Reply to Reviewer #1

(Referee comment on "Correction of temperature and relative humidity biases in ERA5 by bivariate quantile mapping: Implications for contrail classification" by K. Wolf et al. (egusphere-2023-2356), https://doi.org/10.5194/egusphere-2023-2356-RC1, 2023)

We thank the Reviewer for the time she/he spent on the manuscript and for the useful comments. We have addressed all the comments from the Reviewer, which has improved the manuscript.

For better legibility, the Reviewer's comments are highlighted in **bold** and changes in the manuscript are in *italic*.

General Comments:

The paper evaluates the ability of ERA5 to correctly represent temperature and relative humidity in the upper troposphere using data from IAGOS as reference. Biases in both fields are detected and characterised and a correction method is applied that corrects both T and then RHi using the so-called quantile mapping method. It is found that this method indeed is able to reduce the biases. Unfortunately it turns out that in spite of this improvement the prediction of contrail formation or contrail persistence is not improved. But at least, for other quantities that are relevant for the assessment of aviation climate impacts, e.g. the optical thickness of contrails, improved estimation seems possible (not individually, but in a statistical sense).

While the method leads certainly to a significant improvement of the RHi-statistics of ERA5, the investigations and analyses in the paper are somtimes a bit lengthy and, at least to my view, not always necessary. The differences between the original ERA5, and after the application of the QM method of the correction of Teoh et al are often minor and it is not sure whether they are always real or simply caused by statistical noise. But even if they are real in certain cases, it is not always clear to me why the reader needs to know the potential causes of these differences. I have the feeling that the reader can easily get lost in these details and that the straight way from the analysis to the results and conclusions becomes unclear. So, I think, there is potential to make the paper more concise and to clearly convey a message.

Multiple sections and paragraphs of the manuscript have been shortened and revised to make the manuscript more concise. For specific modifications to the manuscript we would like to point the Reviewer to the revised manuscript with track-changes and to the specific answers to major comment number 2 as well as to minor comments numbers 14 and 16.

This paper needs at some places more elaborateness in its formulations. For

instance, the sentence (Page 1) "IAGOS flight trajectories are used to extract co-located meteorological conditions - temperature, relative humidity, and wind speed - and are compared with the IAGOS measurements" is obviously a faulty. In short, it says that IAGOS trajectories are compared with IAGOS measurements, which is evidently nonsense.

Following the Reviewer's comment the sentence has been corrected.

"IAGOS flight trajectories are used to extract co-located meteorological conditions from ERA5, namely temperature and relative humidity, which are compared with the IAGOS measurements."

Also "representation of contrail occurrence in ERA5" (P 1) is misleading, since contrails are not represented in the ECMWF model and reanalysis.

The sentence has been rephrased to be clearer.

"*To estimate potential contrail formation on the basis of ERA5, data points from IAGOS as well as corrected and uncorrected data points from ERA5, temperature and relative humidity, are flagged for contrail formation using the Schmidt-Appleman criterion*."

Therefore I recommend a major revision of this paper to make its messages stronger and clearer.

Major issues:

1.) Reading the abstract it seems that the evaluation merely uses bulk statistics, which I deem insufficient. An good indication of this is figure 5, where, apart from MD, all other measures are very insensitive to the applied corrections. The r^2 hardly change (and these tiny changes may be insignificant on a $p=0.05$ level). The r^2, like the other insensitive statistics, are bulk measures, that say nothing to a point-by-point comparison. The authors acknowledge this in the first sentence of section 3.4.

The Reviewer is right that we primarily use bulk statistics. However, these bulk statistics do rely on a point-by-point comparison after careful sampling of the ERA5 data. The proposed quantile mapping technique is commonly applied to remove biases in data sets, which is the goal of this study. It is not designed to improve the other statistics. Our study is nevertheless an improvement over the state-of-the-art. Earlier correction methods of relative humidity from ERA-interim and ERA5 estimated the contrail occurrence based on a simple scaling technique, which, in principle also aims to compensate a potential bias in the dataset. Our method is a more elaborate way of applying a bias correction as we not only use an established method but we also apply it as a function of temperature.

A more sophisticated method aiming at exploiting more of the ERA5 variables and/or their three-dimensional structure remains to be developed. This is the subject of ongoing work.

