General comment:

Thanks to the reviewer's efforts to identify methodological flaws in the preprint - which should have been addressed by the authors prior initial submission(!) - the key message of this work changed remarkably: While the preprint stated a 17% and 23% RMSE reduction due to oxygen channel and polarization features compared to the baseline, these numbers reduced to 5% with (almost) no additional improvement due to polarization features. A comparison of Fig. 3a between the three manuscript versions raises a question as to which degree the small remaining differences are even statistically significant. At the same time, the authors state that polarization differences have a "critical role" in reconstruction, which is not supported by Figure 3. Given the large change of the results and the misalignment between tables/figures and text I believe that the authors should carefully revise the entire manuscript and further check the statistically significance, especially for the polarization difference effect. The spread of RMSE of the three independently trained models for each experiment should be added to Table 3 and Fig. 3 (e.g. as a shading). This will help to assess whether the effect of input features is statistically significant or not, i.e., are 5% RMSE reduction smaller or larger than the RMSE spread of the three random initializations of each experiment. And does PD really improve the reconstruction or is it just noise that gets interpreted as an improvement?

Response：

We deeply appreciate the reviewer's thorough and insightful feedback, which has significantly improved the rigor and clarity of our manuscript. We sincerely apologize for the methodological flaws in the preprint, including the overstated RMSE reductions (17–23%) and the erroneous claim of a "critical role" for polarization differences (PD), which were not supported by Figure 3. These issues, which should have been addressed prior to submission, led to a misalignment between the text and figures/tables, and we are grateful for the reviewer's diligence in identifying them.

The substantial reduction in reported RMSE improvements (from 17–23% to approximately 5%) resulted from methodological corrections, including optimized training strategies, model enhancements, and the use of three independent training runs to account for training variability. For instance, in coastal precipitation scenarios, Ex14 (mean RMSE = 4.02, spread = 0.043) and Ex26 (mean RMSE = 3.80, spread = 0.024) yield an RMSE reduction of 5.47% ± 1.17%. PD (Ex35) contributes only 0–1.6% RMSE reduction, with curves largely overlapping Ex26's, indicating no statistically significant effect due to comparable RMSE spread.

To address the reviewer's concerns, we are undertaking a comprehensive revision of the manuscript with the following actions:

1. Statistical Significance Evaluation: We have computed the RMSE spread from three independent training runs to confirm that the 5.47% RMSE reduction for dual oxygen channels is not due to model randomness, with an improvement spread of 1.17%.
2. Table 3 Update: We have augmented Table 3 to include the RMSE spread for each experiment across three training runs, enabling direct comparison with reported improvements.
3. Figure 3 Revision: We have modified Figure 3 to incorporate RMSE spread as shading

around the mean RMSE curves for Ex14, Ex26, and Ex35. This will visually demonstrate the overlap between Ex26 and Ex35 curves and contextualize the 5.47% improvement relative to the spread.

4. Text Revisions: We have corrected the overstated PD contribution in the manuscript and reviewed all sections to ensure consistency between text, figures, and tables.

These revisions will provide a clearer and more robust assessment of the input features' effects, particularly the limited role of PD.

Specific comments:

- Line 10-11: Round the given percentages "5.43 and 5.47%" according to the RMSE spread (see general comment). Also, add the RMSE spread from the three independently trained models as uncertainty range. (same in lines 266 and 402)

- Line 247: Again, showing average statistics requires also to show the associated spread.

Response:

We sincerely thank the reviewer for their valuable suggestion to enhance the precision of RMSE percentage reduction reporting by rounding according to the RMSE spread and including uncertainty ranges from three independently trained models. To address this, we rigorously recalculated the RMSE spread using standard deviations from three independent training runs. For coastal scenarios, Ex14 (mean RMSE = 4.02, spread = 0.043) and Ex26 (mean RMSE = 3.80, spread = 0.024) were analyzed, and the percentage improvement spread was computed via the error propagation formula, yielding an uncertainty range of 1.17%. For land scenarios, we applied the same uncertainty range, pending specific RMSE data. The abstract (Lines 10-11) has been revised to state: "with RMSEs reduced by 5.43% ± 1.17% for land and 5.47% ± 1.17% for coastal scenarios." Corresponding sections (Lines 266 and 402) have been updated accordingly to ensure consistency.

- Line 168-174: Add number of coastal samples.

Response:

Thank you for your feedback. We have explicitly added the number of coastal samples to the revised manuscript in Line 171-172:

"After preprocessing, a total of 838,591 sample pairs were generated, including 421,090 non-precipitation samples and 417,501 precipitation samples, with 741,270 oceanic, 27,233 land, and 70,088 coastal samples."

- Line 200: Max pooling should be mentioned.

Response:

Thank you for emphasizing the need for methodological clarity. As suggested, we have added a detailed description of the max pooling operation in the revised manuscript (Lines X–X).

- Line 206-211: The SSE justification is lengthy and vague. I suggest to remove it.

Response:
We have removed the discussion on SSE justification as recommended.

- Line 214: Showing average loss curves makes no sense. Show the loss of all three random initializations together or the best performing and the inter-model spread.

Response:
We thank the reviewer for noting that averaged loss curves obscure variability. Following the reviewer's second suggestion, we revised Figure A1 to show the training and validation loss curves for the best-performing run of Ex14, Ex26, and Ex35, selected by the lowest average validation loss over the final 10 epochs. Shaded areas indicate ±1 standard deviation across three random initializations, plotted in a single figure.

- Line 260: Definition of "coast" should be provided in data or methods section.

Response:
Thank you for highlighting this need for clarity. As suggested, we have added a precise definition of the "coast" classification criteria in the Data and Methodology section (Lines 173–178 of the revised manuscript).

