Response to Reviewer 2

General comment:

The manuscript is generally well written, the figures and tables are presented in a comprehensive and clear way.

However, I have one methodological problem with the evaluation in section 4: The new model (ARMS-gb v2.0) is compared to observations and other fast radiative transfer models (RTTOV-gb and ARMS-gb v1.0 without clouds). What is missing to my opinion is the comparison to a high-accuracy radiative transfer tool, such as MonoRTM or others. Only with that it can be assessed whether the observed biases (see Fig. 6) are due to instrumental issues or due to deficiencies of the model. A further uncertainty source might be biases in ERA-5 data, especially regarding cloud occurrence. As the behavior of YKW-1 and YKW-2 are quite different, I would be very much interested in this comparison, and it would make it possible to attribute biases to either model or observation.

To my mind this is crucial to assess the performance of the new radiative transfer model.

Response:

We sincerely thank the reviewer for the positive assessment of the manuscript and for identifying this important methodological issue. We agree that comparison with a high-accuracy radiative-transfer model is necessary to better assess the origin of the channel-dependent O–B structures. To address this concern, we added MonoRTM clear-sky simulations as a reference in Figure 6. MonoRTM and ARMS-gb were driven by the same clear-sky ERA5 profiles and configured using the same GMR channel settings. For each channel, 256 monochromatic calculations across the nominal bandwidth were averaged to obtain a channel-equivalent brightness temperature. The resulting ARMS-gb and MonoRTM clear-sky baselines show closely matching channel dependence, with maximum differences of approximately 0.38 K at Yan'an and 0.30 K at Karamay. This agreement indicates that deficiencies in the ARMS-gb clear-sky radiative-transfer calculations are unlikely to be the dominant source of the persistent baseline structures. We nevertheless recognize that this comparison cannot uniquely separate uncertainties associated with instrument calibration, channel characterization, observation–ERA5 collocation, and the ERA5 atmospheric backgrounds, because the two models use the same atmospheric inputs and channel configurations. We have clarified this limitation and revised the interpretation of the O–B departures accordingly. More broadly, we expanded the instrument and calibration descriptions, provided additional details on the sensitivity experiments, ERA5 inputs, scene classification, and sample sizes, and reorganized Section 3 to consolidate the hydrometeor-sensitivity analyses. We also revised the discussion of potential error sources, updated the figures and captions, thoroughly edited the manuscript for clarity and readability, and archived the processed O–B datasets used in the evaluation experiments in a publicly accessible repository. These revisions have substantially improved the transparency, reproducibility, and scientific precision of the manuscript.

