

Response to Topical Editor's Comments

Dear Editors,

We sincerely appreciate the reviewer's constructive and insightful comments, which have greatly helped us improve the technical completeness, rigor, and readability of this manuscript. We have carefully addressed all comments point by point, revised the manuscript accordingly, and supplemented relevant experiments, analyses and discussions. The detailed responses in egusphere-2025-5121-track-V5.pdf are as follows.

Reply on RC1:

Comment 1: The authors state that LSM3 slightly outperforms LSM2. However, from a theoretical perspective, weighted ordinary least squares (LSM2) is mathematically equivalent to ordinary least squares with weighted variables (LSM3). Could you please explain why a performance difference is observed in practice?

Response:

Thank you for your insightful comment. It is important to clarify that in our actual implementation, the weight matrices employed by LSM2 and LSM3 are not identical. They are not strictly equivalent in practice due to the differing weight definitions.

a) Different cost functions

In our implementation, the weights $W_j = \sqrt{\tau_j * \tau_j}$ depend on the transmittance τ_j , which is itself a function of the fitted optical depths.

The two weighted least squares methods are theoretically equivalent, yet they differ notably in mathematical formulations. The cost function of LSM2 (Weighted Least Squares Method 1) is $S = \sum_{m=1}^p W_m (y_m - \hat{y}_m)^2$, which directly incorporates linear weights into the residual sum of squares. The cost function of LSM3 (Weighted Least Squares Method 2) is $S = \sum_{m=1}^p W_m^2 (y_m - \alpha_m X_m)^2$, which linearly preprocesses raw data with weights before standard ordinary least squares. Differences in matrix construction, numerical stability and error propagation lead to distinctions in practical solving processes and results. **At line 204, we have added the cost function expression for the second weighted least squares method**, in order to distinguish it from the cost function expression for the first weighted least squares method given at line 177.

b) Different numerical solvers

LSM2 is typically implemented via the normal equations (e.g., Cholesky decomposition of $X^T W X$), which is known to amplify rounding errors when the matrix is ill-conditioned. LSM3, however, first transforms the data and then employs standard OLS solvers based on QR decomposition or SVD, which are far more robust against near-singularity.

c) Propagation of rounding errors in extreme cases

When transmittance values approach machine precision (e.g., very opaque layers), the computed weights may suffer from rounding errors. In LSM2, these errors are squared and directly affect the regression equations; in LSM3, they are only linearly scaled. As a result, the two methods may converge to slightly different fixed points after several iterations, with LSM3 generally closer to the true WLS solution.

In summary, although LSM2 and LSM3 are mathematically equivalent in the exact arithmetic sense, their numerical behaviors may diverge in practice due to differences in finite-precision computations, the formulation of the cost function, and solver robustness. Therefore, the slight superiority of LSM3 observed should be attributed to numerical factors rather than to any statistical distinction. We hope this explanation clarifies the issue, and we once again sincerely thank you for your careful review and valuable comments.

Comment 2: Page 5, Line 156 & Page 6, Line 168: Maybe change 'when ...' to 'When...'

Response: Thank you for your careful proofreading. You are correct. For consistency and proper English sentence structure, the word "When" should start with a capital letter as it begins a new sentence following a period.

We will make the following corrections:

- **Page 5, Line 156:** Change 'when ...' to 'When...' at line 157.
- **Page 6, Line 168:** Change 'when ...' to 'When...' at line 170.

Comment 3: Page 6, Eq. (6): The weight factor is defined as a function of transmittance (i.e., $W \propto \tau$). Is this derived from an iterative optimization procedure, or other physical assumptions? Could you please add a description to clarify the principle behind this specific choice?

Response: Thank you for this insightful comment. Regarding the rationale for choosing the weight factor $W_j = \sqrt{\tau_j * \tau_j}$ in Eq. (6), we provide the following clarification, which will be added to the revised manuscript.