2.) Section 3.4: The selection of the score values is an unlucky choice to my view, since the TN cases dominate in this data set. Thus, if one would ignore everything and always predict "no ISS", one would already get an impressive accuracy and FA rate. The statement "which indicates that the overall performance of ERA5, even in the uncorrected form, is at least similar or even has improved" is thus misleading. It may be that the author made their choice to compare with Tompkins et al. (2007). If the goal of the paper is to provide better data on the occurrence of ISS, this might be ok, although I still think, that scores that downweight the default TN would be better. If the ultimate goal is, however, forecast of contrail persistence, I think, that the authors should better compare with Gierens et al. (2020) and also use the ETS score described there.

We followed the suggestion of the Reviewer and computed the equitable threat score (ETS) based on the four entries in the contingency table. Following this work, we replaced the calculation and discussion of hit rate, false positive rate, and accuracy with the ETS, which is indeed more informative. The calculated ETS is given in Table 5. Accordingly, the entire sections “Along-track contrail formation potential and the effect of applied corrections” and “Analysis of collocated contrail formation potential from ERA5 and IAGOS” were revised.

in the course of the revision of the metrics, we found a mistake in the contrail flagging algorithm, which has now been corrected. This leads to slightly different results in the contrail categorization and the contingency table in Figures 7 and 8 as well as Tables 3 and 5. The quantile mapping itself remains unchanged but the conversion from RHi w.r.t to ice and liquid water was unnecessarily applied in succession, leading to a cumulative error in the conversion of RH between ice and liquid water. The error was dominant for samples with low temperature, where the saturation curves of liquid water and ice are close to each other and small errors had a significant effect on the calculated RH (Ambaum, 2020).

The revised analysis and the evaluation of the impact of the bias correction on the contrail categorization with the ETS now shows an improvement of NPC, PC, and R detection from ERA5 data against the IAGOS reference observations. A similar improvement was identified for the scaling method in Teoh et al. (2022).

We refer the Reviewer to sections “Along-track contrail formation potential and the effect of applied corrections” and “Analysis of collocated contrail formation potential from ERA5 and IAGOS” in the revised manuscript and the manuscript with track-changes.

Minor issues:

1.) Lines 34-38: I suggest that you state that RF is a global (or at least regional) quantity, averaged over a long time period. On first reading, it was not so clear to me whether you refer to single contrails or contrails in general. Furthermore, aren't the quoted values ERF values (in Lee et al.) rather than RF values?

To be consistent within the text, we removed the citation from Lee et al. (2021) and only kept the references from Boucher et al. (2021) and Burkhardt and Kärcher (2011), who give estimates for the radiative forcing.

"The influence of a perturbation, e.g., clouds, aerosols, or gases, on the Earth's atmosphere and its radiative transfer is quantified by the radiative forcing (RF). By definition, RF is defined as the difference in the net irradiance at the top of atmosphere under perturbed and unperturbed conditionsI(Ramanathan et al., (1989)). In the context of climate studies the RF is understood as the difference in the Earth energy budget due to a contributor to climate change (Bickel et al., 2020). For example, the aviation-induced global CO2-related RF is estimated to be around 30 mWm^{-2} (Boucher et al., 2021). Contrail RF is estimated to be stronger, at about 60 mWm^{-2}, but is subject to much larger uncertainties (Burkhardt and Kärcher, 2011)."

2.) L 48 and following: This text mixes up different things which is not good. In order to avoid contrails, one needs a precise PREDICTION of where they WILL occur. Knowledge of their occurrence is insufficient, it can only result in a kind of climatology. Schumann's use of the roof-camera is a bad example for predictive purposes, as it shows contrails that already exist and perhaps exist already quite some time. The same comment can be made for satellite or other observations of already present contrails. Then, contrail simulations in a climate model are not intended for contrail avoidance or prediction. CoCiP is the only example here where contrail prediction is the intention, but of course only, if it is fed with actual weather forecasts, not with ERA5.

We understand the objection of the Reviewer and rephrased the paragraph in the manuscript. We agree that there is only one way to avoid contrails, and this is by using numerical weather prediction models that predict regions that need to be avoided. However, it is useful to obtain a statistical database that documents the spatial and temporal distribution of regions where contrails are likely to form. The paragraph has been rephrased as follows:

"To lower the climate impact of aviation it is important to reduce CO2 as well as non-CO2 effects. An approach to minimize non-CO2 effects is active flight re-routing to avoid areas where contrails are likely to form and persist, which would require accurate numerical weather predictions. A useful prerequisite is to identify and document flight levels and regions of the Earth's atmosphere that are particularly prone to contrail formation due to meteorological and dynamical conditions that favor contrail formation. Such a statistical data base might be obtained in four different ways.

