## Reply to Anonymous Referee #2

Referee comment on "Modeling and verifying ice supersaturated regions in the ARPEGE model for persistent contrail forecast" by S. Arriolabengoa et al. (egusphere-2025-1499, https://egusphere.co-pernicus.org/#RC2, 2025).

We thank the Referee for the time spent reviewing our manuscript and for the valuable comments, which have helped us to enhance the clarity of the paper and to provide a more thorough discussion of certain aspects. The answers to the various remarks are given as follows. For better legibility, the Referee's comments are highlighted with a gray background and changes in the manuscript are in italic.

## Review of the Manuscript

The manuscript presents a modification to the ARPEGE NWP cloud scheme to allow for the representation of ice supersaturated regions (ISSRs), with evaluation against aircraft in situ humidity data and comparisons to the Integrated Forecast System (IFS). The topic is timely and highly relevant to aviation climate impact mitigation strategies. The paper is generally well-motivated and provides a solid contribution toward improved ISSR representation. However, there are several issues that need to be further clarified or elaborated before publication.

### Major comments

1. While the study introduces a practical modification to ARPEGE's cloud scheme, the extent of its novelty compared to earlier approaches (e.g., in IFS or other global circulation models) is not entirely clear. The paper would benefit from a more explicit discussion of how the proposed approach differs from existing contrail cirrus cloud parameterizations used in current GCMs (e.g., ECHAM, CESM).

We agree with the general comment on the value of further discussion on how the proposed approach differs from existing parameterizations of condensation cirrus clouds. The first paragraph of Section 6.1 (L430) has been reworded and fully reorganized to include more precise remarks between similarities and differences concerning ARPEGE and IFS cloud schemes + other GCM models with ISSR representation capabilities.

A modified cloud scheme for ARPEGE NWP global model is presented in this work to enhance the representation of relative humidity w.r.t to ice, and in particular supersaturation, which is a necessary condition for contrail persistence. The modeling is based on a generalization of the Smith cloud scheme currently used in the operational ARPEGE. This modification can be implemented without any major modification and does not imply supplementary computational effort. A notable point is that the methodology developed in this article reworks the Sommeria and Deardorff (1977)

statistical concepts while incorporating ISSR parametrization, which allows for extensions to other atmospheric models using a similar framework. This is the case, for example, of the ICE3 microphysical scheme implemented in the French regional NWP model AROME (Seity et al., 2011, 2012), which uses a turbulence-based Gaussian distribution for the representation of clouds, and which could benefit from these developments.

Some common assumptions are shared with the parameterization implemented in the IFS model (Tompkins et al., 2007), in particular the use of a temperature function based on Koop et al. (2000) to describe the ice nucleation threshold, and the assumption that once this threshold is locally exceeded, local adjustment is instantly obtained back to saturation in the cloud. However, there are major differences between the two models in terms of their respective microphysical frameworks. For example, the cloud fraction is diagnostic in ARPEGE and prognostic in IFS. These different frameworks also imply differences in the representation of subgrid variability: IFS uses a dedicated distribution combining uniform (out-cloud) and Dirac (in-cloud) distributions, compatible with its prognostic cloud fraction parametrization, while ARPEGE uses a triangular distribution. The ARPEGE model also differs from other GCMs that allow supersaturation, such as ECHAM or CAM (see e.g. Bock and Burkhardt, 2016; Neale et al., 2010; Chen et al., 2012). These models are used for climate prediction and include supersaturation w.r.t ice, generally associated with a parametrization of contrail cirrus. In both cases, the triggering of homogeneous nucleation is proposed in a similar manner, mainly driven by a temperature function, but the thermodynamic in-cloud adjustment to saturation is relaxed, within a more complex framework of a 2-moment microphysical scheme.

