The manuscript mainly assembles previously published approaches for angular correction, cloud-gap filling, and downscaling into a production workflow for FY-4A LST. The methodological novelty is limited, several key assumptions are questionable, and the writing quality is low. The stacked machine-learning workflow likely propagates errors, while validation remains geographically limited.

Response: Thank you very much for your thorough review of our manuscript and for your insightful and professional comments, which have significantly improved our work. We apologize that the novelty of this study was not clearly stated and that the main assumptions were not sufficiently presented in the original version. In response, we have undertaken substantial revisions to the manuscript, focusing on the following key aspects:

(1) The Introduction has been rewritten to better highlight the novelty of our hourly, angular-normalized, all-weather, 0.01° LST dataset, which has long been sought by the thermal infrared community.
(2) The main assumptions and their associated limitations have been clearly clarified in the revised manuscript. The ignored nighttime angular effect, imperfect ATC modeling, and isotropic cloud radiative force correction were further clarified and discussed in subsection 5.5.
(3) Pixel-level uncertainties associated with the machine learning models have been explicitly quantified. The average RMSE is 2.49 K for the cloud-filling process and 2.28 K for the spatial downscaling step. In addition, the Results section has been reorganized to more clearly present and distinguish the performance at each processing stage.
(4) A more rigorous temperature-based validation has been conducted by incorporating OzFlux sites as you suggested. Updated results show that the RMSE and MBE is 2.48 K and -0.83 K under clear-sky conditions, as well as 3.65 K and -1.43 K under cloudy-sky conditions.
(5) The manuscript has been comprehensively polished to improve overall clarity and readability.

All comments have been addressed point by point as detailed below. (Red text indicates our responses, and blue text indicates the corresponding revisions in the manuscript.)

Major comments

1. **Incorrect use of a sinusoidal annual temperature cycle**

The method assumes a sinusoidal annual temperature cycle to generate gap-free LST fields and to support downscaling. This assumption may work at the higher latitude sites selected in the study, but it is fundamentally flawed at large spatial scales. Many regions, particularly in monsoon and low-latitude climates, exhibit multiple seasonal temperature peaks or asymmetric annual cycles. Previous studies have shown that it is difficult to design a universally valid ATC model across broad latitudinal gradients. Using a simple sinusoidal ATC as the backbone of cloudy reconstruction and downscaling is therefore physically incorrect.

Response: Thank you for your insightful comments. We're sorry that the uncertainty caused by the imperfect ATC modeling is not fully clarified. Here, we present the limitation of the ATC model, its refinement using machine learning model, and an examination of the LST annual variation in tropical region.

(1) Imperfect modeling of the ATC model

We fully agree that the sinusoidal ATC model performs poorly in tropical regions, where the land surface temperature (LST) typically exhibits twice annual LST peaks. As you pointed out, developing a universal

ATC model applicable across the full FY-4A disk is inherently challenging, and any simplified representation will inevitably introduce uncertainties. To address this limitation, the ATC-derived LST is not directly used as the clear-sky LST or as the final downscaled LST in our product generation process. Instead, we adopted a two-step hybrid strategy to refine the ATC model, which taking ATC-derived LST as an initial approximation followed by residual modeling via machine learning models.

(2) Refinement of the ATC model with machine learning methods

A common solution to refine ATC modeling is estimating its residual (i.e., the bias between ATC-simulated LST and actual LST) via data-driven machine learning methods. In this framework, the ATC model serves primarily to provide thermal texture and a physically meaningful initial temperature field, while the machine learning component captures the residual discrepancies caused by climate fluctuations. This hybrid strategy (i.e., using ATC as an initial approximation followed by residual modeling) has been widely adopted in the thermal infrared remote sensing community as summarized in Table R1.

Table R1. Recent studies using sinusoidal ATC model, and followed by residual modeling.

Research	Description
Zou et al., 2018; Xia et al., 2021; Liu et al., 2019; Quan, 2023	Simultaneous calibrating ATC model and a regression model by incorporating vegetation index and air temperature, which can capture the LST variation caused by vegetation phenology and local climate condition.
Yang et al., 2024; Ding et al., 2022; Zhu et al., 2022; Tang et al., 2024	Taking ATC-simulated as initial value, and then estimate climate fluctuations using auxiliary datasets and machine learning models.

In this study, the ATC-derived LST act as the thermal-infrared texture for reconstructing hypothetical clear-sky LST and downscaled LST as shown in Eq. R1-2. The climate fluctuations were represented using auxiliary reanalysis data. Under this framework, the predicted LST has potential to reflect bimodal annual variations, as the input climate-related variables exhibit bimodal pattern.

$$T_{clr} = f_{clr}\left(DOY, \theta_s, Lat, Lon, LCT, CT, FVC, DEM, T_{ATC,d1}, T_{ATC,d2}, T_{ATC,n1}, T_{ATC,n2}, R_{in}, T_{ERA5-skin}, T_{2m}, D_{2m}\right) \quad \text{(R1)}$$

$$\Delta LST_{0.05} = f_{downscale}\left(T_{ATC,d2}, MAST_{d2}, YAST_{d2}, dx_{d2}, MAST_{n2}, YAST_{n2}, dx_{n2}, T_{2m}, DEM\right) \quad \text{(R2)}$$

(3) Annual variation of final LST product in tropical region

Here, we further examine the annual variation of final hemispherical LST (which is predicted by the ML model) in mid-latitude regions (from 40°N to 50°N) and tropical regions (10°S–10°N) within the FY-4A disk at local noon, as illustrated in Fig. R1. A single-peak pattern is observed in mid-latitude regions, and a pronounced double-peak pattern is observed in tropical regions. The temperature peaks occurring slightly after the equinoxes. As a result, we conclude that the employed machine learning method could effectively refine the ATC model despite its imperfect performance in tropical region. We added a discussion about the influence of imperfect ATC modeling in limitation part in revised manuscript. Thank you again for your detailed comments.