From a physical perspective, from the top of the atmosphere (TOA) to the surface, as the transmittance from a given layer to space decreases, the layer lies closer to the surface. For such layers, the contribution of their absorption capability to the total atmospheric absorption becomes smaller. This is because the overlying atmosphere has already absorbed a significant portion of the incoming radiation, leaving less energy to be absorbed by the lower layers. Consequently, for layers with small transmittance (i.e., near-surface layers with weak signal), their confidence weight in the regression fitting should be reduced accordingly.

Based on this physical understanding, we select a function of transmittance as the basic form of the weight coefficient. After analyzing the characteristics of transmittance variation and comparing the effects of different weight functions (such as τ_j , $|\tau_j|$, τ_j^2 , $\sqrt{\tau_j}$, etc.) on the regression results, we ultimately adopt the total transmittance from the layer to space τ_j as the weight factor in the weighted least squares method, i.e., $W_j = \sqrt{\tau_j * \tau_j}$. This choice ensures that layers with low transmittance (near-surface, small absorption contribution, low signal reliability) receive lower weights, while layers with high transmittance (near TOA, high signal reliability) receive higher weights. This effectively improves the fitting performance of the regression model under heteroscedastic conditions.

We have added the above explanation to the revised manuscript. **This explanation (lines 186–198)** clearly states the physical basis and selection process of the weight factor.

Comment 4: Page7, Lines 202-214 & Lines 215-220: This part needs to be carefully revised. It appears to be a repetition of content, where two consecutive paragraphs convey essentially the same meaning with slightly different wording.

Response: We thank the reviewer for carefully reading the manuscript and pointing out this issue. We agree that the two consecutive paragraphs on page 7 (lines 202–214 and lines 215–220) are largely repetitive and should not have been presented separately. In the revised manuscript, we have consolidated these two paragraphs into a single, concise paragraph. The redundant content has been removed, and the key information is now presented clearly without repetition. **The revised text is shown on page 7, lines 224–229 (revised version).** We apologise for this oversight and appreciate the reviewer's constructive suggestion.

Comment 5: Page 13, Line 331-332: The authors used a semi-empirical threshold of 10^{-4} as the screening criterion. However, it is not entirely clear whether this threshold is specifically selected for the FY-3F HIRAS-II instrument based on dedicated analysis, or adopted from previous studies as a generally applicable value. Could you please provide more quantitative evidence to support this choice if possible?

Response: Thank you for this valuable comment. The 10^{-4} semi-empirical threshold is not an empirical constant borrowed directly from existing literature. Instead, it is quantitatively determined via systematic comparative tests targeting the FY-3F HIRAS-II interferometric spectrometer. Detailed quantitative supporting evidence is summarized as follows:

a) Numerical risk analysis of unphysical negative transmittance

The spectral response function (SRF) of HIRAS-II is characterized by oscillatory side-lobe values oscillating around zero. After convolution with monochromatic transmittance spectra, convolution artifacts readily yield unphysical negative transmittance within strong absorption bands (e.g., intense water vapor absorption channels near 1506.875 cm^{-1}). Our layer-by-layer calculations reveal that once the upward cumulative transmittance from a given atmospheric layer to TOA falls below 10^{-4} , the cumulative transmittance of the underlying layer tends to turn negative. This numerical artifact originates from spectral convolution effects and small-number division operations. These unphysical negative values violate physical laws and need to be filtered out via threshold screening.

b) Quantitative comparison thresholds via LBLRTM reference truth values

We performed comparative calculations using two candidate thresholds (10^{-4} , 5×10^{-5} and 10^{-5}) taking the 1506.875 cm^{-1} water vapor channel as the test case, where high-precision LBLRTM results are defined as true transmittance benchmarks:

- When adopting the 10^{-5} threshold, the fast transmittance deviates obviously from LBLRTM true values owing to severe small-number division errors in regression;
- After raising the threshold to 10^{-4} , the regression precision of layer-wise fast transmittance coefficients is markedly improved and the deviation against LBLRTM is largely suppressed.