2. The evaluation is primarily based on IAGOS aircraft measurements. While these are high-quality and relevant, the representativeness of IAGOS data (limited flight routes, sampling biases) should be more explicitly discussed. Could the conclusions change in regions less well sampled by IAGOS? The authors mention radiosondes as a possible next step, therefore an expanded discussion of current limitations in observational coverage would strengthen the evaluation framework. The authors should provide more detail on the dataset and discuss potential sampling biases.

Regarding the question of whether the conclusions could change in regions less well sampled by IA-GOS, the answer is addressed in the response to Special Comment 3. However, note that using data from a region of high data density, which coincides with a region of strong contrail cirrus, provides a solid foundation on which to evaluate NWP models.

Regarding the general comments on the observation datasets, the discussion on observations on L491 of Sect. 6.2 is reworded and enhanced to take into account the referee's demand:

In future works, other observational dataset could be useful for further verification of the NWP models. IAGOS dataset is open-accessible and quality-checked, relevant for humidity measurements in the UTLS. However, there is an heterogeneity in sampling due to the operational flight paths of participating airlines (Wolf et al., 2025). The highest measurement density is found across the North

Atlantic flight tracks, North-Eastern America and Europe at cruise altitudes (200hPa to 250hPa). In addition, the sampling is biased by avoiding aviation hazards (e.g. deep convective clouds and the outflow of such clouds, turbulent regions) and by route optimization according to the jet stream. Radiosonde observations which mainly sample the atmosphere above continental surfaces represent a valuable complementary database in world regions or weather situations where IAGOS data are less represented. This observation source also makes it possible to assess humidity at levels above or below typical cruising altitudes (see e.g. Bland et al., 2021; Thompson et al., 2024). Also, contrails tracked using satellite imagery [...]

3. The authors note that both ARPEGE and IFS underrepresent the highest values of RHice (> 110–115%). This is an important limitation, since persistent contrails are most strongly associated with highly supersaturated regions. Could the authors elaborate on the physical or numerical reasons why models fail at these extremes? Addressing this could help guide future model improvements.

Sperber and Gierens (2023) suspects that the cut-off in the histograms can be related to the assumption (commonly made in modified ARPEGE and IFS) that once the supersaturation threshold is locally exceeded, local adjustment is instantly obtained back to saturation, which leads that air inside cirrus clouds is assumed to be exactly at ice saturation. Indeed humidity grows progressively larger in the gridbox, and when locally crosses the supersaturation threshold, the local  $RH_{ice}$  is instantly dropped back to 100% while adding the excess vapor directly into solid phase. This is clearly an approximation as we know that in-cloud supersaturation can exist due the slow process of vapor deposition on pre-existing ice. This is also in line with the comment from referee 1 (specific point 3).

To clarify this point, we propose to rephrase the second paragraph from L442 in the Discussion Sect. 6.1:

For example, in the modified scheme, the adjustment in the cloud to ice saturation is assumed to be instantaneous, which is probably a major limitation of the physical description of ice representation in our context and an explanation of the cut-off on the highest supersaturation observed in the  $RH_{ice}$  distribution histogram (Fig. 6a). Indeed, it has been shown that  $RH_{ice}$  decreases with a relaxation time that can exceed several time steps, to finally reach a few percent above saturation, thus allowing local in-cloud supersaturation w.r.t ice. Sperber and Gierens (2023) proposed a cooling-cloud formation-phase relaxation process [...]

#### and to add a comment in the perspectives L583:

During this process [towards operationalization], additional studies will be conducted to better understand the underlying reasons for the underrepresentation of the highest supersaturations and to explore the feasibility and value of including more detailed representation of cirrus cloud physics, such as water vapour deposition, in the ARPEGE microphysical framework.

4. The modification to the cloud scheme is presented as computationally inexpensive and algorithmically simple, but the paper does not sufficiently explore whether introducing supersaturation impacts other parts of the model system. For example, how does it affect cloud microphysics, radiative transfer, or dynamics? Even if these effects are minor, they should be explicitly discussed.

