Comments on “Predicting Ice Supersaturation for Contrail Avoidance: Ensemble Forecasting using ICON with Two-Moment Ice Microphysics” by Maleen Hanst et al.
https://doi.org/10.5194/egusphere-2025-3312

The authors have done important work. While earlier studies attempted to apply correction methods to better represent ice-supersaturated regions (ISSRs), this study aims to implement an improved two-moment ice microphysics scheme within the ICON model framework.

An improved representation of supersaturation in models is essential for accurately forecasting cirrus and, in particular, the potential for contrail formation. Furthermore, a better representation of ice supersaturation and contrail estimation is fundamental for estimating the radiative forcing of cirrus and contrails, enabling flight re-routing and more realistic modeling of the water budget. The proposed two-moment scheme is an important step forward in the development of the ICON model.

In addition to the modified two-moment scheme, this study estimates the benefits of ensemble member simulations for ISSR prediction and uncertainty estimation. Simulated temperature and relative humidity from the operational ICON one-moment ice microphysics scheme and the modified two-moment schemes are compared with radiosonde measurements and, to a lesser extent, IAGOS measurements. The comparison shows that ICON's representation of ice supersaturation improves significantly, when using the two-moment schemes.

To make the results more accessible to a broader community, the structure of the text needs improvement. Please try to get quicker to the point. Restructuring the entire text is also necessary to avoid repetition and establish a logical structure. The structure of a paper should guide readers step by step through the topic and help them understand the subject matter. Please see below for more details.

A possible structure for the beginning of the manuscript could be:

1. Introduction
2. Data
2.1 Model
2.1.1 Operational ICON 1-Mom
2.1.2 Modified ICON 2-Mom and ICON 2-MOM EPS
2.2 Measurements
2.2.1 Radiosonde
    - processing
    - data extraction
2.2.2 IAGOS
    - resolution, uncertainty
    - processing and extraction

3. Methods
3.1 processing of model data
3.2 Validation scores
.....
This structure is only a suggestion, and I will leave it up to the authors to decide how to restructure their text.

A concern along the same line. The paper includes ten different metrics, some of which are not well introduced. You may consider focusing on fewer and the most important metrics. For details please see my major and minor comments. There are also many inconsistent abbreviations and mathematical formulas that must be homogenized to avoid ambiguity.

Most of my comments are suggestions, and I hope the authors find them helpful and do not misinterpret them.

Major comments:

- The text can be shortened and made more concise by removing some parts or simply by restructuring paragraphs. For example, L197-200: "Again, the humidity measurement technology used here combines humidity and temperature sensing elements. In more detail, it consists of a capacitive relative humidity sensor (Humicap-H, Vaisala, Finland) and a platinum resistance sensor (PT100) for the measurement of the temperature at the humidity sensing surface. " This could be shortened:
  "The temperature is measured using a platinum resistance sensor (PT100) and rel. humidity is measured using capacitive relative humidity sensor (Humicap-H, Vaisala, Finland)."
  I do understand that everyone has their own writing style, but since the paper is already quite long, writing concisely will make the content more accessible to potential readers.

  Another example is the repeated mentioning of information, like "reducing the number of ensemble members". This point is first introduced in L122, and then revisited in L146–147 and L166–168.
- Since you are introducing an improved version of the two-moment scheme you may spend a little more time one the actual improvement. Would it be worth to shift it from Appendix A to the main text?
- The analysis makes use of ten different "metrics": confusion matrix, FBI, POD, FPR, precision, MCC, ROC, k-out-of-10, Youden index, and F1. You may want to consider reducing the number of metrics and focusing on the most important ones that can be used throughout the paper. If you decide to keep them all, they must be introduced well, either in the methods section (preferred) or in the text when they are first used. Please provide the possible technical ranges of the indices/metrics, common values, and the desired value so that the reader can interpret the metrics.
- The discussion session reads more like a summary. If introducing a dedicated discussion session, then the results of the analysis should discussed by setting them into context with existing literature. For example to discuss differences or similarities in distributions of ISSR and to explain the causes for potential differences.