The mean RMSE for the mid-wave spectrum (1300–1900 cm^{-1} 465) of LSM1.2 (10^{-5}), LSM1.1 (5×10^{-5}), and LSM1 (10^{-4}) progressively decreases from 0.17 K to 0.12 K and 0.1 K. Figures A1 to A3 have been added to the Appendix to demonstrate the results under different threshold values, respectively. Figures A1 through A3 in Appendix A present the error comparisons between the LSM series models (LSM1, LSM1.1, LSM1.2) with varying thresholds and the LBL line-by-line model (true value) under typical layer, full atmospheric column, and strong absorption channel conditions, respectively. In each figure, the subplots are arranged as follows: the left panel corresponds to the LSM1 model with a threshold of 10^{-4} , the middle panel corresponds to the LSM1.1 model with a threshold of 5×10^{-5} , and the right panel corresponds to the LSM1.2 model with a threshold of 10^{-5} .

c) Trade-off analysis on excessively large thresholds and atmospheric sample loss

Although a threshold larger than 10^{-4} can further reduce small-value numerical anomalies, a higher threshold corresponds to an elevated cutoff altitude (lower cutoff pressure). For strong water vapor absorption channels, the atmospheric layers below the cutoff height will be forced to fix transmittance as unity due to insufficient training samples, which introduces prominent calculation errors for mid-to-lower atmosphere transmittance of dry atmospheric profiles. By contrast, weak absorption channels feature slowly decaying transmittance and large bottom-layer cumulative transmittance, where the 10^{-4} threshold barely damages regression accuracy.

In summary, the 10^{-4} threshold is finalized after quantitative error comparison against LBLRTM results, SRF convolution error diagnosis, and the balance between numerical stability and effective retention of atmospheric samples, which is purpose-oriented for the instrumental features of FY-3F HIRAS-II rather than a universal empirical value cited from former publications. **The revised text is shown on page 14, lines 355–370 (revised version).**

Comment 6: Figure quality: Most of the comparison figures (e.g., Figs. 5, 6, etc.) appear noticeably blurred and contain more than 15×3 lines in a single image, making it difficult for readers to distinguish between different schemes and symbols. To improve readability, I suggest select a limited number of representative profiles (e.g., Profiles 81, 82, 83) to illustrate the key findings, and move others to the appendix or supplementary material.

Response: We appreciate the reviewer’s valuable comments. The blurriness of these figures resulted from resolution loss when the original files were saved. In the revised manuscript, we have enhanced the image quality to improve clarity. Following your suggestion, **we have kept only three representative profiles (81, 82, and 83) in the main figures to present the core results in Section 4. The remaining profile plots have been moved to Appendix A (pages 25–26).**

Comment 7. Figure caption: Different methods are described using words in Fig. 5 (e.g., ‘pentagrams’, ‘circles’, ‘triangles’), while symbols in Fig. 9 (e.g., ‘★’, ‘○’, ‘△’). For reasons of consistency, please unify the notation style (i.e., either textual descriptions or symbols).

Response: We appreciate this valuable comment. Following your suggestion, all figure captions in the revised manuscript have been uniformly revised to adopt textual descriptions to denote marker styles instead of graphic symbols, achieving consistent notation throughout all figures. **Figures 9, 10, 12, and 13 have been revised:**

Figure 9: Comparison of fast transmittance calculated by LSM0 (pentagrams), LSM1 (circles), LSM2 (crosses), and LBLRTM (triangles) across three channels (1380 cm^{-1} , 1423.75 cm^{-1} , 1506.875 cm^{-1}) versus pressure.

Figure 10: Comparison of fast transmittance calculated by LSM0 (pentagrams), LSM1 (circles), LSM2 (crosses), and LBLRTM (triangles) across three channels (1646.875 cm^{-1} , 1730 cm^{-1}) versus pressure.

