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.

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.

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?

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.

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?

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.

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.

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.

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

P2L48 A citation is needed to support this statement.

P3L85 "Basic methods" is rather unclear. You could replace it by "Current workflow of flight operations" for example.

P3L90 This sentence is a repetition of the previous one.

P3L94 Typo “focussing on i.e.” -> “focusing on e.g.”

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.

P5L154 The two parts of this sentence are not related. The sentence should be split in two.

P6L166 The pyTOM model is not described in this section, contrarily to the other two tools.

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.

P7L203 “relatively high false alarm rates and low detection efficiencies” Could you provide a number?

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.

P7L207 Which projects? Or add a citation.

P7L212 Which projects? Or add a citation.

P8L227 Typo “were” -> “where”

P11L266 Why was this value chosen? A reference would be appreciated.

P11L269 Can you provide a reference to support this statement?

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).

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?

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.

P12L295 This sentence is a repetition of the paragraph starting at P12L301. I would simply delete it.

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.

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?

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.

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.

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.

P13L322 This paragraph comes a bit out of the blues. I would recommend moving it, e.g. in section 3.2.1.

P14L338 What is the value of the conversion factor? If it's not constant, can you provide an average value?

P14L343 This paragraph should come much earlier, in the description of the tools.

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.

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.

P16L387 Why was the buffer not necessary? How was this conclusion drawn?

Figure 6. The text on the right panel is difficult to read. If possible, could you use a higher-resolution screenshot?

P17L394 What are the requirements? A list of them would be very valuable.

P18L406 I guess the end of the sentence 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.

P21L487 By forecast uncertainty, I guess you only refer to the uncertainty arising from stability? If so, please state it explicitly.

P21L489 I don't understand the purpose of this paragraph, which is not related to stability. I recommend removing it.

P21L495 The purpose of this paragraph is unclear, in this validation section. It could be moved to the description of the tools.

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

P23L536 Typo, I guess this should be section 4.2.2.

P23L547 Why are the data masked by areas with natural cloud cover? How is natural cloud cover related to TRM?

P25L586 Is this how workload is assessed? If so, please state it explicitly.

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.

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.

P31L739 You only used ATR100 in this paper, so this statement seems inexact.

P31L748 Could you explicitly state that this corresponds to the ATM-led simulation trial to be completely clear?

P31L748 This is a discussion on uncertainties, not sensitivity, thus should be moved above.

P32L774 Should the recommendation emphasize PPC areas rather than high-impact climate sensitive regions, such as the regions of level 4 used in FPO?

P32L783 I would move this sentence in the conclusion section, and focus solely on recommendations.

P33L793 This conclusion could be separated into two separate points, one on measurements and one on the 2-moment cloud scheme.

P33L801 What is the challenge linked to this point? Moreover, this is the first time this is mentioned in the paper.

P33L809 What is the challenge linked to this point?

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?

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.