Sentence L465 has been reworded to be more precise on the potential effects of introducing supersaturation on the model system and some figures have been added in the appendix to illustrate those effects, also in response to referee's special comment 6. The intensity of these impacts, in particular on the dynamics, and whether they benefit ARPEGE or not, cannot be assessed without a complete study, including the retroaction of observation assimilation which is beyond the scope of this paper. This work is currently underway with a view to implementing the changes operationally, and preliminary results show that after recalibrating some parameters of the microphysical scheme, a neutral to positive effect on prognostic variables such as temperature, wind or geopotential is obtained for the various regional domains of the WMO, traditionally used for NWP model verification.

#### Modification of sentence L465:

[...] Indeed, the introduction of supersaturation implies a reduction of the cloud cover and cloud ice contents in the altitudes which are impacted by the modification. This change will have a direct impact on temperature through changes in radiative and latent heat transfers, with an expected cooling effect, mainly in the upper troposphere (see Appendix B). The intensity of these impacts, their consequences for the dynamics, and whether they benefit or not for ARPEGE will be assessed in a dedicated forthcoming study.

## Special comments

1. L102–103: What is the value of  $C_{calib}$  in this study? How do you get this value? Is  $C_{calib}$  expected to change with variations in time and geographic location?

 $C_{calib}$  is a global tuning coefficient attached to the ice nucleation threshold formulation. In our study, it is empirically obtained in Sect. 4 by comparing the forecast  $RH_{ice}$  distribution, with the observed  $RH_{ice}$  distribution with an independent dataset (see Fig. 5 with several values of  $C_{calib}$  empirically tested). In the paper this value is fixed but it could vary among different localizations (eg. tropical, mid-latitudes, high latitudes), or time. Note that for NWP models in an operational context, there is usually a tradeoff between complexity and accuracy to deal with, and changing the calibration across different variables could add substantial complexity.

To have a best understanding on how  $C_{calib}$  is defined, more details are given in sentence L220: The calibration is based on adjusting the saturation ratio coefficient k through the calibration coefficient  $C_{calib}$  (see Fig. 4). The value of  $C_{calib}$  is obtained empirically by comparing the predicted and observed distributions of  $RH_{ice}$  using a dedicated calibration dataset.

In the discussion on the impact of regional domains, we also added a reference to the section describing calibration to help the reader make the connection (see the answer to Special Comment 3).

2. L147–150: To what extent is the cloud fraction sensitive to the choice of probability distribution? What alternative approaches exist for representing probability distributions beyond the normalized centered probability distribution? What rationale did the authors provide for selecting the normalized centered probability distribution over other methods?

In statistical cloud schemes, the probability distribution represents subgrid variability through local distances to saturation within the gridbox. Since cloud cover is obtained by integrating local points above saturation, the characteristics of the distribution will have a direct impact on cloud cover. For example, assuming a pre-defined standard deviation, a triangular distribution will result in a non-zero cloud fraction for lower relative humidity values compared to a uniform distribution, but the increase in cloud fraction relative to the increase in mean relative humidity will be more gradual. This behaviour will be accentuated if a Gaussian distribution is used.

The distribution used in our work is Triangular (Smith scheme - see Fig. 2) and not Normal (I think there may be a misinterpretation here). The mention "normalized centered" in calculation steps L148 does not refer to a Gaussian distribution, and has to be interpreted as: "centered and reduced" which is a preferable formulation to avoid any confusion. We propose to modify this in the article. Note that the reason for the choice of a triangular distribution is due to the fact that the operational ARPEGE uses this formulation that we want to adapt to allow ISSR.

3. L184–185: This study excludes the United States and Asia. How could that influence the results?

North America and Asia are large continental surfaces with large mountain ranges and the repartition between processes leading to ISSR in the atmosphere (e.g. Warm Conveyor Belt, Convection) probably differs from the European / Northern Atlantic climatology, and could modify the results. However, subcontinental variability should be considered. For example, Petzold et al. (2020) shows that the climatology of H<sub>2</sub>O at cruising levels and ISSR occurrence in the Eastern North America (ENA) is rather similar to the North Atlantic and Europe. Regarding subtropical Asia, the ISSR occurrences are rather low because of higher temperature at cruising levels, which may induce adaptation of the models.

