

Dear Editor and Reviewers,

We sincerely thank the editor and both reviewers for their careful reading and constructive comments. We have revised the manuscript in response to each comment. The revised manuscript has been improved in methodological clarity, statistical analysis, physical interpretation, figure presentation, literature context, and language.

Our point-by-point responses and the corresponding locations of the revisions are provided below.

Reviewer 1

1. The manuscript contains numerous non-idiomatic expressions, grammatical inaccuracies, and awkward phrasings that reduce its readability and professionalism. I strongly recommend a thorough language polish by a native English speaker or a professional editing service before resubmission.

We have thoroughly edited the manuscript to improve grammar, idiomatic expression, concision, consistency of tense, and technical terminology. We have also revised the figure captions and corrected non-idiomatic phrasing throughout the manuscript to improve clarity and overall presentation.

Changes in the manuscript: Throughout the manuscript.

2. The discussion section would be strengthened by providing deeper physical interpretations of the observed biases. For example, the pronounced bias at 700 hPa deserves more specific explanation — possibly linked to its role as a key interface for weather systems or the characteristics of the instrument's weighting functions.

The previous relative-deviation index could not directly represent the magnitude of the bias. After replacing it with the mean signed bias, the 700 hPa result was no longer exceptional. We removed the former speculative explanation. The corrected analysis instead identifies pronounced negative biases at 850 and 920 hPa, particularly during daytime.

We added a focused interpretation for these lower levels. Surface-temperature and spectral-emissivity uncertainties may propagate into the retrieval. Broad weighting functions and limited vertical sampling may smooth strong lower-atmospheric temperature gradients and warm anomalies. Water-vapor absorption or residual thin-cloud contamination may reduce infrared sensitivity, especially in humid regions.

Changes in the manuscript: Abstract, Introduction, Sections 2.3.2, 3.2, 4.1, and 4.3, Table 1, and Conclusion.

3. All figure titles and axis labels should be reviewed for clarity and precision. For instance, the Y-axis label in Figure 6 (“LTS from MOD07_L2 (K)”) could be more clearly expressed as “MODIS-derived LTS (K)”.

We reviewed figure titles, axis terminology, units, and captions. The LTS terminology now uses "sounding-derived LTS" and "MODIS-derived LTS", and the temperature-validation captions use standardized wording and clear day/night panel assignments.

Changes in the manuscript: Fig. 7 and captions for Figs. 3–10.

4. While the introduction cites relevant background literature, it could be further strengthened by directly referencing studies most closely related to this work. Incorporating specific regional

validations of MODIS temperature and water vapor profiles in China (e.g., Ouyang et al., 2015 in the Heihe River Basin) would better contextualize the present contribution and highlight its added value.

We expanded the Introduction with regional Chinese remote-sensing studies, including Ouyang et al. (2015) on thermal-infrared land-surface-temperature retrieval in the Heihe River Basin and Zhu et al. (2017) on regional daytime air-temperature retrieval from MODIS products. We then clarify how these surface or near-surface applications differ from the present long-term, nationwide validation of complete MOD07_L2 temperature profiles and derived LTS/TIL metrics.

Changes in the manuscript: Introduction and References.

Reviewer 2

1. Temporal matching window: The choice of $|\Delta t| < 2$ hours is justified by sample size vs. accuracy trade-off, but the sensitivity analysis (Fig. 2) shows noticeable degradation in some regions (e.g., HHH daytime bias -1.77°C). Please add a brief sentence in Section 2.3.2 or 3.1 quantifying the overall impact on national-scale statistics and explicitly state why <2 h remains acceptable despite regional variations.

We expanded the sensitivity analysis to quantify both sample coverage and national-scale accuracy. Extending the window from $|\Delta t| < 1$ h to $|\Delta t| < 2$ h increased the matched records from 554,128 to 894,061 during daytime and from 299,578 to 711,344 at nighttime.

