Reply to Reviewer #2

We would like to thank the reviewer for their constructive comments, which have helped improve the paper significantly. We agree to shorten the paper to focus more on the new results, particularly those for the tropics. While reviewing the paper, we found an error in the analysis tool for the vertical distribution of the mean temperature was found and corrected (Fig. 7 in the preprint), i.e., the mean temperature was plotted with respect to the dynamic tropopause instead of the thermal tropopause. This only affected the interpretation of the results in Sect. 3.1, for which the corresponding findings has been corrected. However, it does not change the overall conclusion on the temperature comparison between ERA5 and IAGOS.

In the following, we reply to the individual comments from the reviewer. We use colour to organize this as follows:

- The questions and comments from the reviewers are marked in blue.
- The replies from the authors are written in black.
- Any changes to the manuscript are written in green.

Major Comments:

Reviewer: This study demonstrates a significant investment of time and effort, application of recognized methods, and extensive knowledge of IAGOS and ERA5, making the foundation of the study solid. However, the analysis itself is overly descriptive. Often minor differences between the old and new results are shown, regardless of their significance. Therefore, I recommend shortening the text and focusing more on relevant aspects.

Authors: Thank you for the feedback. The comparison between ERA5 and IAGOS indeed focuses a lot on confirming earlier studies, which is a way for us to evaluate our analysis approach. Nevertheless, we agree that the paper is perhaps too long. Following the suggestions, we have shortened the paper by focusing more on novel results, i.e. those from the tropics and removing comparisons to literature in the results section. Comparisons to literature are now mainly found in a newly added discussion section (Sect. 4). Furthermore, we have removed Sect. 2.5 on the North Atlantic weather pattern methodology and Sect. 3.6 on the analyses of the North Atlantic weather patterns based on other feedback received in the review process. These sections are now found in Sect. S7 in the supplement. We suggest it as future research in the discussion section. We have also removed Fig. 5 from Sect. 2.4, which showed the global mean annual cirrus occurrence, and Fig. 8 from Sect. 3.2, that showed a 2D histogram of the ERA5 RHi as a function of IAGOS RHi; we found that their key messages were mainly confirming previous results.

Reviewer: I don't think it makes much sense to calculate the ETS for many different situations when the poor agreement between the ISSRs found by ERA5 and IAGOS is due to the adjustment scheme under cloudy conditions. I would recommend repeating this part of the analysis by using a threshold of ERA5 RHi of 90 % or 95 % to determine ISSRs (at least under cloudy conditions).

Authors: We agree that repeating this analysis using different thresholds of ERA5 RHi to determine ISSRs is interesting. However, we find that only showing how the ISSR fraction changes with different thresholds might not give a good idea of whether different RHi thresholds improves the ability of ERA5 to predict ISSRs. This is because of how the ISSR fraction is calculated; we sum the number of counts where either IAGOS or ERA5 finds RHi > threshold and divide by the total number of measurements within the conditions considered (i.e. region, atmospheric layer, and in-cloud conditions). Thus, it does not consider if ERA5 improves at those specific points where it was either not predicting an ISSR or incorrectly predicting an ISSR. This can be considered with the ETS. Hence, we have replaced the analysis in Sect. 3.5 with how the ETS changes with different RHi thresholds under cloudy and clear-sky conditions.

Reviewer: I also have to question the calculation of ETS in the context of discussing different weather patterns. Before discussing the differences in ETS, it would be helpful to examine the fraction of ISSRs for these weather patterns as observed by IAGOS, and determine if there are any differences among them.

Authors: We agree that examining the fraction of ISSRs for each weather pattern as observed by IAGOS and determine differences among them could be helpful. Now that we have removed Section 3.6 and included it as a discussion point, we have decided to include these figures in the supplement (see Sect. S7).

Reviewer: It is an interesting feature that in South Asia you find the minimum temperature 60 hPa below the thermal tropopause. The structure of the lapse rate and cold point tropopause in the tropics can be quite complicated but, normally, if there are multiple tropopauses the cold point tropopause is above the WMO thermal tropopause (also in Muhsin et al., 2018?). I suggest performing a literature review to learn more about this. See for example: https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2011JD016637

Authors: Thank you for this comment and for the references. An error in the plotting tool for the vertical profile of mean temperature was noticed, where it was plotting the vertical profile with respect to the dynamic tropopause instead of the thermal tropopause (Fig. 7 in the preprint). It only affected the interpretation of the results in Sect. 3.1, for which the corresponding findings has been corrected. This has now been corrected to the thermal tropopause. The minima of temperatures in the tropic regions of Southern Trans-Pacific and Africa now appear more distinct. The minimum temperature in South Asia for JJA is no longer visible after fixing the error in the analysis tool. However, it does not change the overall conclusion that ERA5 shows a cold bias in temperature when compared to IAGOS. For example, we still find ERA5 to have a cold bias of 0.5 K on average in the extratropical upper troposphere.

Reviewer: Line 334 and 516: "Therefore, the moist bias appears to arise as a result of the saturation adjustment, which cannot take into account that the (wet) mode of the RHi PDF can be subsaturated in some regions." In my opinion this should not lead to a moist bias but to an alignment, because the (false) saturation adjustment would just not happen. Please explain.

Authors: We appreciate this feedback. We meant to say that it appears that ERA5 cannot consider that the wet mode can be subsaturated in cloudy conditions; under subsaturated conditions, ERA5 still showed a mean value of 100%, which is the value of the saturation adjustment. However, we do not know for certain why the subsaturation is not well estimated. We have rephrased the sentence to:

"This is because the mean RHi in IAGOS for this particular region and season is less than 100% in cloudy and indeterminate conditions, which aligns with the behaviour of the RHi PDF, whereas ERA5 shows a mean of 100%. Therefore, the moist bias appears to arise because ERA5 cannot take into account that the wet mode in cloudy conditions can be subsaturated in some regions. Whether the subsaturation is a result of the specific weather conditions in JJA for South Asia is uncertain and requires further evaluation."

