

Response to Referee #2

Bingying Shi

Philipp Griewank
Jinzhong Min

Florian Meier
Martin Weissmann

Annika Schomburg

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.

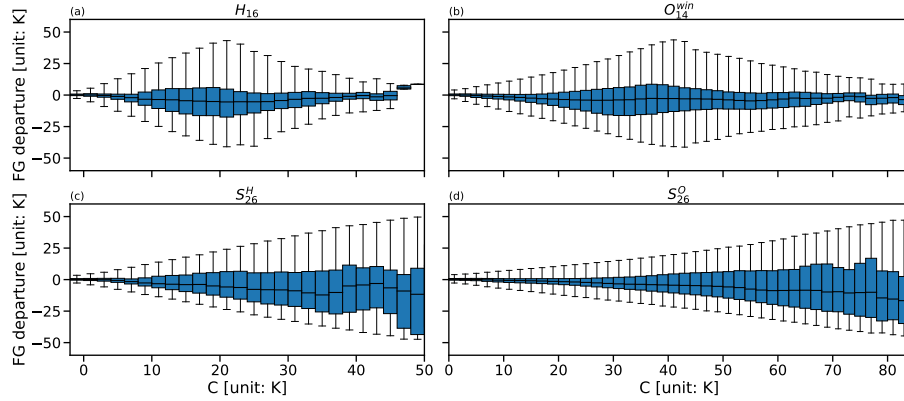


Figure R1: Box plots of Ch7.3 FG departures (verification period) binned for every 2 K of C . C computed with (a) H_{16} method, (b) O_{14}^{win} method, (c) S_{26}^H method and (d) S_{26}^O method. The lower and upper boundaries of the box are the first quartile and third quartile, respectively; the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

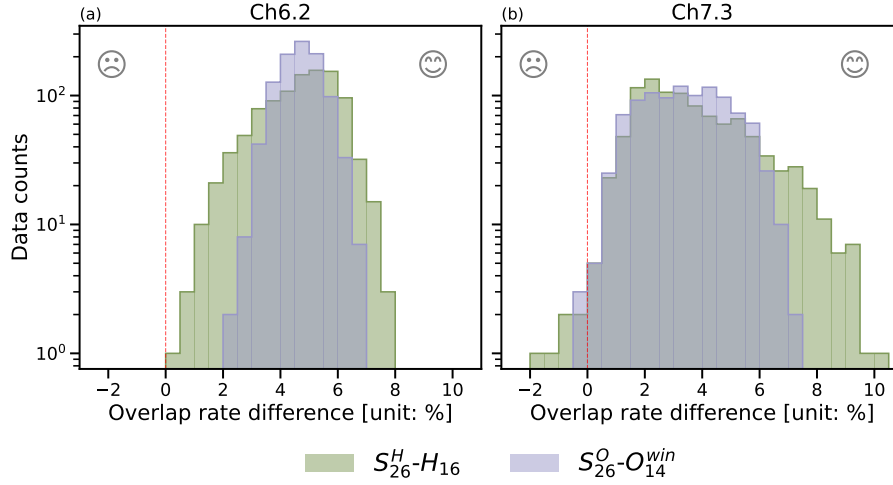


Figure R2: Histogram of overlap rate difference between S_{26}^O and O_{14}^{win} methods (deep indigo) and between S_{26}^H and H_{16} methods (olive gray) for (a) Ch6.2 and (b) Ch7.3. Results are computed based on 1000 bootstrapped runs.

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 #2. 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 #2

Major comment 1. Is the use of maximum cloud effect representative and appropriate for all the cases? A statistical analysis to justify this choice would be essential for the improvement of the manuscript.

To show that the use of maximum cloud effect is representative, we added two features to the manuscript. The first feature is an additional figure Fig. R2 (Fig. 14 in the revised manuscript) showing the histograms of overlap rate differences between the maximum and mean approaches in the 1000 bootstrapped experiments. This figure shows that the max approaches outperform the mean approaches in all 1000 bootstrapped experiments for Ch6.2, and in over 99% of bootstrapped experiment for Ch7.3. Secondly, we added results for data from another modeling system (DWD ICON) and another season (winter), in which the max version also outperforms the mean version for

the Harnisch approach. Unfortunately, we could not apply the Okamoto window channel approach to the ICON data, because the window channel is not included in the DWD system.

Major comment 2. The authors use one month (August 2023) of data. Ideally, 4 months (1 per season) or a full year would be more presentative for the outcomes of the study. Is it something possible? If not, some case studies of other months would be necessary to evaluate the methodology.

As this was a consistent comment by all 3 reviewers, we've added results of another 1000 bootstrapping runs with a new dataset of three winter months (Jan, Feb, Dec) data from the DWD operational model ICON. It generally shows similar results as those with AROME data. Since the new dataset is both from a different season with longer time period and from a different NWP model, it underlines the generalizability of the results.

A description of the ICON data and the results of the error models for it have been added in Sect. 2.3 and Sect. 5.2. Furthermore, we've incorporated the findings into the 'Discussion and summary' section and made some minor revisions throughout the manuscript to make it consistent.

Minor comment 1. Page 2, Line 31: Add the acronym ECMWF & NCEI.

Added.

Minor comment 2. Page 4, Line 121: "... for the BTcld ..." and not BTclr.

Thanks for your careful review. We've corrected it.

Minor comment 3. Same comment for Figure's 1 & 2 captions.

Revised accordingly.

Minor comment 4. Page 6, Line 151: Could you justify your choice of bin width equal to 0.1?

We intended to bin our data in a decimal scale, therefore we choose the bin width of 0.1. Binning the data in a smaller bin width can help to get more an accurate result, but we believe using the bin width of 0.1 is sufficient for our purpose. For example, for Ch7.3 with H_{16} method (Fig. 4b in the manuscript), the overlap rate when using the bin width of 0.1 and 0.01 are 85.216% and 85.172%, respectively. This small difference indicates that changing the bin width only has a minor impact on the calculated overlap rate.

Minor comment 5. Page 7, Line 165: BTclr both of them? Please make sure that BTclr and BTcld are correct throughout the text.

Thanks for your careful review. We’ve thoroughly checked the text and corrected this.

Minor comment 6. Page 10, Line 238: It should be stated where the selection of $C_{\min}=2$ is based on.

Thank you for the suggestion. We use $C_{\min} = 2$ based on the sensitivity evaluation shown in Fig. 7 in the manuscript. The highest overlap rate is obtained when approximately $C_{\min} = 1.5$ for Ch6.2 and $C_{\min} = 1.1$ for Ch7.3. After that, the Gaussianity of distribution decreases gradually but in return the data utilization increases. Due to this trade-off, we think setting $C_{\min} = 2$ is a balanced choice since it can both ensure the distribution near-Gaussian and reserve enough data. For clarification, we’ve added a line after we set $C_{\min} = 2$:

The impact of the value of $C_{\min}$ will later be described in detail in Fig. 7.

Besides, due the quite different results from AROME and ICON model with the same value of $C_{\min}$, we’ve added one item (the fourth item in the revised manuscript) in the last section about the impact of $C_{\min}$:

For the own-channel version of Okamoto et al. (2014), the impact of QC differs for different NWP models and meteorology conditions. We therefore recommend to tune the QC for the respective modeling system.