The first approach builds on ground-based observations. For example, [...]"

3.) L 75: It is a bit surprising that a "fourth approach" is now mentioned. I believe that the mixed-up list from above (L 48 ff) is now continued, but that is not sufficiently clear. Again, this is not an approach to prediction and should thus not be mentioned as something that has to do with contrail avoidance. I find also, that this section interrupts the logic of the argumentation. It would be better if the

paragraph that introduces various correction attempts would directly by followed with tha paragraph explaining the goal of the present paper.

We agree with the Reviewer and rearranged the paragraphs for a better logical flow. Please also see the answer to minor comment 2.
We now clearly state that the primary way to avoid contrails is active re-routing of flights by using numerical weather prediction and by circumventing contrail prone areas. However, creating a statistical database of contrail formation potential can be achieved by the four methods that we listed. (The list is not necessarily exhaustive.) Due to the length of the paragraphs we direct the Reviewer to the updated manuscript and the manuscript with track-changes.

4.) Fig. 2 and corresponding text: it seems that either the information in the figure is useless or something is wrong. For instance, why are there so many 175hPa data for EU, where I would expect a lot of landing and departures (i.e. low altitudes)? How do I have to read the figures? It seems, the interpratation is: on 300 hPa most flights are over EU etc..... and on 175 hPa again most flights are over EU. Is this, because most flights are over EU anyway? Then the information is useless. Should't it rather be the following: In EU most flights are on lower altitudes because of a lower fraction of cruise, similar in US but perhaps not as strong, and over the NA all flights are in cruise, therefore a predominance of high altitudes (or low pressures). It seems, the data should be organised the other way, i.e three panels "EU", "NA", "US" and then showing the fraction of pressure levels in the threee panels.

As described in section “Quantile mapping”, the CDFs have been calculated on individual pressure levels that are determined by the ERA5 pressure levels. Individual levels have been considered because the temperature and humidity biases do not have to be consistent in the vertical (across pressure levels) nor in the horizontal (across sub-domains); see Fig. B1 in the appendix. Considering the spatial dependencies of the biases, it is most important to be aware of the distribution of the IAGOS data.

By separating into individual pressure levels (Fig. 2) we intended to clearly show the respective contributions of each region to measurements for a given pressure layer. It should be noted that the total number of samples is given for each sub-panel of the Figure. The majority of samples are obtained at the 200 hPa level, which corresponds to the cruise level.
The Reviewer pointed to the 175 hPa layer, which she / he assumes to be incorrect. It is true that the majority of IAGOS flights depart and arrive in the EU. However, the fraction of samples in the EU domain on the 175 hPa layer is high because aircraft reach their maximum altitude at the end of the flight (when the fuel tanks are almost empty).

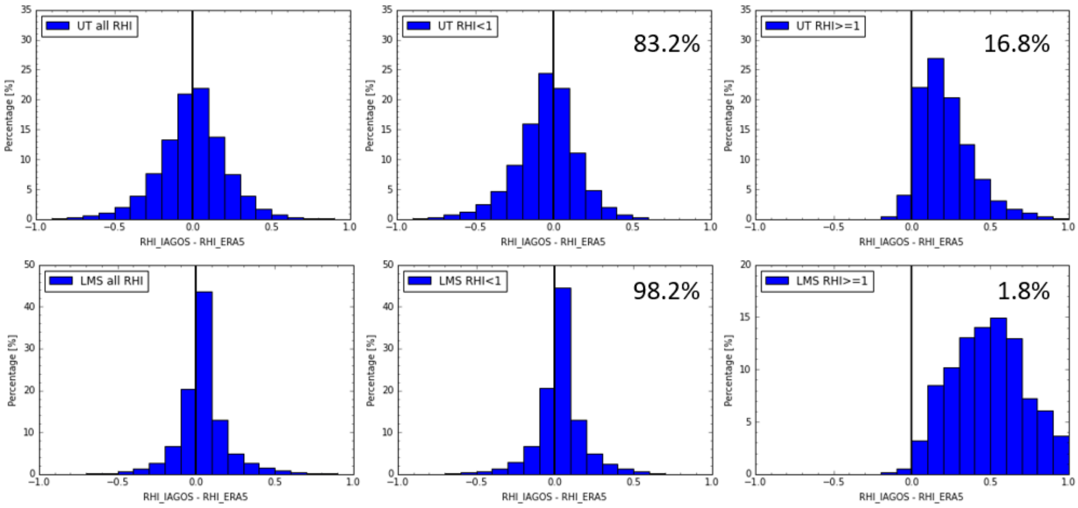
5.) Section 2.1: Please indicate whether all data are used along a flight (it seems so) or a subset to avoid autocorrelation which might spoil the statistics. If autocorrelation has not been avoided, a check should be made whether this affects the results or at least good arguments for this should be given.