We completed sentence L461 by including North America and Asia in the discussion about the impact of selecting different domains. We also referenced the section on calibration to help readers make the connection:

However, we believe that this work could be extended to tropical regions. In this case, it could be calibrated accordingly, for example by modifying the  $C_{calib}$  coefficient (see the calibration procedure,

section 4), given that tropical climatology differs from mid-latitude regions and that the occurrence of ISSR depends on factors such as dominant deep convection. Similarly, North America and Asia are large continental areas with large mountain ranges, and the distribution of processes leading to ISSR in the atmosphere (e.g. warm convection belt, convection) is likely to differ from European/North Atlantic climatology, which could alter the results.

### 4. L186: Please provide an explanation of FL250 and FL450.

The sentence L186 has been reworded: The vertical domain extends from flight levels FL250 to FL450 (i.e. 375hPa to 150hPa), which correspond to altitudes of 25000ft and 45000ft in the ICAO standard atmosphere (ICAO, 1993), encompassing regions favorable to the triggering of persistent contrails.

# 5. L215–218: Will the smoothing and undersampling alter the properties represented in the original IAGOS dataset?

In response to this comment and to specific comment 2 from Referee 1, we have reworked Figure 6 in Section 5.1 by adding IAGOS unfiltered observations to the frequency histogram. As illustrated in Figure 6a, there is no significant difference between filtered and unfiltered histograms, showing that filtering does not alter the properties represented in the original IAGOS dataset.

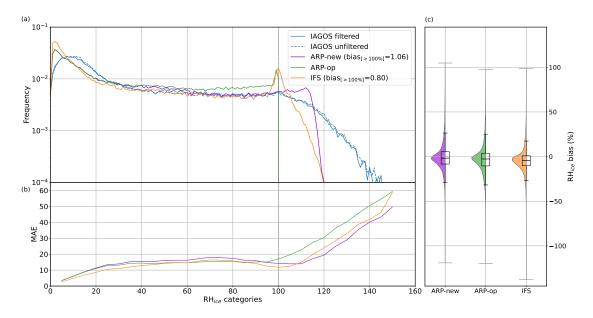


Figure 6: (a) Frequency histogram of  $RH_{ice}$  (1 % bins) with the associated frequency bias on  $RH_{ice}$  > 100% and (b) Mean Absolute Error (MAE) computed for different categories of observed humidity (5 % bins). Results are shown for IAGOS observational dataset (blue), filtered and unfiltered, ARP-new (purple), ARP-op (green) and IFS (orange). (c) Distribution plot of model bias in  $RH_{ice}$ ,

computed against unfiltered IAGOS observations, for the three ARP-new, ARP-op and IFS models. Verification dataset from the 1st July 2022 to the 30th June 2023 within the aerial boundary of 80° W-40° E and 30-75° N, covering North Atlantic and Europe.

Further comments have also been added to discuss the new figure:

L259: Model results are compared with filtered and unfiltered IAGOS observations, the latter being used to compute the distribution plots of direct model errors (Fig. 6c).

L267: We note that there is no significant difference between filtered and unfiltered observation histograms, showing that applying a 100s mean-filter does not alter the properties represented in the original IAGOS dataset.

6. Section 6.1: Providing results that show how ARP-new alters the cloud properties compared to ARP-op would be very helpful.

We have added results showing how the ice water content and cloud fraction of ARP-op have changed in ARP-new. Diagnostics in the horizontal domain show how cloud properties change when supersaturation is allowed in the model, as well as how this changes affect the temperature and the radiation fluxes. These results have been included in the appendix after the calculations relating to the modifications made to the ARPEGE cloud scheme. Below are the figures and comments we have included:

In this section, we analyze how the modifications introduced in ARPEGE to allow supersaturation w.r.t ice affect the cloud properties, which in turn may influence the temperature trend. The diagnostics over the horizontal domain toolbox (DDH toolbox, Piriou, 2025) allows us to quantify and analyze differences between the modified and non-modified models, ARP-new and ARP-op, for several ice cloud-related variables. The results are averaged by latitude bands for different pressure levels. As an illustrative case, we performed a diagnostic for the 6-hour forecast initialized on 1 January 2023 run 00UTC (Fig. B1).

