

Response to Referee #3

Bingying Shi Philipp Griewank Florian Meier Annika Schomburg
Jinzhong Min Martin Weissmann

September 12, 2026

This statement concerns our revision of the Paper #egusphere-2026-1463 submitted to the Atmospheric Measurement Techniques, entitled “*Revisiting error models for the assimilation of all-sky infrared satellite radiances*”, based on the referees’ report.

We would like to thank all reviewers for taking the time to read and review our manuscript. In response to the comments we have added another 1000 bootstrapping runs based on 3 months of data from another modelling system (DWD ICON) data in another season (winter). As we incorporated an additional data set from DWD, we also added Annika Schomburg as additional coauthor from DWD. A description of the ICON data is added in Sect. 2.3. To show the bootstrapping runs more clearly, we split Sect. 5 into two subsections. Results from AROME and ICON are presented in Sect. 5.1 and Sect. 5.2, respectively. Additionally, we incorporated the findings into the ‘Discussion and summary’ section. The overall conclusion, however, remain largely unchanged as the new dataset confirmed the main conclusions.

Other than the large amount of text tweaks due to adding the additional dataset, the biggest change was the inclusion of two additional figures (Figs. R1 and R2), which are Figs. 9 and 14 in the revised manuscript. Figure R1 illustrates that the maximum cloud effect is physically better justified, and Fig. R2 gives a more detailed illustration of the bootstrapping results for every run.

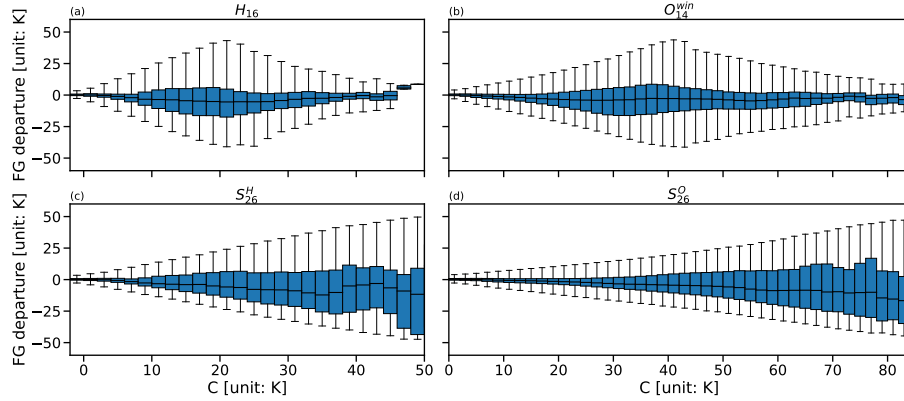


Figure R1: Box plots of Ch7.3 FG departures (verification period) binned for every 2 K of C . C computed with (a) H_{16} method, (b) O_{14}^{win} method, (c) S_{26}^H method and (d) S_{26}^O method. The lower and upper boundaries of the box are the first quartile and third quartile, respectively; the lower and upper whiskers show the 0.5 percentile and 99.5 percentile, respectively.