Minor comments:

- The manuscript focuses specifically on contrail formation. However, the formation of cirrus clouds and supersaturation is also relevant to the water budget in the upper troposphere and

lower stratosphere. You could mention this as an additional motivation for improving the representation of supersaturation.

- L40: Could you provide a reference for interested readers?
- L41: NWP was already introduced in Line 35
- L43: RH_ice was already introduced in Line 35
- L41-47: These lines contain redundant information by mentioning the uncertainty and lack of in situ observations multiple times. Please revise accordingly.
- L52: Maybe "...and returns adjusted values of RH_ice?" instead of "outputs".
- L52: "Wang et al. (2025)...." Please note that ERA5 is suspected to be biased in terms of relative humidity, which causes problems in resolving ISSR. This is due to the fact that RH_ice is clipped to a maximum value, as well as the spatial resolution of the model grid. However, there is no consensus on whether RH_ice is too low or too high at the tropopause level. Several correction methods for ERA5 have been provided, e.g.,

  1) Schumann, U. and Graf, K.: Aviation-induced cirrus and radiation changes at diurnal timescales, J. Geophys. Res.-Atmos., 118, 2404–2421, https://doi.org/10.1002/jgrd.50184, 2013.

  2) Schumann, U., Penner, J. E., Chen, Y., Zhou, C., and Graf, K.: Dehydration effects from contrails in a coupled contrail--climate model, Atmos. Chem. Phys., 15, 11179–11199, https://doi.org/10.5194/acp-15-11179-2015, 2015.

  3) 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, 10919–10935, https://doi.org/10.5194/acp-22-10919-2022, 2022a
- L54: At the end of the sentence: "...when validated against test data"?
- L84: NWP already introduced
- L92: Throughout the manuscript, there are multiple versions of RH_ice. There is: RHice and RHice. Figures use RHi. All of these combinations are sometimes in italics and sometimes not. Choose one version and stick to it throughout the script.
- L96: What are the typical challenges? They are known to you, but since you are making this point, please briefly mention them. Or are they given in the next sentences? If so, I would suggest to write, ... with forecast applications, such as data assimilation and model uncertainties, the interpretation of the resulting forecast..."
- L126: abbreviate "relative humidity of ice"?
- L140: and elsewhere. To facilitate understanding of the paper, clearly define your different model setups and provide unique abbreviations for each configuration. ICON (for the operational 1-Moment scheme), ICON 2-Mom (for the new 2-Moment scheme), and ICON EPS 2-Mom. Then, stick to these abbreviations. The current version uses various

abbreviations and paraphrases. Sometimes, the 2-Mom model is also called the deterministic model. These variations make it difficult to understand the paper's content and cause unnecessary ambiguities.

- L164 and elsewhere: For ranges, e.g., 8.5-12.5, use spaced en-dashes "--" see ACP style guide
- 169 - 170: Why is ICON 2-Mom now cursive? Please see comment above.
- L184: What does "TEMP BUFR" mean? Enter the full name here. Explaining BUFR in the next sentence is insufficient and too late.
- L185: It is usually first the long name, followed by the abbreviation in brackets.
- Fig2: The subpanels and labels are small but still legible. Would it not make sense to split the plot and place the individual plots in the positions where they are discussed? This would also prevent flipping back and forth through multiple pages when referencing back from Section 4. The authors may want to consider this.
- L209: Please explain which model you mean with "dedicated model". ICON 1-Mom, 2-Mom?
- L213: Call it ICON 2-mom?
- L215: Why ICON 1-mom now in italics?
- L224: Is the new paragraph required?
- L224: Is "density tail" the correct term? I have not found it in the literature. Would you call it the "tail of the density distribution"?
- L224: "operational system"? you mean ICON 1-mom?
- L230: "The ICON grid employed features a horizon..." Please check the grammar of this sentence.
- L232-233: Why do you use the closest match to the radiosonde station instead of the actual position of the radiosonde? Radiosondes drift horizontally by several kilometers and may end up in a different grid cell.
- Sec 4.1.1 and 4.1.2: What are the fundamental differences between these subsections? Would it not be better to start with 4.1.2 and explain how the data is extracted, and then analyze and compare the data? Additionally, both could be done in one subsection ending in a less fragmented text.
- L241: "An ICON spin up time of a minimum of 6 hours was required" This does not fit here. Would it be better placed in the introduction of the ICON model?
- L247: "...simple scatter plot." Please explicitly mention which plot you are referring to so that the reader can more easily identify it. In my opinion, showing it at the end of the next sentence is too late. I found myself wondering where to look.
- L262: Why is the first part in brackets and the rest in the subscript?
- L265: You may state the possible range of FBI and that a value close to 1 would be desirable.