Figure B1 shows that the introduction of supersaturation leads to a significant decrease in cloudiness and ice content in the atmospheric layers close to the tropopause. The impact on temperature is mainly reflected in a general cooling at and below the layers affected by this modification. This difference in temperature trend is primarily attributable to three contributions: solar radiation, thermal radiation, and stratiform processes related to solid water, the thermal radiative contribution being dominant. Indeed, the solar radiation diagnostic shows that in regions where cloudiness decreases, more solar radiation is transmitted downward, and less solar radiation is transmitted upward. The signal, however, remains weak comparatively to the thermal radiation diagnostic which reveals the opposite behavior of the dominant greenhouse role of cirrus clouds: with fewer cirrus, the trapping of outgoing terrestrial radiation is reduced, resulting in weaker longwave fluxes in the troposphere. At the same time, since less radiation is absorbed below, more longwave radiation escapes upward, which appears as an increase in fluxes at higher levels. Finally, a decrease in ice water content implies that

less water vapor is undergoing condensation at high levels, and therefore less latent heat is released, as shown in the stratiform solid water diagnostic.

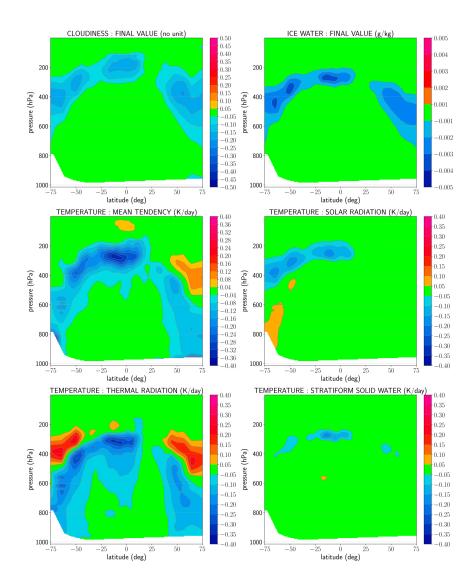


Figure B1: Differences between ARP-new and ARP-op betweenfor several parameters related to cloud processes at T+6h-1 January 2023 run 00UTC. The results are averaged by latitude bands for different pressure levels. A positive or negative difference is highlighted by red (or blue) colours on the scale.

7. L562: Why did the authors choose to emphasize results at a neighborhood tolerance of  $150 \,\mathrm{km}$ ?

We have chosen to focus on the results obtained with a neighbourhood tolerance of 150 km because this distance could be considered a reasonable margin allowing aircraft to avoid ISSR areas horizontally in an operational re-routing or alternative flight planning context. Note that this distance is also consistent with the work of Spichtinger and Leschner (2016), which shows, using the IAGOS dataset, that the decrease in  $RH_{ice}$  to significantly subsaturated values (considered as such when  $RH_{ice} < 70\%$ ) occurs approximately 100 km from the edges of ISSRs. We have reformulated the paragraph from L561 to make a clearer relation between spatial verification results and avoidance: Spatial verification allows to show that a good spatial agreement between forecast and observations is obtained. For example, discrimination capabilities when forecasting ISSRs with the modified ARPEGE at lead terms between 6 and 18 hours show a hit rate of  $\sim 80\%$  and a false alarm ratio of  $\sim 30\%$  when a neighbourhood tolerance of 150 km, in line with avoidance operations, is applied.