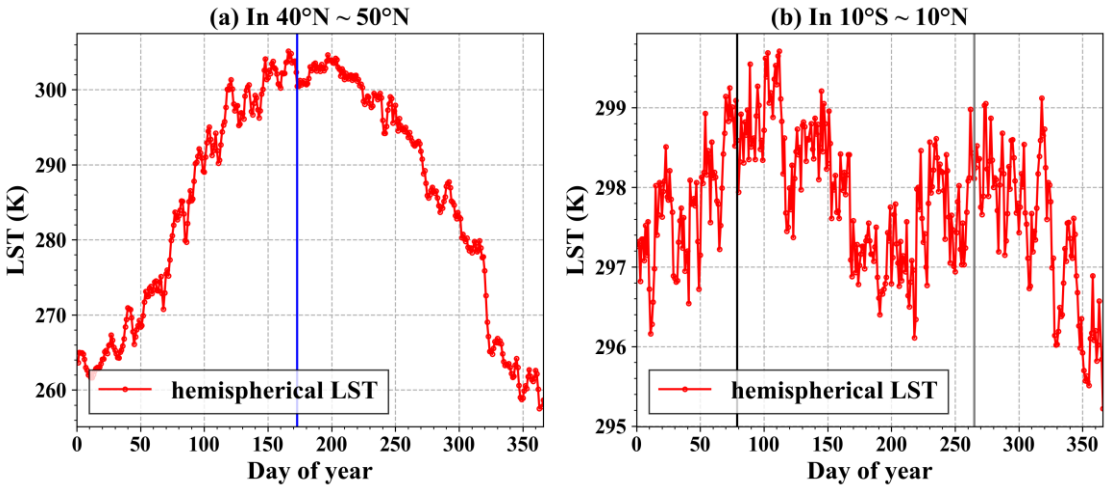

Fig. R1 Annual variation of the final hemispherical LST at noon in (a) mid-latitude regions (from 40°N to 50°N) and (b) tropical regions (10°S–10°N). The blue line in (a) is the summer solstice, the black and gray lines in (b) is spring equinox and autumn equinox, respectively.

Reference:

Ding, L., Zhou, J., Li, Z.-L., Ma, J., Shi, C., Sun, S., Wang, Z., 2022. Reconstruction of Hourly All-Weather Land Surface Temperature by Integrating Reanalysis Data and Thermal Infrared Data From Geostationary Satellites (RTG). IEEE Trans. Geosci. Remote Sens. 60, 1–17. https://doi.org/10.1109/TGRS.2022.3227074

Liu, Z., Zhan, W., Lai, J., Hong, F., Quan, J., Bechtel, B., Huang, F., Zou, Z., 2019. Balancing prediction accuracy and generalization ability: A hybrid framework for modelling the annual dynamics of satellite-derived land surface temperatures. ISPRS J. Photogramm. Remote Sens. 151, 189–206. https://doi.org/10.1016/j.isprsjprs.2019.03.013

Quan, J., 2023. Generating 60–100 m, hourly, all-weather land surface temperatures based on the Landsat, ECOSTRESS, and reanalysis temperature combination (LERC). ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2023.10.004

Tang, W., Zhou, J., Ma, J., Wang, Z., Ding, L., Zhang, Xiaodong, Zhang, Xu, 2024. TRIMS LST: a daily 1 km all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). Earth Syst. Sci. Data 16, 387–419. https://doi.org/10.5194/essd-16-387-2024

Xia, H., Chen, Y., Gong, A., Li, K., Liang, L., Guo, Z., 2021. Modeling Daily Temperatures Via a Phenology-Based Annual Temperature Cycle Model. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 6219–6229. https://doi.org/10.1109/JSTARS.2021.3085342

Yang, Yujia, Zhao, W., Yang, Yanqing, Xu, M., Mukhtar, H., Tauqir, G., Tarolli, P., 2024. An Annual Temperature Cycle Feature Constrained Method for Generating MODIS Daytime All-Weather Land Surface Temperature. IEEE Trans. Geosci. Remote Sens. 62, 1–14. https://doi.org/10.1109/TGRS.2024.3377670

Zhu, X., Duan, S.-B., Li, Z.-L., Wu, P., Wu, H., Zhao, W., Qian, Y., 2022. Reconstruction of land surface temperature under cloudy conditions from Landsat 8 data using annual temperature cycle model. Remote Sens. Environ. 281, 113261. https://doi.org/10.1016/j.rse.2022.113261

Zou, Z., Zhan, W., Liu, Z., Bechtel, B., Gao, L., Hong, F., Huang, F., Lai, J., 2018. Enhanced Modeling of Annual Temperature Cycles with Temporally Discrete Remotely Sensed Thermal Observations. Remote Sens. 10, 650. https://doi.org/10.3390/rs10040650

In Subsection 5.5:

In this study, a sinusoidal ATC model was employed to provide auxiliary thermal texture information. However, this approach may introduce uncertainties in tropical regions, where LST typically exhibits bimodal seasonal cycles with two annual peaks. Previous studies have attempted to address this limitation by incorporating dual-sinusoidal functions (Xing et al., 2020). In future work, the adoption of more accurate and universally applicable ATC models would be beneficial for improving the robustness of cloudy-sky LST estimation and downscaling.

2. **Temporal coverage and model stability are not demonstrated**

The product only spans the period up to 2023, and the training samples appear to be drawn from the same time span. The manuscript does not demonstrate whether the model can maintain stable performance for future years or under changing climatic conditions. Without testing temporal extrapolation or independent-year validation, it is difficult to assess whether the framework can support sustained long-term product generation. The dataset currently stops in 2023 and appears more like a test dataset, which cannot meet the standard of ESSD.

Response: Thank you for your constructive comments regarding the temporal stability of the proposed method. We apologize for not clearly describing the temporal scope of our framework in the original manuscript. **In practice, the angular normalization, cloud-gap filling, and downscaling processes are implemented on a year-by-year basis, rather than relying on a single model trained across multiple years.** Angular normalization was achieved by calibrating kernel-driven model parameters within one single day, cloud-gap filling was done by building separate models every year, and downscaling was achieved by establishing machine learning models for each image. These models are temporal interpolation models within one specific year. For instance, the model developed for 2018 is applied only to data from 2018, without any temporal extrapolation to other years. The detailed temporal configuration and processing frequency of each component are summarized in Table R2.

Table R2. The detailed information about temporal frequency of three steps.

Process	Frequency of models	Model numbers
Angular normalization	Daily	6 years × 365 days
Hypothetical clear-sky LST estimation during cloud-filling	Yearly	6 years
Downscaling	Hourly	6 years × 365 days × 24 hours

As summarized in Table R2, the product can be generated independently within each year and can be updated when additional auxiliary datasets become available in subsequent years. The dataset produced in this study spans the period from 2018 to 2023, **covering the full operational phase of the FY-4A LST product** (note that the official FY-4A LST product has not been generated since March 4, 2024). Thank you again for your detailed comments. In the revised manuscript, we have further clarified the temporal frequency and implementation details of each processing step.

