataset does not include any CoCiP calculation according to Table 1, it only contains raw ICON outputs. Please correct this sentence or Table 1 depending on the variables contained in Clima-4s.
Reply: Corrected.

P13L318 At this point, I don't find this discussion sufficient to explain the major differences between the aCCFs and CoCiP. I only have additional information much later, in section 4.1.3, not accounting for the fact that most of this discussion has been deferred to a future publication. The two discussions should be reunited at the same place in the paper.
Reply:We added more detailed discussion on the major differences between aCCFs and CoCIP and combined the discussion during the restructuring of the manuscript.

Figure 3. Why was the CoCiP output masked with the ensemble mean PPC? What is the scientific basis of doing this? CoCiP already includes a test for contrail formation and persistence. The reason for applying this mask needs to be explained.
Reply: Contrary to Clima-1s in Clima-3s contrail climate responses were calculated individually for ensemble members, which lead to a large spread in areas covered with contrails. We added this information in the text.

P13L322 This paragraph comes a bit out of the blues. I would recommend moving it, e.g. in section 3.2.1.
Reply: We moved the paragraph to section 3.2.1 and some of the details to Table A1.

P14L338 What is the value of the conversion factor? If it's not constant, can you provide an average value?
Reply: Added more detailed explanation how this conversion factor was derived.

P14L343 This paragraph should come much earlier, in the description of the tools.
Reply: True. We give a detailed description of pyTOM in section 4, now.

P14L344 Why are continuous trajectories used? What is the added value? As the trajectory is likely not legal, is this tool used solely for scientific purposes? More importantly, pyTOM and the associated trajectories are not mentioned again in the manuscript. It is really unclear how this tool integrates the workflow, and what is its purpose. This needs to be clarified, or alternatively, the mentions of this tool should be dropped.
Reply: Although the resulting continuously optimized trajectories are not always directly operationally feasible, they provide an important theoretical benchmark for assessing the maximum achievable reduction in climate impact. A comparison between such unconstrained optimal trajectories and legally flyable trajectories was conducted in a number of smaller case studies. While pyTOM was originally intended to enable a large-scale benchmarking analysis of extensive Lido Flight 4D scenario ensembles, the insights and

methodological developments emerging from the 100-flights trial ultimately prevented the implementation of such a comprehensive assessment. Nevertheless, the smaller-scale comparative studies delivered valuable insights into the trade-offs between climate optimality and operational feasibility and contributed significantly to the development, validation, and calibration of the multi-criteria optimisation capabilities using the aCCFs.

P15L356 How is the optimization calculated with this tool? Is it the same as for the Lido optimization (minimization of the cost)? I understand from reading section 4.2.3 that the optimizer may simply avoid PPC=1 areas, but this explanation should come as of section 3.3.2.
Reply: LIDO always uses all climate responses that are available in the WAWFOR-Clima-X data, even in the planning of the TF-100 flights. In contrast, the FPO cloud software has two modes of operation, as described in the paper, PPC-mode (contrail avoidance acc. to the PPC values) and ECO-mode (using all aCCFs).

P16L387 Why was the buffer not necessary? How was this conclusion drawn?
Reply: In planning the exercise, it was foreseen that aircraft will be brought back to its original routes in that buffering area. However, the simulation showed that this buffer was not necessary, since regular coordination procedures with adjacent sectors still prove to be the best solution.

Figure 6. The text on the right panel is difficult to read. If possible, could you use a higher-resolution screenshot?
Reply: This figure will be deleted.

P17L394 What are the requirements? A list of them would be very valuable.
Reply: The text has been rewritten. The old sentence has been deleted and replaced by: "A final representation version will be customized, as the development and definition of an optimal display remain important next steps for the practical implementation of the procedure to avoid persistent contrails." We think that a listing of dedicated requirements is not appropriate in that paper. We will assess them internally at DFS.

P18L406 I guess the end of the sentence is missing?
Reply: Sentence is long, but nothing is missing.

P20L483 Is the ETS the best metric to assess forecast stability? Dean et al. (2025) showed that forecasts are relatively stable when considering patches of PPC=1 regions (or small dilations of such patches) rather than point-to-point comparison.
Reply: There are certainly other possibilities to assess the forecast stability, but in the cited paper (von Bonhorst et al., 2025) ETS (and CSI) was chosen and justification for this choice was given there. Indeed, these quantities count point-to-point comparisons, but the plots in this paper show more: they show that robust prediction is connected to quite organized (large-scale) structures while non-robust predictions are characterized by a patchy mess of super- and subsaturated (or PPC=1 and PPC=0) regions. These characteristics are further confirmed in a just submitted paper (von Koslowski und Gierens, 2026). Perhaps, there are closer agreements with Dean et al. than one would think of by just thinking of the metrics used. In particular, we agree that the planning of contrail avoidance should not be based on pointwise predictions of PPC=1. But, at this location in the paper, this is not the point. Here the point is, how the robustness of forecasts changes in time with forecast length (von Bonhorst et al.) and during the year (von Koslowski and Gierens).
The text has been rewritten to include the new results.