All IAGOS data was filtered for data quality using the provided quality flags provided by the IAGOS post-processing. After filtering, around 90 % of all the data was usable.
The analysis that is presented in the manuscript has been performed with all remaining 90% of the data but also for a second time, where only every fourth data point (one data point approx. every 4 km) was used (second analysis not shown in the manuscript). No significant differences in the results were found.

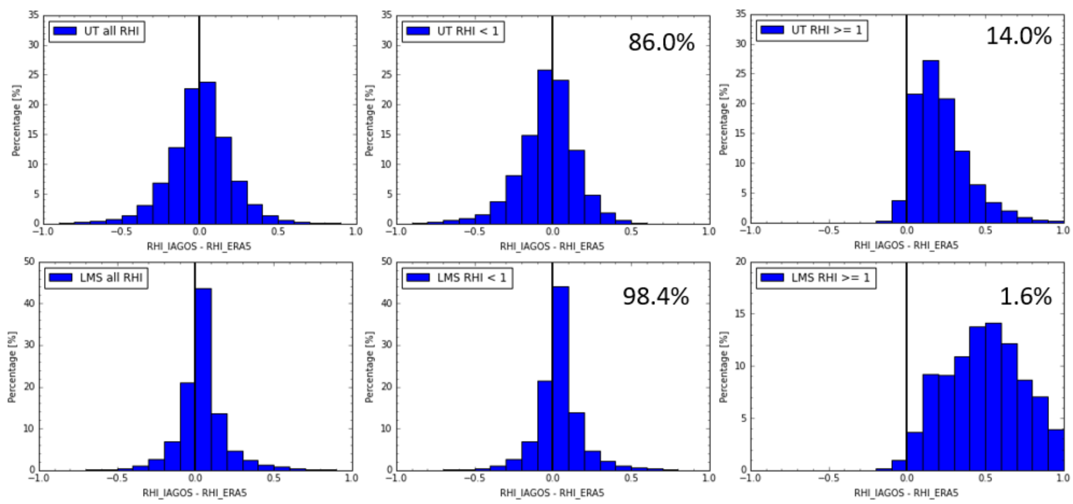
The data extraction method, i.e., selecting and sampling collocated data from ERA5 based on IAGOS flights in an identical manner, might lead to autocorrelation, which then exists in both extracted data sets and, therefore, should not influence the analysis. Autocorrelation might be a problem if trends within a single time series are analyzed. Autocorrelation is expected to be an issue for the temperature field, which is relatively homogeneous, but will not be an issue with the relative humidity field, which is subject to greater tempo-spatial variability.

To further elaborate on the question about the impact of spatial averaging of IAGOS data and potential autocorrelation we prepared the two plots below. The plot at the top shows IAGOS data at the original 4-s resolution and the lower plot shows temporally smoothed data (60 s). The obtained distributions change little between the original and the smoothed data and, thus, we argue that averaging IAGOS only has a second-order impact on the obtained statistics and subsequent analysis.

4s resolution IAGOS



1 min resolution



6.) L 153/154: I suggest to delete this statement since the procedure is better explained below from L 173 on. Anyway, I think the argument is weak (I don't remember Schumann's reasons for this) and refers rather to interpolation of specific humidity than relative humidity.

Interpolation of relative humidity is not straightforward as it depends on the underlying temperature and absolute humidity fields, and is determined on the basis of the exponential Clausius–Clapeyron relationship. Due to the nature of the Clausius–Clapeyron equation, linear interpolation, for example, leads to incorrect values of relative humidity. We rephrased the paragraph and removed the reference to Schumann et al. (2012). The end of the paragraph reads as follows:

“Spatial and temporal interpolation of relative humidity is not done because the relative humidity depends on the underlying temperature and absolute humidity field, which are both related through the Clausius–Clapeyron relationship.. Due to the exponential nature of the Clausius–Clapeyron equation, linear interpolation, for example, would lead to incorrect values of relative humidity.”