In subsection 2.1:

The LST dataset produced in this study spans the period from 2018 to 2023, covering the full operational phase of the FY-4A LST product (note that the official FY-4A LST product has not been generated since March 4, 2024, as it has been superseded by a subsequent operational satellite).

In subsection 3.2:

A total of six ML models were trained on a year-by-year basis, rather than using a single model across multiple years. Each model is applied exclusively to data from its corresponding year, without temporal extrapolation to other periods.

In subsection 3.3:

Then, the $f_{downscale}$ was established individually for each image, covering 157,680 images in total (i.e., 6 years × 365 days × 24 hours × 3 types of LST).

3. Limited methodological novelty

The three core components—angular normalization (Na et al., 2024), cloudy-sky reconstruction (Zhang et al., 2024), and spatial downscaling (Liang et al., 2025)—are from existing studies the authors recently published. The manuscript mainly integrates these methods into a processing chain rather than introducing a simplified or integrated conceptual framework or breakthrough. The claimed novelty is therefore overstated.

Response: Thank you for your insightful comments regarding the novelty of this study. We're sorry that the novelty of this study is not well stated in original manuscript. Here, we clarify the main novelty of our LST product from the following three aspects：

(1) Requirements for physically consistent LST products

The development of LST products should be evaluated against the requirements proposed by Global Climate Observing System (GCOS), which emphasize high accuracy (~ 1 K), fine spatial resolution (~1 km), and high temporal frequency (sub-daily), along with long-term consistency. Although recent products have made progress toward these targets, angular consistency remains largely unaddressed. It's essential to ensure the physical homogeneity of LST as a climate variable. In this sense, the issue is not merely a technical detail, but a structural limitation that affects the compliance of existing products with GCOS standards. **Therefore, the primary contribution of this study is to redefine the target LST variable itself (from a directional quantity to an explicitly angular-normalized, physically consistent LST field)** and to demonstrate that such a variable can be retrieved under all-weather conditions and at high spatial resolution over the full geostationary disk for the first time. The integration of the three components is thus not a simple concatenation of existing methods, but a conceptually unified framework designed to resolve the incompatibility between angular effects, cloud contamination, and spatial scaling.

(2) Ignored angular effect in current cloud-filling and downscaling products

To our knowledge, no existing LST product simultaneously provides angular-normalized, all-weather, and high-resolution (~1 km, hourly) LST at the full-disk scale. Although a variety of clear-sky LST, all-weather LST, and downscaled LST products have been released (Dong et al., 2023, RSE; Jia et al., 2023, ESSD; Li et al., 2024, ESSD; Liu et al., 2023, IEEE TGRS; Tang et al., 2024, ESSD; Zhang et al., 2024, RSE), they generally share a common limitation: **the angular effect is either ignored or implicitly treated, rather**

than being explicitly resolved. Specifically, in many widely used approaches (such as surface energy balance-based cloudy-sky reconstruction or regression kernel-based downscaling methods) **directional LST observations are often treated as proxies for hemispherical LST without rigorous angular correction. This simplification may be acceptable under limited conditions, but it introduces systematic inconsistencies across viewing geometries**, especially for geostationary satellites such as FY-4A. For example, pixels located near the edge of the full disk experience significantly different viewing zenith angles compared to nadir pixels, leading to non-negligible angular-induced biases in both spatial and temporal analyses. Such inconsistencies directly affect the physical interpretability and comparability of LST across regions and time.

(3) Achieving daily-scale angular normalization for the first time

We have developed a variety of semi-physical kernel-driven models (KDMs) for characterizing the angular dependence of land surface temperature (LST) (Na et al., 2024 RS, Qin et al., 2023 RSE, 2025 RSE, Cao et al., 2021 RSE). However, conventional KDMs typically require 3–4 simultaneous multi-angle observations, which limits their practical applicability. The recently proposed time-evolving KDM (TEKDM) further releasing multi-angle data requirements, only needing nonsimultaneous seven observations within a single day for parameter calibration, resulting in a large fraction of clear-sky pixels (with <7 observations) remaining unprocessed. **To enable operational generation of angular-normalized LST products, this study proposed a novel 17-day KDM parameter aggregation strategy, which, to our knowledge, is the first attempt to achieve daily-scale angular normalization for all available pixels.** Results demonstrate that this framework substantially improves data coverage, allowing 93.6% of clear-sky pixels to be successfully processed.

In summary, we would like to clarify that the novelty of this study lies not in incremental improvements to each individual module, but in addressing a long-standing and fundamental gap in current LST products, namely the lack of angular consistency in all-weather, high-resolution LST datasets. We have revised the manuscript to better emphasize this point by (1) clearly distinguishing between module-level reuse and system-level innovation, and (2) explicitly positioning our work as a step toward next-generation LST products that meet GCOS requirements while ensuring angular consistency, rather than as a collection of incremental methodological improvements.

Reference:

Cao, B., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Liu, Q., Du, Y., Lagouarde, J.-P., Huang, H., Li, H., Bian, Z., Hu, T., Qin, B., Ran, X., Xiao, Q., 2021. A general framework of kernel-driven modeling in the thermal infrared domain. Remote Sens. Environ. 252, 112157. https://doi.org/10.1016/j.rse.2020.112157

Dong, P., Zhan, W., Wang, C., Jiang, S., Du, H., Liu, Z., Chen, Y., Li, L., Wang, S., Ji, Y., 2023. Simple yet efficient downscaling of land surface temperatures by suitably integrating kernel- and fusion-based methods. ISPRS J. Photogramm. Remote Sens. 205, 317–333. https://doi.org/10.1016/j.isprsjprs.2023.10.011

Jia, A., Liang, S., Wang, D., Ma, L., Wang, Z., Xu, S., 2023. Global hourly, 5 km, all-sky land surface temperature data from 2011 to 2021 based on integrating geostationary and polar-orbiting satellite data. Earth Syst. Sci. Data 15, 869–895. https://doi.org/10.5194/essd-15-869-2023

Li, B., Liang, S., Ma, H., Dong, G., Liu, X., He, T., Zhang, Y., 2024. Generation of global 1 km all-weather

instantaneous and daily mean land surface temperatures from MODIS data. Earth Syst. Sci. Data 16, 3795–3819. https://doi.org/10.5194/essd-16-3795-2024

Liu, W., Cheng, J., Wang, Q., 2023. Estimating Hourly All-Weather Land Surface Temperature From FY-4A/AGRI Imagery Using the Surface Energy Balance Theory. IEEE Trans. Geosci. Remote Sens. 61, 1–18. https://doi.org/10.1109/TGRS.2023.3254211