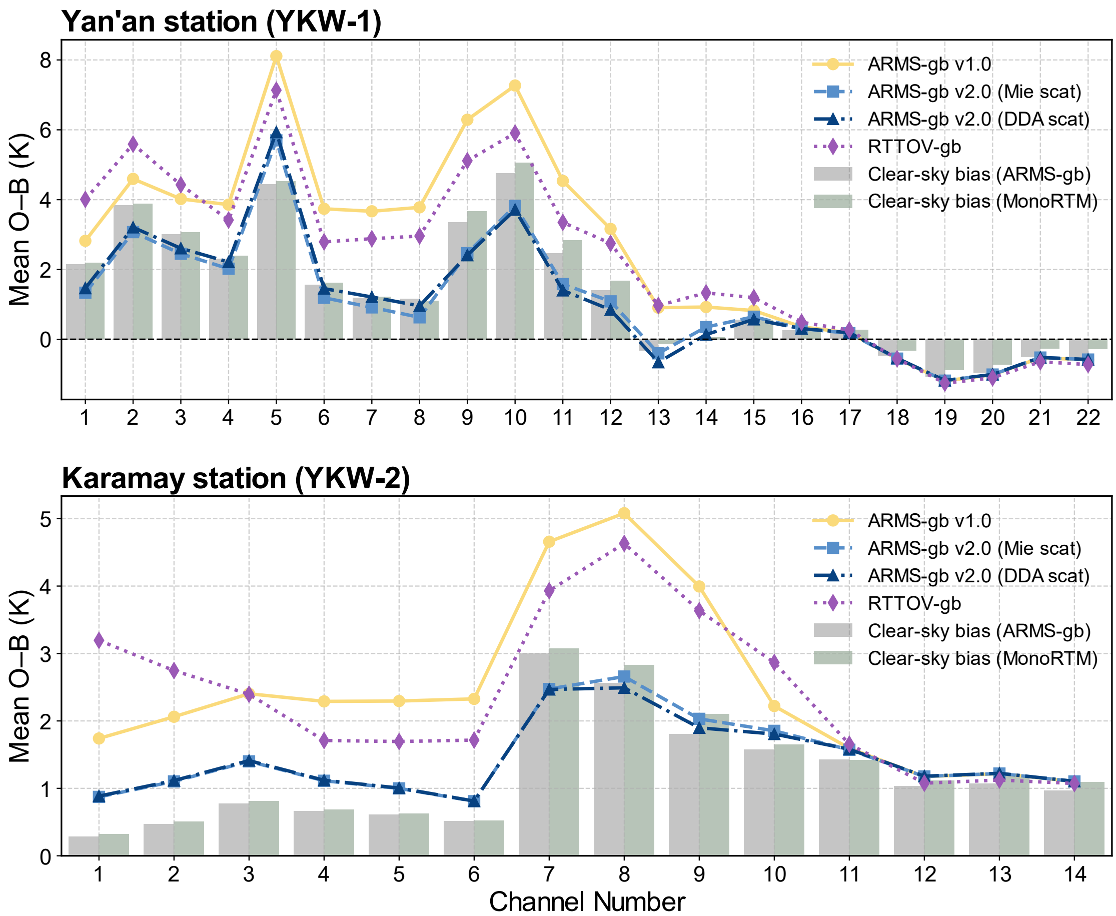

Figure 6: Mean O-B departures for each channel at Yan'an (YKW-1) and Karamay (YKW-2) during April–June 2025. Lines show cloudy non-precipitating simulations, while bars show ARMS-gb and MonoRTM clear-sky baselines.

Comment 1:

Line 29: especially the vertical structure of the PBL is important for extreme weather

Response: We agree with the reviewer and have revised the relevant sentence to emphasize the importance of the vertical thermodynamic structure of the PBL in extreme-weather development. The revised text in Lines 31–32 reads:

"In recent years, extreme weather events have become more frequent and variable, and the vertical thermodynamic structure of the planetary boundary layer (PBL) plays a central role in their formation and evolution."

Comment 2:

Line 41-42: better: "profiles with moderate vertical resolution, and decreasing with height"

Response: We thank the reviewer for this helpful suggestion. The sentence has been revised to clarify that the vertical resolution of GMR retrievals decreases with height. The revised text in Lines 44–46 reads:

"Ground-based microwave radiometers (GMRs), operating in an upward-looking configuration, provide continuous retrievals of temperature and humidity profiles under all-sky conditions, with high temporal resolution and moderate vertical resolution that decreases with height within the PBL".

Comment 3:

Line 45: better: "local" than "regional"

Response: We agree with the reviewer and have replaced "regional representativeness" with "local representativeness" to more accurately describe the spatial representativeness of GMR observations. The revised text in Lines 48–50 reads:

"Compared with radiosondes, GMRs can offer high-frequency, continuous observations that are not affected by wind advection and therefore show stronger local representativeness, which enables long-duration all-weather monitoring within the PBL."

Comment 4:

Line 47: Absorption is dominated by water-vapor absorption only in K-Band (not V-Band)

Response: We agree with the reviewer and have corrected the spectral description to distinguish the dominant absorbers in the K- and V-bands. The revised text in Lines 50–52 reads:

"Under clear-sky conditions, the atmosphere is nearly transparent in the window channels; absorption in the K-band is dominated by water vapor, whereas the V-band is primarily controlled by oxygen absorption."

Comment 5:

Line 120: Please provide more information about the instruments (YKW-1 and YKW-2) From which manufacturer were they built? Are there references to these instruments?

Response: We thank the reviewer for this suggestion. We have expanded the instrument description to specify the manufacturers, channel configurations, frequency coverage, and operational calibration information for YKW-1 and YKW-2. The revised text in Lines 123–129 reads:

"Two types of GMRs are used in this study: the 22-channel YKW-1 and the 14-channel YKW-2. Both are ground-based microwave radiometer systems manufactured in China, with YKW-1 produced by Beifang Tianqiong and YKW-2 produced by Airda. Their channel settings follow those of the MP-3000A and HATPRO radiometers used in Zheng et al. (2026), respectively, with YKW-1 using a 22-channel configuration and YKW-2 using a 14-channel configuration. These channels cover the water-vapor absorption band near 22–31 GHz and the oxygen absorption band near 51–59 GHz, and their center frequencies are listed in Tables 1 and 2."