7.) L 210/211: The statement "The CDFs describe the probability that a certain quantity, for example temperature or relative humidity, exists in the underlying data set" is wrong. I think, we know which quatities are in the data sets, therefore the probability is either zero or one. Please consult a textbook on probability and correct the sentence or leave it out. The CDFs describe the probability that a certain value of a quantity, for example temperature or relative humidity, exists in the underlying data set.

We agree with the Reviewer that the sentence was poorly phrased. The sentence has been rephrased to the following:

“*The CDFs describe the probability that the value of a quantity (or random variable) X, for example temperature or relative humidity, has a value that is lesser or equal to x.”*

8.) L 239/240: You should delete the second part of the sentence. Contrail formation takes place a few tenths of a second after exhaust. It has nothing at all to do with the vortex phase which starts at, perhaps 20 seconds. I also dislike the next sentence. The SAC has been tested on many flights long ago, and it works excellently. There is a figure in the 1999 IPCC report that shows this (I think, the figure has been taken from a paper by Kärcher).

The Reviewer is right and we rephrased the sentence to be more precise.

“The SAc is based solely on thermodynamic principles and has been tested to be a valid approximation although it does not inform on the fate of the contrail, which is a more complicated function of the ambient conditions but also the interactions of the vortex phase with the environment.”

9.) Section 2.4: I miss information on your choice of an overall propulsion efficiency.

This is correct and the propulsion / engine efficient is now mentioned in the text.

“Calculations are performed for kerosene with a fuel specific energy Q = 43.2 MJ kg^{-1} and an emission index of water EIH20 = 1.25. The overall engine efficiency η is set to a typical value of 0.3 (Rap et al., 2010).”

10.) L 272: is there really a general decrease of r_ice with p? Why then occurs ISS mostly directly below the tropopause?

In general, absolute humidity decreases with altitude (e.g. Kiemle et al., 2012; Kaufmann et al., 2018, Kruegger et al 2022). However, local maxima of relative humidity can occur just below the tropopause, e.g., in regions with strong vertical updraft or advection of humidity, which then leads to cloud formation. We followed the suggestion of the Reviewer and rephrased the sentence as following:

“With the general decrease in absolute humidity and possible intrusion of dry air from the stratosphere, the first mode becomes more and more pronounced with decreasing p, while the second mode flattens and almost vanishes.”

11.) L 297/298: it is not clear to me why the T22 correction cannot modify the shape of the pdf. Please explain.

The sentence has been rephrased and is now less definitive. We intended to say that scaling all humidity values above a certain value by a constant increases the values as a whole and shifts the peak of the distribution to higher relative humidity values.

“Furthermore, differences in the second mode in relation to the IAGOS observations remain as the T22-correction only scales values above a certain threshold, which primarily shifts the bulk of data points from 100 % to higher r_{ice}.“

12.) L 370 ff: To my view the discussion of the differences between the QM and T22 is not convincing. For instance, the quoted thresholds from CoCiP are very very low, so they are not really a constraint. To me, the first question is, whether these differences are statistically robust. If a subset of four out of the five years is used, how large is the change of the quoted values? If it is much smaller than the difference between QM and T22, then a real difference seems more plausible. Second, it might be that you compare apples with oranges, i.e. for instance contrail distance with relative frequency of occurrence.

To answer the first question of the Reviewer, we want to make clear that we gave the estimate from Teoh et al. (2020) to provide a reference for the reader and to set our results in the context of existing literature. The differences between the results using our approach and the approach by Teoh et al. (2020) are clearly stated in the manuscript. Differences in approaches can partly explain the differences. We do agree with the Reviewer that the absolute numbers given in our manuscript and in Teoh et al. (2022) might be subject to spatial and temporal effects. At the same time the differences between our calculations of potential contrail occurrence and the calculations from Teoh et al. (2022) must not be overinterpreted. They only indicate the uncertainty due to different approaches.

Regarding the second question of the Reviewer, we think that the comparison is valid as the study by Teoh et al. (2020) entitled “Mitigating the Climate Forcing of Aircraft Contrails by Small-Scale Diversions and Technology Adoption” provides the percentage of flights that form contrails rather than the fraction of time it produces contrail. We direct the Reviewer to Table 1 in Teoh et al. (2020).

13.) L 408 ff: It is not clear to me how you interpret the SAc. Why does an rcrit occur? As far as I understand it, the SAc gives a threshold temperature which is the maximum T where a contrail could in principle form (at 100% RH). Is rcrit the RHi values along the tangential mixing line between the T threshold (rcrit=100% RH) and the lower temperature below which contrails are always formed (rcrit=0%). This is not completely clear to me.