Across the nine pressure levels, daytime r , B , and RMSE changed from 0.961, -1.00°C , and 3.22°C to 0.963, -0.91°C , and 3.18°C . Nighttime values changed from 0.958, -0.46°C , and 2.76°C to 0.968, -0.37°C , and 2.73°C . Although errors increased in a few regions, the national-scale statistics did not show a nationwide deterioration. Instead, the national-mean error metrics decreased slightly. We therefore retained $|\Delta t| < 2$ h criterion while explicitly acknowledging the regional daytime sensitivity in the revised manuscript.

Changes in the manuscript: Section 3.1 and Fig. 2.

2. LTS calculation: Clarify whether surface pressure from MOD07_L2 is used directly for potential temperature calculation or if any quality control/filtering is applied (e.g., cloud contamination flags). This is important because LTS is a derived variable sensitive to near-surface values.

We clarified that surface pressure is taken directly from the MOD07_L2 Surface_Pressure data set (hPa) at 5 km resolution. Fill values and invalid records are excluded. No external surface-pressure product or additional pressure-specific correction is used. Cloud screening follows the standard clear-sky screening inherent in MOD07_L2.

For sounding profiles, the near-surface layer is defined as the lowest-altitude, highest-pressure valid observation. These clarifications make the calculation of near-surface potential temperature and LTS reproducible.

Changes in the manuscript: Sections 2.2.2 and 2.3.4.

3. In abstract “absolute retrieval accuracy was generally superior at nighttime” — consistent with results, but specify that this holds mainly for absolute temperature, while LTS performs better daytime. Add one short sentence on the 700 hPa bias, as it is highlighted as noteworthy in the discussion. “For TIL, MOD07_L2 displayed limited detection capability, achieving an overall detection rate of 6.6% during daytime and 13.1% at nighttime.” — Excellent; no change needed.

The abstract now distinguishes the two findings: MOD07_L2 temperature-profile retrievals are generally more accurate at nighttime, whereas MODIS-derived LTS is more accurate during daytime.

The previous layer-to-layer comparison used a relative-deviation index, $R = \frac{\sum T_{ture}}{\sum T_{measured}} - 1$. Because this ratio cannot directly represent the magnitude of an absolute temperature bias or be compared across pressure levels in the former manner, we replaced it with the mean signed bias and recalculated all statistics. We also added the formulas for the statistical metrics in the Methods section.

After correcting the bias metric from a relative-deviation index to the mean signed bias and recalculating all statistics, the 700 hPa level no longer exhibits an exceptional deviation, daytime/nighttime B values are $-0.52/-0.40$ °C and RMSE values are $2.69/2.57$ °C. By contrast, 850 and 920 hPa have consistently larger errors. During daytime, their B values are -2.03 and -2.14 °C and their RMSE values are 3.74 and 4.11 °C. At nighttime, B is -1.62 and -1.37 °C and RMSE is 3.42 and 3.18 °C. We therefore removed the former statement and explanation concerning 700 hPa and instead identify the pronounced negative biases at 850 and 920 hPa.

Changes in the manuscript: Abstract, Sections 2.3.2, 3.2, and 4.1, Table 1, and Section 5.

4. Equation (2) and (3): Clearly define all variables the first time they appear (e.g., Γ as lapse rate). In 2.3.3, specify the exact formula used for potential temperature (include the $\kappa = R/C_p$ value explicitly). TIL detection (Eq. 6): Define “co-detected” more precisely — does it require the same pressure level or just presence in the same profile?

We now define γ as the temperature lapse rate with respect to pressure and define the adjacent sounding pressure levels, their temperatures, the target MOD07_L2 pressure level, and the interpolated sounding temperature immediately after Eqs. (2) and (3).