Figure 12: Comparison of fast transmittance calculated by LSM0 (pentagrams), LSM1 (circles), LSM2 (crosses), LSM3 (**diamonds**) and LBLRTM (triangles Δ) across three channels (1380 cm^{-1} , 1423.75 cm^{-1} , 1506.875 cm^{-1}) versus pressure.

Figure 13: Comparison of fast transmittance calculated by LSM0 (pentagrams), LSM1 (circles), LSM2 (crosses), LSM3(**diamonds**) and LBLRTM (triangles) across two channels (1646.875 cm^{-1} , 1730 cm^{-1}) versus pressure.

Reply on RC2:

Comment 1: The manuscript would benefit from a clearer explanation of the physical and numerical basis for the 10^{-4} transmittance threshold. The current description is plausible, but a compact sensitivity test using a few nearby threshold values would make the choice more convincing. It would be useful to show whether the retained sample numbers, transmittance errors, and brightness-temperature errors remain stable when the threshold is modestly varied.

Response: We sincerely thank the reviewer for raising this critical concern, which aligns closely with the comments from Reviewer #1. We fully acknowledge that the selection of the 10^{-4} threshold requires robust justification. To further validate this choice, we have added **Figures A1 to A3 in Appendix A**, which respectively illustrate the error comparisons among the LSM series models under different threshold settings: **LSM1 at 10^{-4} (left panels)**, **LSM1.1 at 5×10^{-5} (middle panels)**, and **LSM1.2 at 10^{-5} (right panels)**. These comparisons are conducted across a typical layer (Fig. A1), the full atmospheric column (Fig. A2), and strong absorption channels (Fig. A3). The results unequivocally demonstrate that adopting a threshold of 10^{-5} leads to significant deviations from the LBLRTM truth values, primarily caused by numerical truncation errors in the regression division operations. In contrast, when the threshold is raised to 10^{-4} , the numerical stability of the regression coefficients is remarkably enhanced, and both transmittance and brightness temperature errors are suppressed to extremely low levels across the entire spectral band.

While applying a threshold larger than 10^{-4} (e.g., 10^{-3}) could further minimize numerical anomalies, it results in an overly elevated cutoff altitude. For strong water vapor absorption channels, the atmospheric layers below this cutoff height lack sufficient training samples, forcing the transmittance to be artificially fixed to unity. This leads to prominent simulation errors in the mid-to-lower atmosphere, particularly for dry atmospheric profiles, and fails to accurately represent the FY-3F HIRAS-II's capability to sense the lower troposphere. Consequently, the 10^{-4} threshold represents the optimal balance between eliminating non-physical numerical artifacts, ensuring regression stability, and maximizing the retention of effective atmospheric training samples.

In summary, the 10^{-4} threshold was rigorously determined through error comparisons against LBLRTM, physical SRF diagnostics, and a comprehensive trade-off analysis. It is an instrument-specific optimization for FY-3F HIRAS-II rather than a generic empirical value. The corresponding quantitative discussions and revisions have been systematically incorporated into the revised manuscript on **page 14, lines 355–370**.

We truly appreciate the reviewer's constructive suggestions on improving the robustness of our methodology.

Comment 2: The weighted least-squares schemes LSM2 and LSM3 are promising, but the mathematical description could be made easier to reproduce. Please clarify exactly how the weight factor is defined and whether it is applied to the predictors, the dependent variable, or both before solving the regression. A short algorithmic summary of LSM0-LSM3 would help readers follow the differences among the schemes.

Response: In response to the reviewer's comments, we have thoroughly revised and strengthened the mathematical descriptions of the weighted least-squares schemes LSM2 and LSM3 in Section 2.3 of the revised manuscript. Specifically, we have clearly delineated the differentiated weighting principles of LSM2 and LSM3, rigorously defined the exact formulas for their respective weight factors, and explicitly stated that the weighting strategy in our study is applied to both the independent and dependent variables, thereby ensuring the physical consistency of the regression fitting.