- L277: "...the POD increases from about 0.4 for ICON 1-Mom..." I guess this is for the 100% threshold? You should mention that.
- L282: "..., also known as sensitivity,..." Does this refer to the probability of detection? If so, please introduce the term "sensitivity" when defining the POD.
- L232: "specificity" is not defined. Do you mean sensitivity?
- L301: "Another way..." if I'm not mistaken, no method has been mentioned so far to address the imbalance. At least, none has been mentioned explicitly. Please check and revise the script accordingly.
- L306 here and elsewhere: minus signs with $-$.
- L321-322: You may also mention that humidity varies greatly in the atmosphere, and radiosondes have a much higher spatial and temporal resolution than models.
- L324-327: What do you mean by "post-processing"? Please specify. As I understand it, you are proposing a new two-moment scheme that allows for supersaturation. Your previous analysis showed that the two-moment scheme performs better than the one-moment scheme. Does the new two-moment scheme just need more refinement or adjustment instead of additional post-processing steps? Alternatively, please explain what is meant by "post-processing."
- L329-330: " distinguish between ISSR and non-ISSR conditions (or higher supersaturation)," Should it be "non-ISSR and ISSR (or higher supersaturation) conditions"?
- L330-332: Why is non-ISSR not an option? As I understand it, the critical point is non-ISSR versus ISSR. At some point, the degree of supersaturation might no longer be relevant in determining whether a contrail can form, but rather, the duration that the ISSR and a potential contrail persist.
- Fig4 (b) has no title. Since you provided one for the inserted plot, you should also provide one for the main graph.
- L345: Would it be appropriate to start the following paragraphs with a new heading, such as: "Categorical Scores of ICON 2-Mom ENS". You mentioned investigating the continuous values first and then switching to the metrics. This would also keep it consistent with section 4.1. It's just a suggestion.
- Figure 5 goes over half of the page. You may want to reduce its size and incorporate some of the explanations into the text.
- Fig 7 : Please explain "ROC stratification". This has not been done in the text.
- L444: Earlier, you used a space "\," between the number and the percent sign. From the ACP Submission Guidelines: Spaces must be included between number and unit (e.g. 1 %, 1 m). Please revise the manuscript accordingly.
- L452-456: "When comparing [...] In conclusion, even when the model exhibits high confidence, as reflected by a low standard deviation, the histogram still displays intermediate supersaturation. This suggests that certain ISSRs can be well predicted."
  I have a difficult time following this line of reasoning. Please explain better and potentially

rewrite. Do you mean that the peak around RH_ice = 100 % is not fully resolved and only closely resembles the radiosonde observations?
Please define what you mean by "certain." Does this refer to cases with high, medium/intermediate, or low supersaturation? Which ISSR would be missed and which are well represented?

- L461: What is the ROC stratification approach? Stratification did not appear before, only in the caption of Fig. 7, which is insufficient.
- L467: I would avoid using the abbreviation "MTFs." It is only used one more time.
- L483: Again, it's just one paragraph in a new subsection. Couldn't it be discussed together with the radiosonde observations? A more important question: Do you need the comparison with the IAGOS data at all? What additional information does 4.2.4 provide compared to the radio-soundings? If you want to take advantage of the coverage over the Oceans and at flight levels, where most of the commercial aircraft are operating, then you should spend more time on this analysis and explain this. But subsection 4.2.4 is very brief and does not provide new conclusions
- L488: "...radiosonde data (Fig.8)" may suggest that the plot contains radiosonde data, but none is plotted. Perhaps rephrase. Alternatively, add the radiosonde data to the plot. This may make a visual comparison easier.
- L662: units not in italics
- In the entire reference section the DOIs are missing.