P21L487 By forecast uncertainty, I guess you only refer to the uncertainty arising from stability? If so, please state it explicitly.
Reply: No. We refer to the fact that ISSRs are not always predicted at the location and time when they occur. "Stability" in this respect only implies that the ensemble forecasts may show something different from the deterministic one and that this changes with forecast horizon and with season (which is a new aspect).

P21L489 I don't understand the purpose of this paragraph, which is not related to stability. I recommend removing it.
Reply: The paragraph has been removed as suggested.

P21L495 The purpose of this paragraph is unclear, in this validation section. It could be moved to the description of the tools.
Reply:We rephrased the paragraph, adding more details on major differences between aCCFs and CoCIP and moved it to section 3.2.1.

P22L525 How is this sensitivity used in the workflow? Is it used in operations?

Reply: No. In the operations, LIDO only uses the mono-criterial optimization, that is, both climate costs and operational costs are minimized together with a fixed relation between these to. The bi-criterial optimization is used only as a research tool. There is now a statement of this in the new sect. 4.2.1.

P23L536 Typo, I guess this should be section 4.2.2.
Reply: Indeed, this was wrongly numbered. It has been resolved during the restructuring of the manuscript.

P23L547 Why are the data masked by areas with natural cloud cover? How is natural cloud cover related to TRM?
Reply: This is described in the paper by Gierens (2023). It is a quite complicated procedure and there is no simple formula. One can say, TRM is a stochastic sum of cloud cover times their individual transmissivity, taking into account random overlap for detached cloud layers and assuming maximum overlap for adjacent cloud layers. As this paper is available, there is no need to describe the details in this concept paper.

P25L586 Is this how workload is assessed? If so, please state it explicitly.
Reply: Sentence is changed into: "Furthermore, briefing materials for controllers have been prepared, as well as questionnaires before, during and after each simulation run. These were used to assess workload and situational awareness, as well as to collect general feedback on the simulation."
In detail, we use a modified tool of Instantaneous Self-Assessment (ISA from NATS) to record workload and situational awareness every 5 minutes during simulation run.

Figure 13. Are the results calculated using ERA5 or Clima-1? In any case, could you present the results with both datasets? Moreover, why is the CiC for PPC scenario that high, while many trajectories were rerouted in a PPC=0 area? This needs to be discussed more extensively in the text. See also major comment 1.2.
Reply: Results are calculated using ERA5, as tool for calculating climate responses can't use ICON-data.
It should also be noted that a detailed investigation revealed significant discrepancies in the

representation of PPC areas between the DWD ICON-EU and ECMWF ERA5 models, likely driven by differences in model physics, input data (e.g., humidity), and calculation approaches (DFS PPC tool vs. aCCF-model).

P29L672 Could you provide a recommendation of which contrail detection algorithm to use, based on your results? At the moment I don't know what to conclude on this.
Reply: We are sorry the reviewer is confused. We have added a sentence to ***Sect. 5.2*** *meaning that we prefer a higher POD (recall) even if paired with a high false alarm rate because for avoidance flights it is crucial to observe contrails if there are some. False alarms must then be excluded in another way, e.g. considering the temporal evolution of these objects.*

P31L739 You only used ATR100 in this paper, so this statement seems inexact.
Reply: We agree and we revised this paragraph and only generally state, as known from the literature, that short-term effects as from contrails gain importance over long term effects as from CO2, when looking at shorter time horizons. Nevertheless, pointing out this relation to the reader seems worthwhile.

P31L748 Could you explicitly state that this corresponds to the ATM-led simulation trial to be completely clear?
Reply: We assume that the comment corresponds to the sentence: "Since uncertainties in the input data were not assessed, a realistic error margin could cross the zero line, making a beneficial overall outcome uncertain. For future evaluations the analysis framework must be improved and uncertainties should be fully assessed." This general remark applies to all climate impact assessments within the D-KULT project

P31L748 This is a discussion on uncertainties, not sensitivity, thus should be moved above.
Reply: In this case, the wording is, to our view correct. The section is about sensitivities (eg. sensitivity of results when changing metrics), but the sentence starting in L748 simply says that the analysis was not complete, since there are uncertainties in the system which have not been considered. While we agree that there is subtle difference between the notions "sensitivity" and "uncertainty", we believe that the given description is correct. We prefer to leave the text as it is.