Furthermore, to help readers better understand the differences among the four schemes, we have added a detailed algorithmic summary of **LSM0–LSM3 in lines 241–248 of the revised manuscript**. Concurrently, we have comprehensively **revised Table 1 to systematically outline** the core differences in weighting definitions, implementation logic, and applicable scenarios among

the ordinary least-squares method (LSM0), the threshold least-squares method (LSM1), and the two improved weighted least-squares methods (LSM2 and LSM3). The combination of this unified text description and the detailed table presentation enables readers to intuitively grasp the specific implementation steps and distinctive features of each algorithmic scheme.

Comment 3: Please clarify the RTTOV baseline used for comparison. The title refers to RTTOV v13.1, whereas the validation section mentions RTTOV v13.2. It would be helpful to state the coefficient set, predictors, surface-emissivity treatment used in the baseline comparison so that readers can interpret the magnitude of the reported improvement.

Response: We apologize for the version inconsistency caused by manuscript revision negligence. Our primary algorithm development, coefficient training, and parameter tuning were all completed based on RTTOV v13.1, which is why this version was cited as the baseline in the title and original draft. During the subsequent validation phase, we briefly tested the officially released RTTOV v13.2 update. However, after detailed comparison, we confirmed that v13.2 only includes bug fixes related to file reading and memory management, and does not alter the core optical-depth coefficient sets, predictor parameterization, or radiative transfer physics of the v13 main branch. We have unified the RTTOV baseline version throughout the manuscript as RTTOV v13.1 and revised the title, abstract and all relevant text to eliminate version confusion. Meanwhile, **we have supplemented detailed baseline configuration parameters in the revised manuscript at lines 471-477**, including the official standard coefficient set matching HIRAS-II, unified predictor variable settings, and the fixed surface emissivity processing scheme adopted in the comparison experiment. All baseline parameters are now explicitly stated to ensure the comparability and interpretability of the experimental results.

Comment 4: The observational comparison with FY-3F HIRAS-II is valuable. Because OMB statistics can also be influenced by clear-sky screening, ERA5 profile errors, surface emissivity, calibration uncertainty, and residual cloud contamination, I suggest adding a short uncertainty discussion. This would make the interpretation of the approximately 0.1 K improvement near the absorption peak more balanced.

Response: We sincerely thank the reviewer for this insightful and constructive suggestion. Following your advice, we have added a systematic uncertainty discussion in the revised manuscript to provide a more balanced and objective interpretation of the approximately 0.1 K improvement achieved by the LSM3 model.

We quantitatively assessed the uncertainty from five major sources: **(1)** limitations of clear-sky screening; **(2)** ERA5 reanalysis profile errors; **(3)** surface emissivity uncertainties (where the standard deviation over land is notably higher than over ocean, as shown in Table 2; **(4)** thin cloud contamination, which tends to introduce a cold bias.

We believe the added uncertainty discussion provides a more balanced physical interpretation of the 0.1 K improvement near the absorption peak. The corresponding revisions have been incorporated into the revised manuscript on **page 23, lines 526-556**. We thank the reviewer for helping us improve the clarity and rigor of our validation.

Comment 5: The clear-sky selection method based on collocated MERSI cloud mask products is a sensible choice. Please provide a little more information on the number of retained HIRAS-II observations, land-ocean distribution, and whether the main conclusions are sensitive to the strict requirement that all MERSI pixels within the HIRAS-II FOV be clear.

Response: We sincerely thank the reviewer for this constructive suggestion. In the revised manuscript, we have supplemented detailed statistical information on the retained clear-sky samples in Tabel 2. After applying the strict criterion that all MERSI pixels within the HIRAS-II FOV must be clear, we retained a total of approximately 120,000 valid HIRAS-II clear-sky observations. It should be noted that, since both the RTTOV and LSM3 models utilize the exact same set of strictly screened clear-sky observations, the sample counts for both models are identical for each corresponding period. Regarding the distributional differences, the ocean and land sample sizes are nearly balanced in August 2024. In contrast, for February 2015, the number of land samples (82,102) is significantly larger than that of ocean samples (37,766). If you need to cite these data in the main text, it is recommended to directly use these precise figures.