Na, Q., Cao, B., Qin, B., Mo, F., Zheng, L., Du, Y., Li, H., Bian, Z., Xiao, Q., Liu, Q., 2024. Correcting an Off-Nadir to a Nadir Land Surface Temperature Using a Multitemporal Thermal Infrared Kernel-Driven Model during Daytime. Remote Sensing 16, 1790. https://doi.org/10.3390/rs16101790

Qin, B., Cao, B., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Ermida, S.L., Bian, Z., Du, Y., Hu, T., Li, H., Xiao, Q., Chen, S., Liu, Q., 2023. A thermal radiation directionality correction method for the surface upward longwave radiation of geostationary satellite based on a time-evolving kernel-driven model. Remote Sens. Environ. 294, 113599. https://doi.org/10.1016/j.rse.2023.113599

Qin, B., Chen, S., Cao, B., Yu, Y., Yu, P., Na, Q., Hou, E., Li, D., Jia, K., Yang, Y., Hu, T., Bian, Z., Li, H., Xiao, Q., Liu, Q., 2025. Angular normalization of GOES-16 and GOES-17 land surface temperature over overlapping region using an extended time-evolving kernel-driven model. Remote Sens. Environ. 318, 114532. https://doi.org/10.1016/j.rse.2024.114532

Tang, W., Zhou, J., Ma, J., Wang, Z., Ding, L., Zhang, Xiaodong, Zhang, Xu, 2024. TRIMS LST: a daily 1 km all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). Earth Syst. Sci. Data 16, 387–419. https://doi.org/10.5194/essd-16-387-2024

Zhang, H., Tang, B.-H., Li, Z.-L., 2024. A practical two-step framework for all-sky land surface temperature estimation. Remote Sens. Environ. 303, 113991. https://doi.org/10.1016/j.rse.2024.113991

In Introduction:

It should be noted that this study does not aim to introduce incremental improvements to individual components, as the angular normalization, cloudy-sky reconstruction, and spatial downscaling modules have been previously developed. Instead, the novelty lies in establishing a unified processing framework that explicitly resolves angular inconsistency and enables their synergistic integration toward a conceptually consistent LST. By doing so, this work moves beyond conventional LST product generation and represents a step toward next-generation LST datasets that simultaneously satisfy the requirements of Global Climate Observing System (GCOS, https://gcos.wmo.int/site/global-climate-observing-system-gcos/essential-climate-variables/land-surface-temperature), including ~1 km spatial resolution, sub-daily temporal frequency, and climate-level consistency.

4. **Fragmented pipeline and potential error accumulation**

The workflow relies on multiple sequential ML-based steps (bias correction, angular normalization, gap-filling, and downscaling). Such a chained structure inevitably accumulates errors, yet the manuscript provides no uncertainty propagation analysis or ablation study to quantify how each module contributes to the final product accuracy.

Response: Thank you for your insightful comments. We fully agree that rigorous quantification of product uncertainty is essential for end users, which was insufficiently addressed in the original manuscript. Here, the uncertainty related to ML methods was analyzed comprehensively.

In the bias correction and angular normalization steps, we adopt widely used linear correction approaches and physical angular normalization methods (Ermida et al., 2018, 2017; Qin et al., 2025, 2023; Wei et al.,

2025), which have been demonstrated to provide stable and consistent accuracy improvements in previous studies. These two steps do not involve machine learning techniques, therefore, the derived FY-4A nadir LST (T_{nadir}) is not affected by additional data-driven uncertainty. Here, our uncertainty analysis primarily focuses on the ML-based cloud-filling and ML-based downscaling components.

The employed two-step cloud-filling method including hypothetical clear-sky LST estimation and cloud radiation force correction, and the ML model was only used in the first step. We quantify ML-induced uncertainty using the root mean square deviation (RMSD) at the pixel level over full FY-4A disk, calculated between the ML–estimated hypothetical nadir LST (T^{ML}_{nadir}) and clear-sky T_{nadir}. As shown in Fig. R2(a), the mean RMSD of the cloud-filling results is 2.49 K. Higher uncertainties are observed over regions such as the Qinghai–Tibet Plateau and the Indonesian Peninsula, which is comparable with recent all-weather LST estimation studies (Zhang et al., 2024, Li et al., 2025).

The spatial downscaling includes the fitting of 0.05° all-weather LST ($T^{0.05}_{nadir,all\text{-}weather}$) using ML model, predicting at 0.01° resolution, and residual redistribution. The uncertainty associated with the ML-estimated LST at 0.05° resolution ($T^{ML,\ 0.05}_{nadir,all\text{-}weather}$) is illustrated in Fig. R2(b), where the RMSD is computed between $T^{ML,\ 0.05}_{nadir,all\text{-}weather}$ and input $T^{0.05}_{nadir,all\text{-}weather}$. The average RMSD is 2.28 K, with higher values predominantly observed in mountainous regions, which exhibits similar accuracy with recent studies (Dong et al., 2023, 2020).

For cloud-contaminated pixels, the final LST incorporates uncertainties from both the cloud-filling and downscaling processes. The overall uncertainty for cloud-covered pixels is therefore estimated based on error propagation rules:

$$\sigma_{all} = \sqrt{\sigma^2_{CloudFilling} + \sigma^2_{Downscaling}} \tag{R3}$$

where σ_{all} is the overall uncertainty caused by ML models, $\sigma_{CloudFilling}$ is the uncertainty caused by cloud-filling, $\sigma_{Downscaling}$ is the uncertainty caused by downscaling. The resulting pixel-level uncertainty for cloud-contaminated LST is presented in Fig. R2(c). The propagated uncertainty from the machine learning components reaches an overall value of 3.4 K, with relatively higher values observed in mountainous regions.

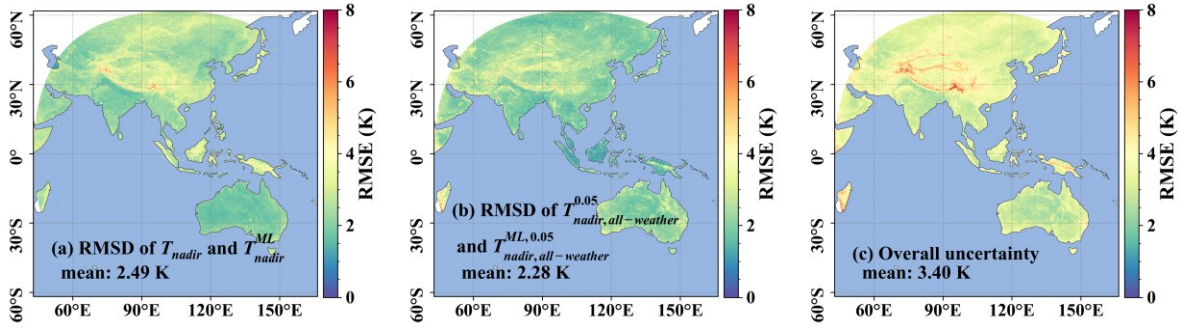

Fig. R2 The fitting accuracy of machine learning models in (a) cloud filling, (b) downscaling process, and (c) the overall accuracy.