Specific comments:

Reviewer: Line 123: In IAGOS RHL > 100 % is flagged invalid.
Authors: Thank you for pointing this out, this is a useful information.

Reviewer: Line 131: Why does sampling every minute result in 2.5% of all IAGOS measurements?
Authors: We can see that the structure of the paragraph was misleading. Indeed, it is not the sampling that results in 2.5% of all IAGOS measurements. We meant to present the percentage of measurements left after applying the criterion from Table 2 and the sampling together, but we realize the value presented is also wrong. This paragraph has been restructured, and the value has been updated. See the new paragraph below:
"Lastly, IAGOS records the measurements every four seconds. To avoid autocorrelation affecting our analysis, it is chosen to sample a measurement approximately every minute. This is done using a uniform random number generator ranging from 1 to the maximum number of measurement points from IAGOS, to avoid systematic bias. The sampling and application of the criterion from Table 2 results in using 3.8% of all IAGOS measurements between 01/07/2011 and 31/12/2022."

Reviewer: Line 182: You find a higher frequency of indeterminate and cloudy conditions compared to IAGOS because of the limitations of the BCP (see Petzold et al., 2017).
Authors: We agree, thank you for the confirmation.

Reviewer: Line 202: Yes, physically, relative humidity partially depends on temperature, but the ICH sensor from IAGOS measures RHL, so, this is the water variable which is independent from the temperature.
Authors: We have removed the sentence "This is important to consider, as relative humidity partially depends on temperature (Reutter et al., 2020)". However, we still consider temperature an important aspect to compare as it directly influences the selection of points considered for ice supersaturation, as mentioned in Section 3.2.

Reviewer: Line 205: Please specify how big the ERA 5 cold bias in the extratropics is.
Authors: We have specified how large the ERA5 cold bias is in the extra tropics.

Reviewer: Line 212: "extratropical seasonal cycle": in the Northern Hemisphere
Authors: This change has been incorporated.

Reviewer: Line 250: it does not provide good quality results in dry conditions in the lower stratosphere due to the loss of sensitivity as a result of the adiabatic compression effect (Konjari et al., 2025)
Authors: Thank you for this clarification. We have adapted the sentence to include "in the lower stratosphere".

Reviewer: Line 302: Please rephrase: "The tropopause layer is not completely dry (Petzold et al., 2020; Reutter et al., 2020)"
Petzold et al. state: "a tropopause layer characterized by mean RHice of 60% almost independent of the season..."
Reutter et al. state: "In the tropopause layer, still a significant amount of the data is exceeding values of RHi > 100 %, both in the in situ data as well as in the ERA data set."

Authors: We agree that this statement was not clear. At the same time, we have restructured this section to shorten the paper, but we still implement one of the suggestions. The suggestion has been included as follows:
"The UT has more moisture compared to the TROP and LS, allowing for higher probabilities of ISS occurrence. However, the TROP still shows high ISS occurrence, both in the tropics and extratropics. This is in line with Reutter et al. (2020), where it was found that there was a significant amount of in-situ and ERA-interim data that exceeded the RHi = 100% threshold near the tropopause in the North Atlantic flight corridor."

Reviewer: Line 314: "For clear-sky conditions, there is a smaller probability of low RHi, with increased probability of higher RHi, and a more equal probability across all observed RHi values." I don't understand this sentence.
Authors: We have rephrased this sentence to, while at the same time combining the analysis of the UT and TROP:
"For UT and TROP clear-sky conditions, the probability of RHi > 100% is higher, and there is a more uniform probability across the observed RHi range compared to the LS."

Reviewer: Line 341: typo: "ERA5 may predict less ISSRs compared to ERA5"
Authors: Thank you for pointing this out. With shortening this paper, this sentence has been removed.

Reviewer: Line 364: Delete the word "but" in: "On the other hand, Reutter et al. (2020) found a maximum of 40% using IAGOS when considering all seasons, but with the maximum also occurring 30 hPa below the tropopause and a reduction of the ISSR fraction to zero in the lower stratosphere.
Authors: We have restructured this paragraph, for which the 'but' has also been removed.

Reviewer: Line 367: "ISSR fraction is highly sensitive to…": Maybe better without the word "highly"
Authors: We agree, and we have removed the word 'highly' from this sentence.

Reviewer: Line 402: Could the reason for higher ETS in North America be, that there are more data assimilated for ERA5?
Authors: Thank you for asking this question. We have looked at the WMO WDQMS webtool to get an idea. Looking at this tool, for upper-air land observations from Near-real-time NWP monitoring of the Global Observing System networks, it does appear that more data is assimilated in North America compared to the North Atlantic and the tropical regions defined. However, it is hard to make that distinction between North America, Europe and North Asia. Hence, we do not think we can appropriately make this conclusion; we would require more accurate knowledge of how much data is assimilated for ERA5 in each region.
We contacted an expert focusing on data assimilation and differences between Europe and the United States, who mentioned that it is also a difficult question. The expert mentioned that there is more aircraft data over the United States compared to Europe and Asia. Furthermore, ECMWF also uses more METAR data in the US compared to Europe. Hence, the reason for a higher ETS in North America could be that more data is assimilated for ERA5, but we do not have enough information to guarantee this.

Reviewer: Line 407: Please rephrase: "the maximum difference between seasons is approximately 10%, at most. In North America, we find that the variation can be up to 20%."
Authors: Due to the restructuring and shortening of the paper, this sentence has been removed.

Reviewer: Line 410: “tend to find”
Authors: This change has been incorporated.