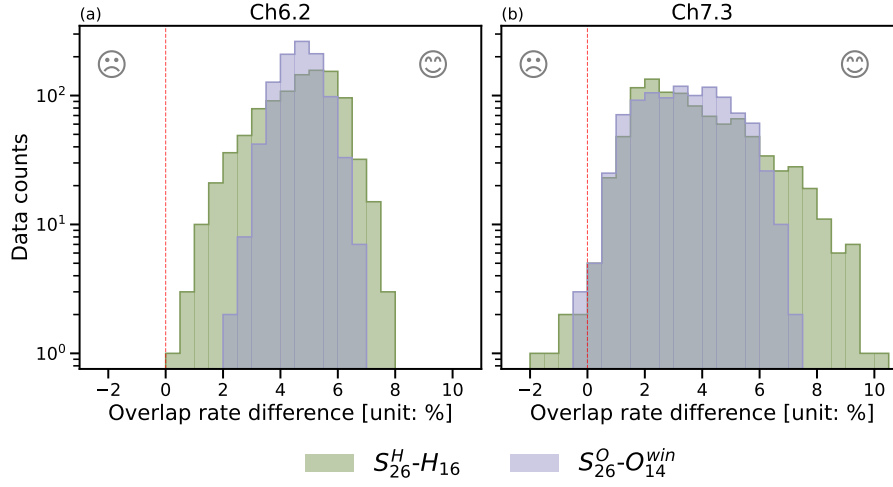


Figure R2: Histogram of overlap rate difference between S_{26}^O and O_{14}^{win} methods (deep indigo) and between S_{26}^H and H_{16} methods (olive gray) for (a) Ch6.2 and (b) Ch7.3. Results are computed based on 1000 bootstrapped runs.

In addition, we corrected one error in L247-250, where we explain which data is removed by QC. We can not know whether the FG departure is twice the cloud effect or minus twice the cloud effect based on the conditions we provide. Thus, we changed the text into: *"We can understand which data is removed by looking at when the absolute FG departure is about twice the cloud effect. This can occur when no clouds are present in the model ($BT_{\text{cld}}^M = BT_{\text{clr}}^M$), resulting in an $|FG \text{ departure}|$ exactly twice the cloud effect, or when $BT^O \approx BT_{\text{clr}}^M$, resulting in an $|FG \text{ departure}|$ approximately twice the cloud effect."*

We also corrected some typos throughout the text, and replaced some full names (such as brightness temperature) with their abbreviations for consistency.

Below we respond to the comments of referee #3. Text in *italics* is the new text in the paper. We have not included here all the modified text from the manuscript as it would make this long document even much longer.

Comments by Referee #3

1. Section 1: This study is largely motivated by the nonlinear non-Gaussian effects associated with all-sky IR statistics. Please include the following relevant references on the subject.

Lei, L., Ju, H., Fu, K., Anderson, J. L., Zhou, L., & Tan, Z. M. (2026). Quantile-Conserving Ensemble Filters for All-Sky Infrared Radiance Assimilation. *Monthly Weather Review*, 154(1), 99-117.

Zhou, L., Lei, L., Tan, Z. M., Zhang, Y., & Di, D. (2023). Impacts of observation forward operator on infrared radiance data assimilation with fine model resolutions. *Monthly Weather Review*, 151(1), 163-173.

Chan, M. Y. (2026). Improving convection-permitting all-sky infrared radiance ensemble data assimilation through mitigating deleterious non-Gaussian artifacts. *Quarterly Journal of the Royal Meteorological Society*, 152(774), e70039.

Chan, M. Y., Chen, X., & Anderson, J. L. (2023). The potential benefits of handling mixture statistics via a bi-Gaussian EnKF: Tests with all-sky satellite infrared radiances. *Journal of Advances in Modeling Earth Systems*, 15(2), e2022MS003357.

Chan, M. Y., Zhang, F., Chen, X., & Leung, L. R. (2020). Potential impacts of assimilating all-sky satellite infrared radiances on convection-permitting analysis and prediction of tropical convection. *Monthly Weather Review*, 148(8), 3203-3224.

Thank you for providing the valuable suggestion. We've added the first and fourth references in the discussion of the possible solution to address non-Gaussianity. The third one has been added when stating that non-Gaussianity is one of the major challenges for assimilating all-sky infrared radiances. We also did some small tweaks in introduction section to make the motivation of applying cloud-dependent observation error more clear.

2. L47: The sentence beginning with "This method was first proposed..." should be the start of a new paragraph.

Revised accordingly.

3. L101: "peak around 350 hPa and 500 hPa" -- please include a reference for this.

Thank you for your advise. We've included Schmetz et al. 2002 showing the weighting functions for all three channels we used in our manuscript. We added the citation at the end of the sentence "*Ch10.8's weighting function peaks near the surface in clear-sky conditions*" in L109 in the revised manuscript.