The T-based validation indicates that the RMSE of the final LST product under cloudy-sky conditions is 3.65 K, which is comparable to the theoretical estimate of 3.4 K. Incorporating explicit error quantification enhances our understanding of uncertainty accumulation within the machine learning framework. We have

revised the manuscript accordingly and added the corresponding RMSD maps to support this analysis. Thank you again for your constructive suggestions, which have significantly improved the quality and rigor of our manuscript.

Reference:

Ermida, S.L., DaCamara, C.C., Trigo, I.F., Pires, A.C., Ghent, D., Remedios, J., 2017. Modelling directional effects on remotely sensed land surface temperature. Remote Sens. Environ. 190, 56–69. https://doi.org/10.1016/j.rse.2016.12.008

Ermida, S.L., Trigo, I.F., DaCamara, C.C., Pires, A.C., 2018. A Methodology to Simulate LST Directional Effects Based on Parametric Models and Landscape Properties. Remote Sens. 10, 1114. https://doi.org/10.3390/rs10071114

Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S.K., Zhao, L., 2020. Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. ISPRS Journal of Photogrammetry and Remote Sensing 169, 44–56. https://doi.org/10.1016/j.isprsjprs.2020.08.018

Dong, P., Zhan, W., Wang, C., Jiang, S., Du, H., Liu, Z., Chen, Y., Li, L., Wang, S., Ji, Y., 2023. Simple yet efficient downscaling of land surface temperatures by suitably integrating kernel- and fusion-based methods. ISPRS Journal of Photogrammetry and Remote Sensing 205, 317–333. https://doi.org/10.1016/j.isprsjprs.2023.10.011

Li, C., Wu, P., Duan, S.-B., Jia, Y., Sun, S., Shi, C., Yin, Z., Li, H., Shen, H., 2025. LFSR: Low-resolution Filling then Super-resolution Reconstruction framework for gapless all-weather MODIS-like land surface temperature generation. Remote Sensing of Environment 319, 114637. https://doi.org/10.1016/j.rse.2025.114637

Qin, B., Cao, B., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Ermida, S.L., Bian, Z., Du, Y., Hu, T., Li, H., Xiao, Q., Chen, S., Liu, Q., 2023. A thermal radiation directionality correction method for the surface upward longwave radiation of geostationary satellite based on a time-evolving kernel-driven model. Remote Sens. Environ. 294, 113599. https://doi.org/10.1016/j.rse.2023.113599

Qin, B., Chen, S., Cao, B., Yu, Y., Yu, P., Na, Q., Hou, E., Li, D., Jia, K., Yang, Y., Hu, T., Bian, Z., Li, H., Xiao, Q., Liu, Q., 2025. Angular normalization of GOES-16 and GOES-17 land surface temperature over overlapping region using an extended time-evolving kernel-driven model. Remote Sens. Environ. 318, 114532. https://doi.org/10.1016/j.rse.2024.114532

Wei, R., Duan, S.-B., Liu, X., Liu, N., Min, X., Li, Z.-L., 2025. Angular effect correction of remotely sensed land surface temperature by integrating geostationary and polar-orbiting satellite data. Remote Sens. Environ. 325, 114788. https://doi.org/10.1016/j.rse.2025.114788

Zhang, H., Tang, B.-H., Li, Z.-L., 2024. A practical two-step framework for all-sky land surface temperature estimation. Remote Sensing of Environment 303, 113991. https://doi.org/10.1016/j.rse.2024.113991

In Subsection 4.2:

Spatial heterogeneity in LST can lead to variations in model performance. The pixel-level RMSE of the CatBoost model for 2020 was evaluated by comparing the CatBoost-predicted hypothetical nadir LST with the FY-4A nadir LST. The nadir configuration is well suited for quantifying the performance of ML model, as it minimizes systematic uncertainties associated with angular effects. As shown in Fig. 8, the mean RMSE of the cloud-filling results is 2.49 K. Higher uncertainties are observed over regions such as the Qinghai–

Tibet Plateau and the Indonesian Peninsula, which can likely be attributed to persistent cloud cover. In such regions, the reduced availability of clear-sky observations limits the training samples, leading to increased uncertainty in the machine learning models.

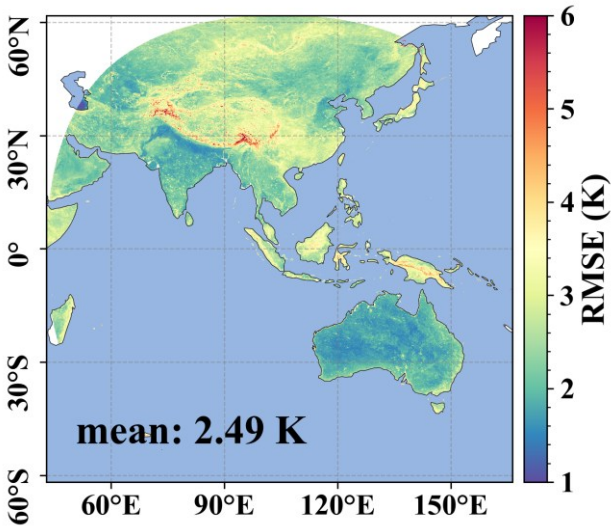

Figure 8: Pixel-level prediction RMSE of the CatBoost model in 2020.

In Subsection 4.3:

The spatial distribution of prediction accuracy for the machine learning model during the downscaling process is presented in Fig. 12, where the RMSE is calculated between the CatBoost-estimated LST and the input all-weather nadir LST. The overall RMSE is 2.28 K, with higher values predominantly observed in mountainous regions, likely due to enhanced surface heterogeneity and elevation-induced variability. It should be noted that clear-sky pixels remain unchanged in the all-weather LST product; therefore, the ANCFDS-LST for these pixels is affected only by uncertainties introduced during the downscaling step. In contrast, for cloud-contaminated pixels, the final LST incorporates uncertainties arising from both the cloud-filling and downscaling processes.