The critical threshold r_{crit} is always 100 % (w.r.t liquid water). This was incorrectly formulated in the last version of the manuscript. In course of the manuscript revision the sentence in line 408 was also removed.

14.) Figure 7 and corresponding text: This text contains a lot of details and numbers that I find not necessary. What is the take-away message that the reader learns from these details? Again, how robust are these values, if you leave one your out, then another one, etc.?

Following the suggestion by the Reviewer the text has been shortened by removing unnecessary sections. Figure 7a is discussed in detail and the influence of the QM correction (Fig.7c) is presented. We focused on Fig.7 a and c as both are essential to see if the improved correction has an impact on the estimated contrail formation conditions. Furthermore, the discussion of the differences between the two methods is used to explain the rationale behind Fig. 7. Figure 7b and d are only briefly mentioned but not discussed in detail. We further highlight that Fig. 7 and the given numbers should be interpreted in a qualitative and not quantitative way, as they can change slightly between years. The section now ends with a summary (take-away message) that states the robustness of the ERA5 estimated PC occurrence with respect to the temperature and humidity correction. The minor impact of the corrections on the potential PC occurrence implies that ERA5 is suited to estimate regions that are prone for contrail formation. We would like to direct the Reviewer to the revised manuscript to see the changes.

15.) L 482: "Applying ... more correctly DETECTED...", please reformulate. The QM method does not change any detection.

The sentence has been rephrased and now reads as:

“The application of the QM-correction modifies the distributions of temperature and relative humidity in such a way that locations of PC are detected more often where they are supposed to occur and were missing before, resulting in an increase in ETS from 0.27 to 0.36.”

16.) L 493-506: This is an interesting consideration. But it comes too late. Perhaps a lot of not-so-important numbers could be saved if this consideration would be presented earlier.

We agree with the Reviewer that this is an important point to mention. Therefore, the paragraph has been moved earlier in the section. The paragraph, which mentions the potential mismatches and that investigates the 3h time resolution, is now used to explain the baseline and sensitivity of the scores on the temporal-spatial misalignment. This provides a reference for the scores that have been calculated on the different correction methods. Differences that exceed the baseline are thus truly attributable to incorrect values of temperature and relative humidity in ERA5 or the corrections, and are therefore of relevance.
We also followed the Reviewer’s major comment 2 and calculated the ETS. The ETS is provided in Table 5 and we also added a paragraph to the section, where we briefly mention the impact of the corrections on the ETS.
We would like to direct the Reviewer to the revised manuscript to see the changes in Section “Analysis of collocated contrail formation potential from ERA5 and IAGOS”.

17.) Section 3.5 and Figure 9: please state whether you use for the analysis original or QM corrected ERA5 data.

The section and the figure are based on QM-corrected ERA5 values as we intended to evaluate the remaining errors after the correction that are associated with a certain classification. The following sentences have been added at the beginning of the section. The caption of the Figure has also been modified.

“Subsequently, we aim to quantify the mean differences in temperature and relative humidity that remained after the QM correction and that contribute to the misclassification of potential contrail formation. Within the following section all ERA5 values are QM-corrected.”

18.) L 536-539: If that happens only on 225 hPa, is it then really likely that small scale variations cause these differences? Don't such variations occur on the neighbouring pressure levels as well? In which sense are T and rice "more homogeneous" on 200 hPa? What does that mean and from where do you derive such a conclusion?

Thank you for pointing this out. The text has been revised and rephrased as follows:

“Least frequent are the misclassifications ’NoC–PC’ (light blue, 0.3 %) and PC–NoC (dark blue, 0.5 %). These two groups are subject to the largest ΔT and Δrice. Samples in these categories were only found at the 250 and 225 hPa p-level, while the PC–NoC (dark blue) is not found at the 200 hPa level. It is likely that data points in the two categories result from small scale variations captured by IAGOS that are not represented by ERA5 due to temporal and spatial resolution.”

19.) L 539/540: Isn't this a trivial statement? Did you expect something else? Perhaps there are more interesting results in this section that could be more elaborated. For instance, which error usually dominates, is it the error in T or the error in rice. It would also be interesting to see how the differences in absolute humidity contribute to the misclassifications. What is here the conclusion? What needs to be fixed in the ECMWF model more urgently, is it T or the water vapour field? What is the effect of the QM correction in this analysis? I suggest you add a second set of points (squares or empty points), perhaps with arrows, to show how the QM correction shrinks the errors. Please think also on getting rid of the "default" class in the figure. It seems as if a couple of dots are hidden behind the big black dot in the middle.