References

Schmetz, Johannes et al. (2002). "An introduction to Meteosat second generation (MSG)". In: *Bulletin of the American Meteorological Society* 83.7, pp. 977–992. DOI: 10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2.

4. L101: "For Ch10.8, it peaks near..." -> "Ch10.8's weighing function peaks near..."

Revised accordingly.

5. L107: "Austria in 2023,. AROME" -> "Austria in 2023. AROME"

Revised accordingly.

6. L110: "is still currently" -> "is currently"

We've deleted the word "still".

7. L111: "but the GeoSphere plan to switch to a 1km grid spacing some time in 2026" -- this tidbit does not contribute towards your story. Please eliminate that.

We've removed sentence.

8. L134: "which cloud be" -> "which could be"

Revised accordingly.

9. L135: "low/medium clouds, or it could be that with less" -> "low/medium clouds. Alternatively, with less"

We revised the text to: *Additionally, there are more simulated than observed BTs around 230-235 (250-258) K at Ch6.2 (Ch7.3). This could be due to the model overestimating the occurrence of low/medium clouds. additionally, less high clouds can lead to a larger influence of low/medium clouds on the simulated BTs.*

10. L142: "where $\sup_x$ is the" -> "where the $\sup_x$ operator finds the"

Revised.

11. L147: "it's" -> "it is".

Revised.

12. Eq (2): It took me a while to understand how this quantity is a measure of the two distribution's overlap. Am I correct in interpreting this quantity as the ratio of (a) the intersection of two pdf's areas under curve versus (b) the union of the two pdf's areas under curve? I recommend adding a graphical illustration to help readers understand this quantity.

Yes, that's exactly how we calculate the overlap rate. As we've already had lots of figures illustrating two distribution's overlap, we decided to use Fig. 4 as an example when explain how we compute overlap rate. We've added the the following text in L165-166 in the revised manuscript to help readers understand the quantity:

For example, as shown in Fig. 4a, the overlap rate is the ratio of the the dark blue area versus the combined total dark blue, light blue, and red area.

13. L160: Add an explicit statement of the purpose for having a threshold BT. My impression is that the threshold BT's purpose to separate clear-sky BTs from cloudy-sky BTs.

Thanks for your suggestion. We've added in L177 in the revised manuscript: *The purpose of BT_{lim} is to separate cases where clouds have an impact on the BT from those where their impact can be neglected.*

14. L161: "For deriving BT_{lim} , we" -> "We"

Revised.

15. L165: "note the logarithmic scaling" -> "note the logarithmic color scale"

Revised.

16. Eq (3): To help readers with this equation, break up the right hand side into two fractions. The first fraction is the model term, and the second fraction is the observaiton term.

Thanks for the suggestion. We've adjusted the equation accordingly.

17. Eq (4): I cannot locate the definitions of C_{min} and C_{max} . Did I miss them?

Sorry for missing their definitions. We've added them in the two paragraphs following Eq. (4).

18. L193: The term "model error" is vague as it can refer to either model representation errors (epistemic uncertainties) or background uncertainties (aleatoric uncertainties). Please use a more precise term instead of "model error". This term crops up a few times in your paper.

We have difficulties following this statement as in the context of data assimilation model error is a common, essential and clearly defined term. Disentangling the different contributions to model error is often impossible and may lead to speculation. Thus we kept the term "model" error, but we went through the manuscript and tried to be as precise as possible when speaking about model error. We rephrased the sentences when this term first appears in L210-215 in the revised manuscript:

In this context, it should be noted that the Std of FG departures is the square root of the sum of observation error and model error in observation space, where the model error refers to the uncertainty of the background. The approach we used ignores the contribution of model error, which leads to a slightly inflated observation error estimate. Sensitivity tests of Harnisch et al. (2016) using the ensemble spread as proxy for model error in observation space, however, showed that ignoring model error only has a minor effect on the estimate.