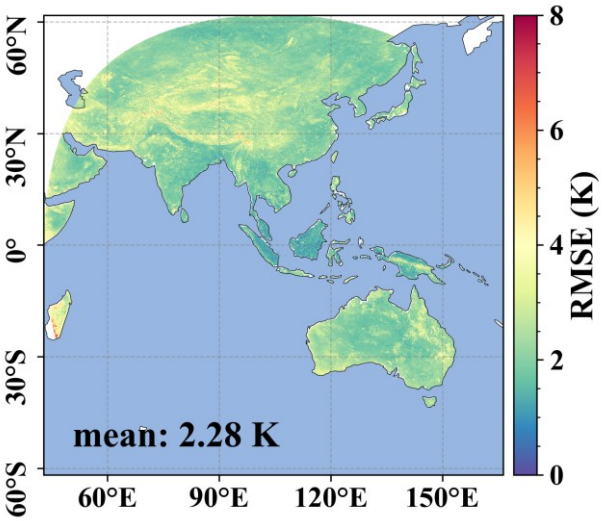

Figure 12: Spatial distribution of prediction RMSE for nadir LST over the test set in 2020.

5. **Problematic nighttime assumption**

The study assumes that nighttime LST has negligible angular effects. This assumption may hold for deep-night polar-orbiting observations, but hourly geostationary observations shortly after sunset can still exhibit directional thermal anisotropy due to soil–vegetation temperature contrasts. The validity of this assumption is not demonstrated.

Response: Thank you for your insightful comments regarding the directionality of nighttime LST. We agree that angular effects can still persist in geostationary observations shortly after sunset. Many existing studies assume that nighttime LST is primarily governed by the gap-fraction effect, and adopt the same gap-fraction kernel coefficient as in the daytime, following Vinnikov et al. (2012), Ermida et al. (2018, 2017), and Gao et al. (2023). However, this assumption may introduce additional uncertainties, as highlighted by more recent work (Snyders et al., 2025). They suggest that angular normalization can be improved by modeling daytime and nighttime TRD effects separately, which is an approach not yet implemented in the current TEKDM framework. Actually, this is a promising direction that warrants further in-depth investigation within the TEKDM model (Qin et al., 2025, 2023). In the revised manuscript, we explicitly acknowledge this limitation and discuss the implications of uncorrected nighttime directional effects. Thank you again for your professional and constructive comments, which have helped us improve the clarity and rigor of our study.

Reference:

Ermida, S.L., DaCamara, C.C., Trigo, I.F., Pires, A.C., Ghent, D., Remedios, J., 2017. Modelling directional effects on remotely sensed land surface temperature. Remote Sens. Environ. 190, 56–69. https://doi.org/10.1016/j.rse.2016.12.008

Ermida, S.L., Trigo, I.F., DaCamara, C.C., Pires, A.C., 2018. A Methodology to Simulate LST Directional Effects Based on Parametric Models and Landscape Properties. Remote Sens. 10, 1114. https://doi.org/10.3390/rs10071114

Gao, Y., Zhu, S., Zhang, G., Xu, Y., 2023. Accuracy Evaluation of the FY-4A AGRI Land Surface Temperature Product. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 9967–9976. https://doi.org/10.1109/JSTARS.2023.3326956

Jiang, L., Zhan, W., Tu, L., Dong, P., Wang, S., Li, L., Wang, Chunli, Wang, Chenguang, 2022. Diurnal variations in directional brightness temperature over urban areas through a multi-angle UAV experiment. Build. Environ. 222, 109408. https://doi.org/10.1016/j.buildenv.2022.109408

Qin, B., Cao, B., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Ermida, S.L., Bian, Z., Du, Y., Hu, T., Li, H., Xiao, Q., Chen, S., Liu, Q., 2023. A thermal radiation directionality correction method for the surface upward longwave radiation of geostationary satellite based on a time-evolving kernel-driven model. Remote Sens. Environ. 294, 113599. https://doi.org/10.1016/j.rse.2023.113599

Qin, B., Chen, S., Cao, B., Yu, Y., Yu, P., Na, Q., Hou, E., Li, D., Jia, K., Yang, Y., Hu, T., Bian, Z., Li, H., Xiao, Q., Liu, Q., 2025. Angular normalization of GOES-16 and GOES-17 land surface temperature over overlapping region using an extended time-evolving kernel-driven model. Remote Sens. Environ. 318, 114532. https://doi.org/10.1016/j.rse.2024.114532

Snyders, L., Blommaert, J., León-Tavares, J., Degerickx, J., 2025. Identification, Characterization and Correction of Directional Anisotropy in High-Resolution Land Surface Temperature through a Comparison of Sharpened Sentinel-3 with ECOSTRESS Data. Recent Adv. Remote Sens. 1–20. https://doi.org/10.62880/rars25008

Vinnikov, K.Y., Yu, Y., Goldberg, M.D., Tarpley, D., Romanov, P., Laszlo, I., Chen, M., 2012. Angular

anisotropy of satellite observations of land surface temperature. Geophys. Res. Lett. 39, L23802. https://doi.org/10.1029/2012GL054059

In Subsection 5.5:

The nighttime TRD effect was neglected in this study, as LST is generally more homogeneous within a pixel and angular effects are typically weak under nighttime conditions (Na et al., 2024a). However, TRD effect may persist in geostationary observations shortly after sunset (Jiang et al., 2022). Modeling daytime and nighttime TRD effects separately is a promising approach for improving nighttime angular normalization; however, this strategy has not yet been implemented in the current TEKDM framework. Therefore, the development of an advanced TEKDM that explicitly accounts for nighttime TRD effects warrants further investigation in future studies.

6. Weak physical justification for cloudy-sky directional consistency

The framework extends clear-sky directional normalization to cloudy-sky conditions without demonstrating whether directional effects persist, weaken, or change under cloud cover. The manuscript effectively transfers the clear-sky logic to cloudy conditions without physical justification.

Response: Thank you for your comment regarding the angular effect under cloudy conditions. We agree that the behavior of thermal directional anisotropy may change in the presence of clouds and that this aspect requires clear physical justification.