We agree with the Reviewer that the sentence is trivial, but we meant to express that it is worth identifying whether the misclassification in ERA5 with respect to IAGOS is most often due to biases in temperature or in humidity. The sentence and paragraph have been rephrased and extended in this regard. The paragraph now reads as follows:

“It is worth identifying whether the misclassification in ERA5 with respect to IAGOS is most often due to biases in temperature or in humidity. Focusing on the PC representation in ERA5, the primary reasons for a misclassification after the correction is the deviation in rice. This is visualized by the proximity of the violet and green dots to the y-axis (small ΔT), while the differences in rice are larger than ±20 %. Hence, the underestimation (green dot) or overestimation (violet dot) of potential contrail formation is primary related to the underlying humidity field in ERA5.”

20.) L 575: "So from a statistical perspective, that the original ERA5 model output is able to adequately represent contrail formation." First, the sentence is gramatically incorrect. Second, please explain what means that contrail formation is adequately represented FROM A STATISTICAL PERSPECTIVE? Assume, the model would only predict ISS and PC over Antarctica and nowhere else, but with the correct frequency of occurrence averaged globally, then one could also state that the formation is represented adequately from a statistical perspective. So my question is, whether your statement makes sense.

In the course of revising the manuscript this sentence has been removed.

21.) L 590/591: Isn't it the other way, that is, the special bias combination leads to the corresponding entry in the contingency table?

Following the suggestion of the Reviewer the sentence has been removed as it is a trivial statement. The sentences that follow in the paragraph are sufficient.

22.) Section 4: should not be entitled "Summary and discussion", since it does not contain any discussion.

“Discussion” has been removed from the subtitle.

Miscellaneous

23.) Line 3: upper tropospheric

The sentence has been rephrased as follows*:*

“The skill of the atmospheric reanalysis ERA5 from the European Centre for Medium-Range Weather Forecasts (ECMWF) at simulating temperature and relative humidity in the upper troposphere and lower stratosphere is assessed by using five years of In-service Aircraft for a Global Observing System (IAGOS) observations.”

24.) L 19: underlying

The typo has been corrected.

25.) L 30: delete "bonds"

“Bonds” was deleted.

26.) L 31: What is WC?

This was a typo and is now corrected to “WV”, which is defined on the line before.

27.) L 41-42: The word "defined" is too strong. In fact, this was never defined. There is only a vague understanding that contrails that survive the vortex phase are somehow persistent.

The sentence was rephrased as follows:

“For a contrail to be persistent (with the common meaning that is has a lifetime longer than 10 minutes), the ambient air has to fulfill the SAc and must also be supersaturated with respect to ice.”

28.) L 67: "due to the high temporal and spatial distribution of WV" ???

The sentence was rephrased as follows:

“Slightly less accurate is the prediction and re-analysis of relative humidity, which is generally challenging due to the high temporal and spatial variability of WV.”

29.)L 73: What is "contrail estimation"?

To be clearer the sentence has been rephrased as follows:

“To mitigate the dry bias under conditions close to ice-supersaturation in ERA-interim and ERA5, studies have applied either multiplication factors (Schumann et al., 2013, Schumann et al., 2015) or parameterized corrections (Teoh et al., 2022).”

30.) L 126: Replave "multiple" by "many". I believe "multiple years" is nonsense.

Multiple has been replaced by many.

31.) L 199: The expression "minimizing the C... test" is confusing. How can a test be minimized. I think this sentence can be dropped since it does not provide essential information.

We meant that the test statistic was minimized. However, we followed the Reviewer’s suggestion and removed the sentence as details about how the parameters were detected can be found in Teoh et al. (2022).

32.) L 241: criterion (singular).

Thank you for pointing this out. The word has been corrected.

33.) L 244: slow

The typo was corrected.

34.) L 335: r_ice IS

The typo was corrected.

35.) L 433: remove one instance of correction

The word was removed.

36.) Table 6: the caption says "TO07" while the table headline says "T07"

Based on the suggestion from the Reviewer we use the equitable threat score to compare the measurements. Doing so removed the reference to Tompkins.

37.) L 513: delete comma after ERA5

The comma was removed.

38.) L 658: require

The typo was corrected.

39.) L 671: showS and remainS

Both words have been corrected.