## **Bibliography**

- Jake Bland, Suzanne Gray, John Methven, and Richard Forbes. Characterising extratropical near-tropopause analysis humidity biases and their radiative effects on temperature forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(741):3878–3898, 2021.
- Lisa Bock and Ulrike Burkhardt. The temporal evolution of a long-lived contrail cirrus cluster: Simulations with a global climate model. *Journal of Geophysical Research: Atmospheres*, 121(7): 3548–3565, 2016.
- Chih-Chieh Chen, Andrew Gettelman, Cheryl Craig, Patrick Minnis, and David P Duda. Global contrail coverage simulated by cam5 with the inventory of 2006 global aircraft emissions. *Journal of Advances in Modeling Earth Systems*, 4(2), 2012.
- ICAO ICAO. Manual of the icao standard atmosphere [: extended to 80 kilometres (262 500 feet). Third Edition, Technical Report Doc 7488-CD, 1993.
- Thomas Koop, Beiping Luo, Athanasios Tsias, and Thomas Peter. Water activity as the determinant for homogeneous ice nucleation in aqueous solutions. *Nature*, 406(6796):611–614, 2000.
- Richard B Neale, Chih-Chieh Chen, Andrew Gettelman, Peter H Lauritzen, Sungsu Park, David L Williamson, Andrew J Conley, Rolando Garcia, Doug Kinnison, Jean-Francois Lamarque, et al. Description of the near community atmosphere model (cam 5.0). NCAR Tech. Note Near/tn-486+STR, 1(1):1–12, 2010.
- Andreas Petzold, Patrick Neis, Mihal Rütimann, Susanne Rohs, Florian Berkes, Herman GJ Smit, Martina Krämer, Nicole Spelten, Peter Spichtinger, Philippe Nédélec, et al. Ice-supersaturated air masses in the northern mid-latitudes from regular in situ observations by passenger aircraft: vertical distribution, seasonality and tropospheric fingerprint. Atmospheric chemistry and physics, 20(13):8157–8179, 2020.
- Jean-Marcel Piriou. Ddh toolbox. https://github.com/UMR-CNRM/ddhtoolbox, 2025.
- Yann Seity, Pierre Brousseau, Sophie Malardel, G Hello, P Bénard, F Bouttier, Ch Lac, and V Masson. The arome-france convective-scale operational model. *Monthly Weather Review*, 139(3): 976–991, 2011.

- Yann Seity, Christine Lac, Francois Bouyssel, Sebastien Riette, and Yves Bouteloup. Cloud and microphysical schemes in arpege and arome models. In *Proceedings of the Workshop on Parametrization of Clouds and Precipitation (ECMWF)*, Reading, UK, pages 5–8, 2012.
- G Sommeria and James W Deardorff. Subgrid-scale condensation in models of nonprecipitating clouds. *Journal of the Atmospheric Sciences*, 34(2):344–355, 1977.
- Dario Sperber and Klaus Gierens. Towards a more reliable forecast of ice supersaturation: concept of a one-moment ice-cloud scheme that avoids saturation adjustment. *Atmospheric Chemistry and Physics*, 23(24):15609–15627, 2023.
- Peter Spichtinger and Martin Leschner. Horizontal scales of ice-supersaturated regions. Tellus B: Chemical and Physical Meteorology, 68(1):29020, 2016.
- Gregory Thompson, Chloé Scholzen, Scott O'Donoghue, Max Haughton, Roderic L Jones, Adam Durant, and Conor Farrington. On the fidelity of high-resolution numerical weather forecasts of contrail-favorable conditions. *Atmospheric Research*, 311:107663, 2024.
- Adrian M Tompkins, Klaus Gierens, and Gaby Rädel. Ice supersaturation in the ecmwf integrated forecast system. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 133(622):53–63, 2007.
- Kevin Wolf, Nicolas Bellouin, Olivier Boucher, Susanne Rohs, and Yun Li. Correction of era5 temperature and relative humidity biases by bivariate quantile mapping for contrail formation analysis. *Atmospheric Chemistry and Physics*, 25(1):157–181, 2025.