The potential-temperature method now gives $R = 287.05 \text{ J kg}^{-1} \text{ K}^{-1}$, $C_p = 1004 \text{ J kg}^{-1} \text{ K}^{-1}$, $\kappa = R/C_p \approx 0.286$, and $p_0 = 1000 \text{ hPa}$. We also clarify that co-detection is profile-based: both matched profiles must contain at least one inversion, but the inversions are not required to occur at the same pressure level.

Changes in the manuscript: Sections 2.3.1, 2.3.4, and 2.3.5.

5. Section 4.1 and 4.2 are strong, but condense repetitive explanations of why lower layers perform worse (surface influence, terrain, clouds). Merge similar sentences. In 3.2 and 3.3, some sentences are repetitive (e.g., “better in upper layers”). Streamline.

We rewrote Sections 3.2 and 3.3 to report the principal numerical results and spatial patterns directly, with emphasis on the 850 and 920 hPa findings. We also condensed Sections 4.1 and 4.2 by removing repetitive explanations and organizing the discussion around surface effects, terrain, vertical-gradient smoothing, moisture, and nighttime thermal structure.

Changes in the manuscript: Sections 3.2, 3.3, 4.1, and 4.2.

6. Figures 3–6: The density scatter plots are useful, but consider adding a summary table of national-averaged r , B , RMSE for each pressure level (day/night) to improve readability. **Fig. 7–10:** Legends and color scales should be consistent across day/night panels. **Fig. 1:** The agricultural zoning link is given — ensure the figure clearly labels all nine zones (NEP, NAS, etc.). **All figures:** Improve caption completeness (include what “ ΔP ” exactly means). Consider moving some detailed spatial maps to supplementary material if the journal has length limits.

We added Table 1 with national-scale r , mean bias B , and RMSE for all nine pressure levels during daytime and nighttime. We also standardized legends, units, panel organization, and color-scale use in paired day/night figures.

Fig. 1 now labels all nine agricultural zones. ΔP is defined at first occurrence as $P_{surface} - P_{target}$, and the captions clarify the day/night panels and displayed metrics. The spatial maps illustrate regional heterogeneity, which is central to the nationwide validation. The associated text was condensed to control length.

Changes in the manuscript: Table 1, Figs. 1 and 3–10, and Sections 3.3 and 3.4.

7. Several awkward phrasings and minor grammatical issues (common in non-native writing). Examples: can represented key temperature” → “can represent” “20203.3” (typo in Figure 4 caption) “the B in some regions became larger: during daytime...” — improve flow. “no inversion layers were detected on the TP and in the SB” — “were not detected”. **Consistent terminology:** Use “MOD07_L2” uniformly; avoid switching between “retrievals”, “retrieved temperatures”, etc.

We completed a thorough language revision throughout the manuscript. The cited grammar and caption errors were corrected, the time-window description was rewritten using complete sentences and regional B /RMSE values, and the inversion statement was revised. We also standardized the use of MOD07_L2, MOD07_L2 temperature retrievals, MODIS-derived LTS, and sounding observations.

Changes in the manuscript: Throughout the manuscript, including Sections 2.2.2, 3.1, and 3.5 and figure captions.

8. Some citations appear slightly outdated or generic (e.g., several 2025 papers). Ensure all are correctly formatted and accessible.

We reviewed the relevance and accessibility of the cited sources, updated the MODIS algorithm document to the Collection 6, Version 7 ATBD, removed references that did not directly support the revised interpretation, and added directly relevant retrieval studies. The Introduction now also includes Ouyang et al. (2015) and Zhu et al. (2017) to better position this nationwide profile validation in the context of regional studies in China.

Changes in the manuscript: Introduction, Discussion, and References.

9. Check DOIs and formatting consistency. A few recent citations (2025) should be verified for final publication status.

We checked DOI links and reference-format consistency, corrected inconsistent hyperlink formatting, removed a duplicated DOI in the Barlow et al. (2015) entry, and completed the bibliographic information and DOI links for the newly cited studies.

Changes in the manuscript: References.

Sincerely,

Yuanjian Yang

On behalf of all authors