(1) The physical interpretation of angular effect under cloud-sky condition

Thermal infrared (TIR) directional anisotropy is primarily governed by surface thermal heterogeneity and canopy structure, which are strongly driven by solar illumination and the resulting temperature gradients between soil and vegetation components. Previous studies have demonstrated that the magnitude of anisotropy is closely linked to the intensity of solar radiation and the associated thermal contrast within the surface–atmosphere system (Trigo et al., 2008; Zhan et al., 2025). Under cloudy conditions, two key physical changes occur: (1) the reduction of direct solar radiation, and (2) the increase of diffuse radiation and atmospheric longwave emission. These processes tend to homogenize surface temperature distributions and reduce vertical and horizontal thermal gradients within the canopy–soil system. As a result, the magnitude of TIR directional anisotropy is expected to decrease rather than increase under cloudy skies (Zhan et al., 2025).

(2) Representation of cloud-sky TRD effects and associated modeling challenges

Based on the above understanding, we can conceptually establish a transformation for TRD intensity (TRDI) between clear-sky and cloudy-sky conditions (i.e., $TRDI_{cld}=\alpha\times TRDI_{clr}$), where α is scaling factor related to cloud optical property, duration time, and humidity. It is difficult to physically estimate the exact value of α; therefore, we implicitly estimate $TRDI_{cld}$ using machine learning models and reanalysis data. Observing and modeling TRD effects under cloudy-sky conditions is a promising direction that warrants further study. We apologize that the original manuscript did not sufficiently state the cloud-sky TRD effect. In the revised version, we modified manuscript accordingly to clarify (1) the expected TRD effect under cloudy conditions, (2) the implications for the proposed framework, and (3) the limitations and future works for comprehensive understanding cloudy-sky TRD effect.

Reference:

Trigo, I.F., Monteiro, I.T., Olesen, F., Kabsch, E., 2008. An assessment of remotely sensed land surface temperature. J. Geophys. Res. Atmospheres 113, 2008JD010035. https://doi.org/10.1029/2008JD010035

Zhan, W., Du, H., Liu, Z., Li, J., Chakraborty, T.C., Huang, F., 2025. Harnessing Satellite Data Alone for Mapping Global Thermal Anisotropy. Geophys. Res. Lett. 52, e2024GL113479. https://doi.org/10.1029/2024GL113479

In Subsection 5.5:

Thermal infrared directional anisotropy is primarily governed by surface thermal heterogeneity and canopy structure, both of which are strongly driven by solar illumination and the resulting temperature gradients between soil and vegetation components. Under cloudy conditions, direct solar radiation is reduced, while diffuse radiation and atmospheric longwave emission increase, which tends to weaken the magnitude of TIR directional anisotropy. However, the weakened TRD effect is strongly influenced by cloud optical properties, cloud duration, and atmospheric humidity, making it difficult to model using physical approaches. Future studies are encouraged to incorporate multi-angle observations under cloudy-sky conditions to better characterize TRD effects under cloud cover (Jia et al., 2024).

7. Heavy dependence on reanalysis and auxiliary variables

The models rely extensively on reanalysis radiation, air temperature, and other auxiliary variables from reanalysis (even dewpoint T?). This raises the question of whether the final product genuinely reflects satellite observations or largely reproduces reanalysis-driven temperature fields. A semi-physical framework can, in principle, leverage multiple data sources and physical constraints, but this advantage is not fully utilized here. In practice, the model appears to be dominated by reanalysis inputs while still relying on an oversimplified SEB parameterization, making the choice of a semi-physical approach difficult to justify.

Response: Thank you for your comment regarding the role of reanalysis variables in the all-weather LST estimation. We agree that the balance between satellite observations and auxiliary data is critical for ensuring the physical interpretability and added value of the final LST product.

(1) In hypothetical clear-sky LST estimation, reanalysis data were employed to constrain the ill-posed nature of cloudy-sky reconstruction. Ablation experiment shows reanalysis data is not the dominate factor for final LST.

We agree the concern that the model may "reproduce reanalysis temperature fields", but this typically occurs only when satellite information is weak or absent. In our case, the primary observational constraint remains the official FY-4A LST products. The clear-sky LST values remain unchanged, and also act as the training labels for cloud-sky LST estimation. The auxiliary variables (e.g., air temperature, radiation, and dewpoint temperature) are introduced to constrain the ill-posed nature of cloudy-sky reconstruction and abundant the spatial texture that was not captured by satellite sensors, rather than to replace satellite information. This strategy is consistent with many previous studies (Jia et al., 2024; Wu et al., 2021). Furthermore, we acknowledge that the current manuscript does not explicitly quantify the balance between satellite and reanalysis inputs. Here, an ablation experiment was added, in which key reanalysis variables (i.e., air temperature, income radiation, and dewpoint temperature) are removed from the model. The results show that the model performance degrades from 2.34 K to 2.57 K, with an RMSE reduction of 0.23 K (9.8%).

This demonstrates that reanalysis data play a supporting but not dominating role.

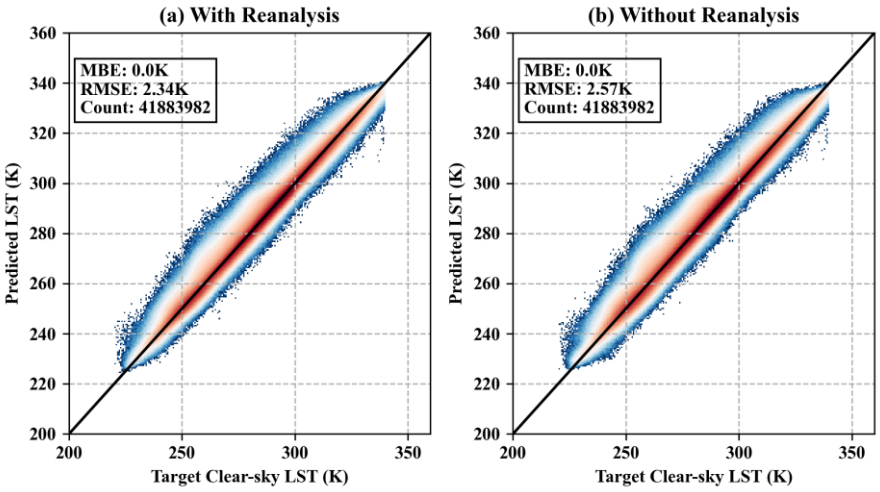

Fig. R3 The fitting accuracy of ML models (a) with and (b) without auxiliary reanalysis data as input.

(2) For semi-physical framework in cloud radiation force (CRF) correction, our intention is not to fully resolve the surface energy balance (SEB), but to introduce physically interpretable constraints into a data-driven model.