Zheng, Q., Sun, W., Liu, Z., Mao, J., He, J., Li, J., and Jiang, X.: Direct assimilation of ground-based microwave radiometer observations with machine learning bias correction based on developments of RTTOV-gb v1.0 and WRFDA v4.5, Geoscientific Model Development, 19, 731–754, https://doi.org/10.5194/gmd-19-731-2026, 2026.

Comment 6:

Lines 194-195: Please rephrase – I would rather say, "….plays a fundamental role in the microwave radiative transfer and thus in the resulting brightness temperature.

Response: We thank the reviewer for this suggested wording and have revised the opening sentence of Section 3 accordingly. The revised text in Lines 199–200 reads:

"The interaction between microwave radiation and atmospheric hydrometeors plays a fundamental role in microwave radiative transfer and thus in the resulting brightness temperatures observed by GMRs."

Comment 7:

Lines 247-248: How many profiles did you use? Daily? Hourly?

Response: We thank the reviewer for pointing out this omission. The analysis uses hourly ERA5 profiles rather than daily mean profiles. We have clarified the temporal resolution, analysis period, and number of retained profiles. Following the reorganization of Section 3, the analysis originally presented in Figure 5 is now presented in Figure 4. The relevant revisions in Lines 232–243 read:

"Hydrometeor effects were further examined under realistic atmospheric conditions using hourly ERA5 profiles collocated with GMR observations from January to June 2025."

"After scene classification and quality-control screening, 2,135 hourly profiles representing cloudy non-precipitating conditions were retained for Yan'an and 2,845 for Karamay."

Comment 8:

Line 368: Here you write about instrument calibration issues. Can you specify the calibration accuracy for both instruments? How often were these radiometers calibrated and using which method? This should be part of the instrument section.

Response: We thank the reviewer for this suggestion. We have added the calibration method, frequency, and nominal accuracy for both instruments to the instrument description. The revised text in Lines 129–130 reads:

“Both YKW-1 and YKW-2 are subject to routine operational calibration using liquid nitrogen every six months, with a nominal calibration accuracy better than 1 K.”

Comment 9:

Figure 5: Do you average over all cases? How would the difference look for only cloudy cases, excluding all clear sky scenes?

Response: We thank the reviewer for raising this point. The analysis does not average over all available cases. Clear-sky and precipitating scenes were excluded before averaging, and only cloudy non-precipitating hourly profiles were retained. Following the reorganization of Section 3, the original Figure 5 is now Figure 4. The scene-classification procedure and retained sample sizes are described in Lines 236–243:

“Each hourly sample was assigned to one of three mutually exclusive categories: precipitating, clear-sky, or cloudy non-precipitating. Samples with a radiometer rain-sensor flag were classified as precipitating. Among the remaining samples, clear-sky cases were identified following Shi et al. (2025a). The screening used sky infrared temperature, the 10 min brightness-temperature standard deviation in the window channel near 31 GHz, and the sum of ERA5 total-column cloud liquid water and cloud ice water. All remaining non-precipitating samples were subsequently classified as cloudy non-precipitating cases. After scene classification and quality-control screening, 2,135 hourly profiles representing cloudy non-precipitating conditions were retained for Yan’an and 2,845 for Karamay.”

The revised Figure 4 caption in Lines 258–261 further states:

“The results are based on ARMS-gb v2.0 simulations driven by 2,135 and 2,845 hourly cloudy non-precipitating ERA5 profiles, respectively. Bars show the responses to cloud liquid water, rainwater, cloud ice, and snow included separately, together with the combined response when all four species are included. ΔBT is calculated relative to the corresponding clear-sky simulation.”

Comment 10:

Data availability: Please provide the datasets from GMRs on an openly accessible platform.

Response: We thank the reviewer for this important request. The original operational GMR observations used in this study are managed by the China Meteorological Administration (CMA), and access requires application and approval through CMADaaS in accordance with the CMA data policy. We are therefore unable to redistribute the complete raw observation records through a publicly accessible repository. To facilitate reproducibility, we have archived the processed O−B datasets used in the model-evaluation experiments on Zenodo at https://doi.org/10.5281/zenodo.21262924. The Code and data availability section has been revised to distinguish clearly between access to the original GMR observations and the openly available processed O−B datasets.