Furthermore, to test the sensitivity, we relaxed the strict screening criterion to allow up to 5% cloudy MERIS pixels within the FOV. The comparative results show that although relaxing the threshold leads to slightly larger simulation biases for both schemes, the relative improvement of LSM3 over RTTOV remains unchanged. This stability is primarily because the fast transmittance model is specifically designed and validated for clear-sky, low-to-mid latitude conditions. At high latitudes, additional complexities arise from both observation geometry and surface emissivity, which introduce larger uncertainties. Consequently, for low-to-mid latitude regions, stricter cloud-screening criteria consistently yield smaller bias between simulations and observations. Therefore, even if the screening threshold is relaxed, the impact on our core conclusions remains negligible. These tests unequivocally confirm that the simulation accuracy improvements and the core conclusions drawn in our study are robust and insensitive to the strictness of the clear-sky selection. **The corresponding revisions have been incorporated into the revised manuscript at lines 557-567.**

Comment 6: AI-based approaches are increasingly being explored in fast radiative-transfer and gas-absorption calculations (e.g., <https://doi.org/10.1016/j.jqsrt.2026.109920>). It would be interesting if the authors could briefly discuss whether the threshold and weighting strategy proposed here might be combined with AI-based approaches in future work to further improve computational efficiency, robustness, or adaptability for strong-absorption-channel simulations.

Response: We appreciate this forward-looking suggestion. We have added a discussion on future interdisciplinary research in the conclusion section of the revised manuscript. **We pointed out that the physical threshold screening and adaptive weighted regression strategies proposed in this study can be perfectly combined with AI and machine learning algorithms at lines 626-634.**

Comment 7: Please correct grammatical and typographical issues throughout. Examples include 'poorly profiles', repeated sentences around the weighting function, and missing spaces before some section headings.

Response: We have conducted a full systematic proofreading of the entire manuscript. We corrected all grammatical errors, typos and inappropriate expressions including the incorrect phrase "poorly profiles" **at lines 153-154**. We deleted duplicate sentences related to the weighting function derivation and description **at lines 234-235**, and standardized the formatting of all section headings to fix the missing space problem. The overall writing fluency and standardization of the manuscript have been significantly improved.

Comment 8: Please ensure consistent terminology for HIRAS, HIRAS-II, FY-3E, FY-3F, RTTOV v13.1, and RTTOV v13.2.

Response: We have comprehensively sorted and unified all professional terminology and instrument/software version names throughout the manuscript. We completely eliminated the confusion of RTTOV, and fully unified the **RTTOV version as v13.1** in all sections to ensure consistent and standardized terminology in the full text. Because FY-3D uses HIRAS-I whereas FY-3E/F use HIRAS-II, inconsistencies exist between generations. However, since our method is applicable to the entire FY-3 HIRAS series, we have not differentiated between the two generations in some places.

Comment 9: The rows in Table 1 are not visually well separated, which makes it somewhat difficult to distinguish the different experimental schemes at a glance. The authors may consider improving the table formatting, for example by adding clearer horizontal rules, alternating row shading, or slightly more spacing between rows.

Response: We have fully optimized the typesetting format of **Table 1**. We added standard horizontal dividing lines and appropriately increased the row spacing. The optimized table has a clear hierarchical structure and intuitive visual effect, which facilitates readers to quickly distinguish and compare different LSM scheme parameters and characteristics.

We have made the following revisions to the manuscript (version 6). We believe these revisions fully address your concerns and significantly improve the clarity and precision of the manuscript, especially for readers of a specialist journal.

We have uploaded the revised manuscript with these changes clearly marked.

Thank you again for your guidance and for handling our manuscript.

Sincerely,

The Authors