19. L206: "the observed profile" -> "the observed BT"?

Thank you for the comment, but what we refer to here is the profile of all physical variables (e.g. temperature, humidity, pressure, hydrometeors) that can affect the value of observed BT, not the observed BT itself. Since "the observed profile" may differ due to model error, the "observed" clear-sky BT would be different from BT_{clr}^M , making the cloud effect in observation space inaccurate when using BT_{clr}^M as the clear-sky reference.

To avoid confusion, we've rephrased the text: *However, BT_{clr}^M is only a near-perfect reference for the cloud effect in model space, while the "observed" clear-sky BT may differ from BT_{clr}^M due to inaccurate model fields caused by model error.*

20. Eq (5): Same as comment 16.

Revised.

21. L220: "It produces" -> " O_{14} produces"

Changed.

22. L220: "doesn't" -> "does not"

Revised.

23. L223: "by computing the" -> "by comparing the"

We rephrased the sentence to: ... *when computing the observed cloud effect by comparing observed BT against the clear-sky model BT*, ...

24. L225-229: I agree with you that if Std is always equal to $2 C_j$ for clear-clear situations, then the two-point Bernoulli distribution manifests. However, it is unclear to me why that equivalence exists. Could you elaborate on why that equivalence exists?

The underlying reason is that the used reference (BT_{clr}^M) contains model error. We added this statement in the text in L241-242 in the revised manuscript: *"The underlying reason why the O_{14} approach fails when no clouds are present is that the reference for estimating the cloud effect ($BT_{\text{clr}}^M = BT_{\text{clr}}^M$) contains model error."*

Hence, the cloud effect can be written as Eq.(6) in the manuscript, which means $|FG \text{ departures}| \sim 2C_{BT_{\text{clr}}^M}^{\text{mean}}$. The value of $C_{BT_{\text{clr}}^M}^{\text{mean}}$ is close to a constant value C_j within a small bin bin_j . Assuming that the FG departures are unbiased, and if the sample size in bin_j is big enough, the Std_j of FG departures in the j -th bin bin_j for cloud-free condition can be written as:

$$\text{Std}_j = \sqrt{\frac{\sum_{i=1}^n (\text{FG departures} - \overline{\text{FG departures}})^2}{n-1}} \quad (1)$$

$$= \sqrt{\frac{\sum_{i=1}^n \text{FG departures}^2}{n-1}} \quad (2)$$

$$\simeq \sqrt{\frac{\sum_{i=1}^n (2C_j)^2}{n-1}} \simeq 2C_j \quad (3)$$

Thereby, the Std is always equal to $2C_j$ for clear-clear situations.

25. L235-236: Remove the sentence beginning in "For this modification".

Removed.

26. L238: "Here, $C_{\text{min}}=2$ is set for futher analysis" -> "Here, we set $C_{\text{min}} = 2$."

Revised accordingly.

27. Eq (8): The sheer number of maximum operators make this equation hard to read. Maybe use "" instead of "()" for the outer-most maximum operator, and use "[]" for the middle maximum operator? Also, add some spaces here and there to help with reading your equation (this would be the \, command in LaTeX).

Thanks. We've changed "()" in Eq (8) following your suggestion, and we've added a space between the parentheses of the max operator and the variables. We've also made the same modification to all max operators all over the manuscript accordingly.

28. L400: "We are convinced that the revised approach" -> " The revised approach"

Changed.

29. I just want to mention here that I did think about whether this study needs to have some DA experiments to confirm that their revised methods result in better results than the original methods. This is something that RC1 did note. However, the Acknowledgements suggest that this study was done during the lead author's visit to the U of Vienna. As such, I suspect that DA experiments might be beyond the resource constraints of the lead author. Therefore, I decided against making that a major comment.

Thanks for your understanding. Yes, to evaluate the impact of observation error on analysis and forecast, a dedicated set of sensitive experiments within a long time period is needed, which is currently impossible for us. But we hope to perform such an evaluation in future projects as stated above.