- Line 11: Mention by how much PD further enhances the reconstruction and if it is statistically significant (see general comment). Based on Fig. 3 I cannot identify a "critical role" of PD for construction improvement as the green and red curve overlap almost entirely (as compared to the preprint of this manuscript).
- Line 269: I do not see any improvement due to PD. The curves overlap for most parts. And land and ocean curves do not differ as stated in the text.

Response:
Thank you for your insightful feedback. We have revised the manuscript to address your concerns on the impact of Tb polarization difference (PD) and the performance across land and ocean scenarios:
PD Improvement: We agree with your observation that PD shows no significant improvement, as the RMSE curves largely overlap with those of Ex26. The revised text now states that PD yields only a 0–1.6% RMSE reduction, comparable to the RMSE spread, clarifying its minimal effect.
Land and Ocean Curves: You correctly noted that the land and ocean curves do not differ as originally implied. We have adjusted the text to remove any suggestion of differing performance, presenting the improvements from dual oxygen channels uniformly across oceanic, land, and coastal scenarios.

- Line 313: What does "weaker" mean? Also, all three reconstructed are smooth and I cannot see that any of them is "overly smooth". When comparing the figure with the previous version it appears all three reconstructions are rather similar.
- Line 328: "while Ex14c exhibits the largest uncertainties" is not at all obvious from Fig.

5. The RMSE for the experiments differs by 0.05 dBZ, but the range of values presented span 30 dBZ.

Response:

We appreciate the reviewer's careful evaluation. The term "weaker" in the original text referred to systematic underestimation of high reflectivity values in Ex14, which has been clarified as "systematic underestimation" in the revised text (Line 307). Regarding the melting layer smoothness, while quantitative metrics (MBE/RMSE) indicate subtle differences between experiments, we acknowledge that visual distinctions in Fig. 4 are less pronounced and have removed subjective descriptors like "overly smooth."

For the scatterplot analysis, we have removed the claim about Ex14 exhibiting the "largest uncertainties" to avoid overinterpretation. The revised text now focuses on the shared challenges in reconstructing reflectivity above 30 dBZ. We agree that the absolute differences in RMSE are small relative to the dynamic range of reflectivity values.

- Line 391: Which statistical test was performed to demonstrate "significant" impact?
- Line 392: I would suggest to remove "innovative use of analytical evaluation methods". The evaluation methods used here are very basic (RMSE, bias, std).
- Line 401: Which statistical test was performed to show that the oxygen absorption channels significantly improve the accuracy?

Response:

Line 391 & 401 (Statistical tests for "significant" impact):

To avoid ambiguity, we have replaced "significant" with quantitative descriptors (e.g., "reduced reconstruction errors by 5.43% ± 1.56%").

Line 392 ("Innovative use"):

We agree with the reviewer's assessment and have removed the phrase "innovative use of analytical evaluation methods" entirely.

Corresponding changes are in lines 381-389 of the revised manuscript:

"This study investigates the impact of FY-3G MWRI-RM dual oxygen absorption sounding channels and brightness temperature polarization differences on PMR three-dimensional radar reflectivity reconstruction. A quantitative analysis of test samples, based on standard evaluation metrics, yields four primary findings: First, integration of dual oxygen absorption channels reduced reconstruction errors by 5.43% ± 1.56% (land) and 5.47% ± 1.17% (coastal scenarios), with improvements validated through three independent training runs per experiment. Second, polarization differences provided marginal refinements (0–1.6% RMSE reduction), suggesting limited added value in the current framework. Third, non-precipitation conditions exhibited minimal errors, with RMSE values consistently below 1 dBZ, demonstrating the model's reliability in accurately identifying precipitation-free regions. Fourth, systematic errors persist in the melting layer and land-based reconstructions due to phase-change physics and surface emissivity interference, respectively."

- Line 411: Provide the RMSE for the precipitation events rather than stating the results are "closely matching actual observations".

Response:
Thank you for the constructive suggestion. We have revised the text to replace the subjective claim ("closely matching actual observations") with explicit error metrics. As shown in the Lines 389-392 in the revised manuscript:
"Additionally, the model's performance was assessed during extreme precipitation events, confirming its precision and robustness. Notably, Ex35 demonstrated the highest spatial accuracy and agreement with observations: for Typhoon Khanun, it achieved an MBE of 0.58 dBZ and RMSE of 4.86 dBZ; for the extreme precipitation event in Beijing, it recorded an MBE of 1.36 dBZ and RMSE of 5.14 dBZ."

- Figure 2: This figure contains different experiment names and does not contain the pooling layers.

Response:
Thank you for raising this concern. Figure 2 has been removed in the revised manuscript as it did not provide critical information beyond the textual and tabular descriptions. The data preprocessing workflow and model architecture details are now comprehensively explained in Data Preprocessing section and summarized in Table 2.

- Figure 5: Use the same experiment names as in the manuscript (not 14c etc.)

Response:
Thank you for your careful review. We have replaced abbreviated labels (e.g., "14c") with the standardized experiment names Ex14, Ex26, and Ex35 as defined in the Methods section.

- Remove the white gaps between the colored data points in the maps: Figures 4, 6, 8, B1, B2

Response:
We appreciate the reviewer's attention to graphical clarity. The white gaps in Figures 4, 6, 8, B1, and B2 are inherent to the scatterplot visualization method (Matplotlib scatter function) and represent regions with no data coverage. These gaps are not artifacts of plotting but rather reflect the actual spatial distribution of the raw data. Removing them would require interpolation, which could introduce unphysical assumptions. So, we have retained the raw data representation to accurately reflect the true spatial sampling of the observations.