P32L774 Should the recommendation emphasize PPC areas rather than high-impact climate sensitive regions, such as the regions of level 4 used in FPO?
Reply: Good point! We have now replaced "PPC-areas" with "climate-sensitive regions". This should suffice for the paper. But it should be considered that PPC-areas are much more variable in space and time than aCCF-fields of gaseous species. Thus contrail-avoidance or avoidance of PPC=1 regions will in operations much more often come into the reach of tactical avoidance (when the PPC-areas have changed significantly), where ATC and FPO are the relevant actors, while regions with high aCCFs of gases can much more often been tackled by pre-tactical planning. These considerations will not be pursued in this paper.

P32L783 I would move this sentence in the conclusion section, and focus solely on recommendations.
Reply: We think, with the new structure of the paper, this sentence fits best in the new

section 6, where open issues are discussed.

P33L793 This conclusion could be separated into two separate points, one on measurements and one on the 2-moment cloud scheme.
Reply: We agree and split the text in two separate points.

P33L801 What is the challenge linked to this point? Moreover, this is the first time this is mentioned in the paper.
Reply: Perhaps it was not clearly enough expressed. It is a result of the quoted study of Peter et al. It is a central point and it is now emphasized in the last part of the abstract as well. We found that, say, difficulties and inconsistencies arose when, e.g. flight planning was based on WAWFOR-Clima-X, but the assessment of climate-benefits was then performed using ERA5 data. Similar problems arise, when different airlines plan with data derived from different NWP models. They see a different "reality". We hope the new text is now clearer and emphasize, that this a indeed a major issue.

P33L809 What is the challenge linked to this point?
Reply: Sentence is rephrased: "Climate effect estimates depend on the selected input weather data. Inconsistencies may arise when different sources are used."

P33L813 This is the first time in the paper such a trial is mentioned. Could it be mentioned before, e.g. in section 4.2?
Reply: Thank you for the hint. We have now mentioned the trial already in the abstract and in Section 2. Note, however, this TF100 trial is not in the central focus of the current paper.

References

Chevallier, R., Shapiro, M., Engberg, Z., Soler, M., and Delahaye, D.: Linear Contrails Detection, Tracking and Matching with Aircraft Using Geostationary Satellite and Air Traffic Data, Aerospace, 10, 578, https://doi.org/10.3390/aerospace10070578, 2023.

Dean, T. R., Abbott, T. H., Engberg, Z., Masson, N., Teoh, R., Itcovitz, J. P., Stettler, M. E. J., and Shapiro, M. L.: Impact of forecast stability on navigational contrail avoidance, Environ. Res.: Infrastruct. Sustain., 5, 045008, https://doi.org/10.1088/2634-4505/ae1da5, 2025.

Meijer, V. R.: Satellite-based analysis and forecast evaluation of aviation contrails, Thesis, Massachusetts Institute of Technology, 2024.

Ng, J. Y.-H., McCloskey, K., Cui, J., Meijer, V. R., Brand, E., Sarna, A., Goyal, N., Van Arsdale, C., and Geraedts, S.: Contrail Detection on GOES-16 ABI With the OpenContrails Dataset, IEEE Transactions on Geoscience and Remote Sensing, 62, 1–14, https://doi.org/10.1109/TGRS.2023.3345226, 2024.

Platt, J. C., Shapiro, M. L., Engberg, Z., McCloskey, K., Geraedts, S., Sankar, T., Stettler, M. E. J., Teoh, R., Schumann, U., Rohs, S., Brand, E., and Arsdale, C. V.: The effect of uncertainty in humidity and model parameters on the prediction of contrail energy forcing, Environ. Res. Commun., 6, 095015, https://doi.org/10.1088/2515-7620/ad6ee5, 2024.

Teoh, R., Engberg, Z., Schumann, U., Voigt, C., Shapiro, M., Rohs, S., and Stettler, M. E. J.: Global aviation contrail climate effects from 2019 to 2021, Atmospheric Chemistry and Physics, 24, 6071–6093, https://doi.org/10.5194/acp-24-6071-2024, 2024.

Wang, Z., Bugliaro, L., Gierens, K., Hegglin, M. I., Rohs, S., Petzold, A., Kaufmann, S., and Voigt, C.: Machine learning for improvement of upper-tropospheric relative humidity in ERA5. weather model data, Atmospheric Chemistry and Physics, 25, 2845–2861, https://doi.org/10.5194/acp-25-2845-2025, 2025.

Wolf, K., Bellouin, N., Boucher, O., Rohs, S., and Li, Y.: Correction of ERA5 temperature and relative humidity biases by bivariate quantile mapping for contrail formation analysis, Atmospheric Chemistry and Physics, 25, 157–181, https://doi.org/10.5194/acp-25-157-2025, 2025.