40.) Table A1: eta should be labelled "overall propulsion efficiency". It is not necessary (and not good) to introduce new notions.

We agree with the Reviewer and added “overall propulsion efficiency” in the text and unified it in the table.

41.) Figure A1(e): The y-label contains [*2e6]. Please check.

The label is correct and scales the y-axis, which shows the number of individual IAGOS measurements (every 4 seconds) within a given month.

References

Appleman, H.: **The formation of exhaust condensation trails by jet aircraft**, *B. Am. Meteor. Soc.*, 34, 14–20, 1953.

Busen, R. and Schumann, U.: **Visible contrail formation from fuels with different sulfur contents**, *Geophys. Res. Lett.*, 22, 1357–1360, 1995

Heymsfield, A., Baumgardner, D., DeMott, P., Forster, P., Gierens, K., and Kärcher, B.: **Contrail microphysics**, *B. Am. Meteor. Soc*., 91, 465–472, 2010.

Jensen, E., Toon, O., Kinne, S., Sachse, G., Anderson, B., Chan, K., Twohy, C., Gandrud, B., Heymsfield, A., and Miake-Lye R.: **Environmental conditions required for contrail formation and persistence**, *J. Geophys. Res*., 103, 3929–3936, 1998

Kärcher, B., Peter, T., Biermann, U. M., and Schumann, U.: **The initial composition of jet condensation trails**, *J. Atmos. Sci*., 53, 3066–3083, 1996.

Kaufmann, S. / Voigt, C. / Heller, R. / Jurkat-Witschas, T. / Krämer, M. / Rolf, C. / Zöger, M. / Giez, A. / Buchholz, B. / Ebert, V. / Thornberry, T. / Schumann, U. , **Intercomparison of midlatitude tropospheric and lower-stratospheric water vapor measurements and comparison to ECMWF humidity data**, 2018, *Atmos. Chem. Phys*. , Vol. 18, No. 22, p. 16729-16745

Kiemle, C. / Schäfler, A. / Voigt, C. , **Detection and analysis of water vapor transport.**

2012, *Atmospheric Physics: Background -- Methods – Trends,* Springer Berlin Heidelberg: Berlin, Heidelberg, p. 169-184

Krüger, K. / Schäfler, A. / Wirth, M. / Weissmann, M. / Craig, G. C., **Vertical structure of the lower-stratospheric moist bias in the ERA5 reanalysis and its connection to mixing processes**, 2022, *Atmos. Chem. Phys.* , Vol. 22, No. 23, p. 15559-15577

Penner, J. E., Lister, D. H., Griggs, D. J., Dokken, D. J., and McFarland, M. (Eds.): **Aviation and the global atmosphere,** *Cambridge University Press*, Cambridge, UK, 1999

Schmidt, E.: **Die Entstehung von Eisnebel aus den Auspuffgasen von Flugmotoren**, Schriften der Deutschen Akademie der Luftfahrtforschung, 44, 1–15, 1941.

Schumann, U., Ström, J., Busen, R., Baumann, R., Gierens, K., Krautstrunk, M., Schröder, F. P., and Stingl, J.: **In situ observations of particles in jet aircraft exhausts and contrails for different sulfur-containing fuels**, *J. Geophys. Res.*, 101, 6853–6869, 1996.

Schumann, U., Arnold, F., Busen, R., Curtius, J., Kärcher, B., Kiendler, A., Petzold, A., Schlager, H., Schröder, F., and Wohlfrom, K.-H.: **Influence of fuel sulfur on the composition of aircraft exhaust plumes: The experiments SULFUR 1–**7, *J. Geophys. Res.*, 107, 4247, doi:10.1029/2001JD000813, 2002.

Schumann, U.: **Formation, properties and climate effects of contrails,** *C. R. Phys*., 6, 549–565, 2005.

Teoh, R. / Schumann, U. / Majumdar, A. / Stettler, M. E. J., **Mitigating the climate forcing of aircraft contrails by small-scale diversions and technology adoption**, 2020, *Environ. Sci. Technol.* , Vol. 54, No. 5, p. 2941-2950

Teoh, R., Schumann, U., Gryspeerdt, E., Shapiro, M., Molloy, J., Koudis, G., Voigt, C., and Stettler, M. E. J.: **Aviation contrail climate effects in the North Atlantic from 2016 to 2021**, *Atmos. Chem. Phys*., 22, 10 919–10 935, 10.5194/acp-22-10919-2022, 2022.