Current LST cloud-filling methods are generally classified into interpolation-based, SEB-based, PMW-based, and simulation-based methods. Among these, interpolation-based methods neglect the CRF effect and therefore cannot recover realistic cloud-sky LST. PMW-based methods cannot provide hourly observations, and are further limited by orbital gaps, coarse spatial resolution, and inconsistencies in physical properties between PMW and TIR signals. Simulation-based methods often suffer from systematic biases between model simulations and TIR observations (Jia et al., 2024). In contrast, hybrid methods that incorporate SEB constraints based on reanalysis data are widely adopted due to their ability to account for the CRF effect, support hourly applications, and exhibit reliable performance (W. Du et al., 2025; Firozjaei et al., 2024; Jia et al., 2021, 2023; Liu et al., 2023; Wu et al., 2025; Zhang et al., 2024). We acknowledge that the SEB representation in this framework is simplified, and this point has been clarified in the revised manuscript to avoid overstating its physical rigor.

Finally, we have revised the manuscript to (1) further clarify the role of auxiliary variable, (2) explicitly discuss the potential limitation of reanalysis data, and (3) better justify the use of a semi-physical framework. Considering reanalysis variables are important for stabilizing the reconstruction under cloudy conditions and enhancing spatial completeness, the final product remains unchanged for reanalysis factors. Thank you again for your reconstructive suggestions.

Reference:

Du, W., Li, Z.-L., Qin, Z., Fan, J., Liu, X., Zhao, C., Cao, K., 2025. Reconstruction of Cloudy Land Surface Temperature by Combining Surface Energy Balance Theory and Solar-Cloud-Satellite Geometry. IEEE Trans. Geosci. Remote Sens. 63, 1–13. https://doi.org/10.1109/TGRS.2025.3532446

Firozjaei, M.K., Mijani, N., Kiavarz, M., Duan, S.-B., Atkinson, P.M., Alavipanah, S.K., 2024. A novel surface energy balance-based approach to land surface temperature downscaling. Remote Sens. Environ. 305, 114087. https://doi.org/10.1016/j.rse.2024.114087

Jia, A., Liang, S., Wang, D., Ma, L., Wang, Z., Xu, S., 2023. Global hourly, 5 km, all-sky land surface temperature data from 2011 to 2021 based on integrating geostationary and polar-orbiting satellite data. Earth Syst. Sci. Data 15, 869–895. https://doi.org/10.5194/essd-15-869-2023

Jia, A., Liang, S., Wang, D., Mallick, K., Zhou, S., Hu, T., Xu, S., 2024. Advances in Methodology and Generation of All-Weather Land Surface Temperature Products From Polar-Orbiting and Geostationary Satellites: A comprehensive review. IEEE Geosci. Remote Sens. Mag. 2–43. https://doi.org/10.1109/MGRS.2024.3421268

Jia, A., Ma, H., Liang, S., Wang, D., 2021. Cloudy-sky land surface temperature from VIIRS and MODIS satellite data using a surface energy balance-based method. Remote Sens. Environ. 263, 112566. https://doi.org/10.1016/j.rse.2021.112566

Liu, W., Cheng, J., Wang, Q., 2023. Estimating Hourly All-Weather Land Surface Temperature From FY-4A/AGRI Imagery Using the Surface Energy Balance Theory. IEEE Trans. Geosci. Remote Sens. 61, 1–18. https://doi.org/10.1109/TGRS.2023.3254211

Wu, P., Yin, Z., Zeng, C., Duan, S.-B., Gottsche, F.-M., Ma, X., Li, X., Yang, H., Shen, H., 2021. Spatially Continuous and High-Resolution Land Surface Temperature Product Generation: A review of reconstruction and spatiotemporal fusion techniques. IEEE Geosci. Remote Sens. Mag. 9, 112–137. https://doi.org/10.1109/MGRS.2021.3050782

Zhang, H., Tang, B.-H., Li, Z.-L., 2024. A practical two-step framework for all-sky land surface temperature estimation. Remote Sens. Environ. 303, 113991. https://doi.org/10.1016/j.rse.2024.113991

In Subsection 3.2:

It should be noted that the reanalysis variables are introduced to abundant the spatial texture that was not captured by satellite sensors, rather than to replace satellite information. The primary observational constraint remains the official FY-4A LST products. The clear-sky LST values remain unchanged, and also act as the training labels for cloud-sky LST estimation.

…

Considering that our aim is not to fully resolve the surface energy balance (SEB), a simplified analytical correction equation (Eq. 14) proposed by Zhang et al. (2024) was adopted for the CRF correction of T_{hemi}.

In Subsection 5.5:

This study extensively employs machine learning methods driven by reanalysis products. Although the primary observational constraints are derived from remote sensing products, uncertainties in the inputs may still propagate through the cloudy-sky LST estimation and downscaling processes. In future studies, developing robust error quantification techniques would aid in tracing how uncertainties in input parameters influence the final 0.01° all-weather LST, thereby further refining the ANCFDS-LST product (Li et al., 2024b).

8. Questionable value of the downscaling step

The product is downscaled from 0.05° (~4 km) to 0.01° (~1 km), but the resulting improvement is minimal. The downscaling is not helpful but may introduce additional uncertainty. Air temperature is one of the dominant predictors in the downscaling model, yet it is obtained from simply interpolated reanalysis fields. If the downscaled LST strongly depends on interpolated air temperature rather than satellite observations, the added scientific value of the product becomes questionable.

Response: Thank you for your insightful comments regarding the use of air temperature in the downscaling

process. The primary objective of downscaling in this study is to enhance spatial detail without compromising accuracy, while meeting the Global Climate Observing System (GCOS) requirements for LST products (i.e., hourly temporal resolution and 1 km spatial resolution). Accordingly, we have incorporated additional metrics in Section 4.3 of the revised manuscript to quantify improvements in spatial detail.

Regarding the use of ERA5 surface air temperature (SAT), our analysis indicates that it introduces little systematic bias into the downscaling process. In addition, it also helps regularize spatial patterns in LST fields and improves the overall fitting performance.

(1) Improved spatial consistency and fitting accuracy by incorporating SAT

Cloud cover typically induces localized cooling effects that cannot be captured by clear-sky remote sensing regression kernels alone. As illustrated in Fig. R5 (a), the 0.05° all-weather LST exhibits a distinct strip-shaped cooling pattern associated with cloud presence. This feature is successfully reproduced when SAT is included as an input (Fig. R5 (b)), whereas the model without SAT yields an overly homogeneous spatial distribution (Fig. R5 (c)). Therefore, SAT plays a critical role in preserving physically meaningful spatial variability in all-weather LST fields.

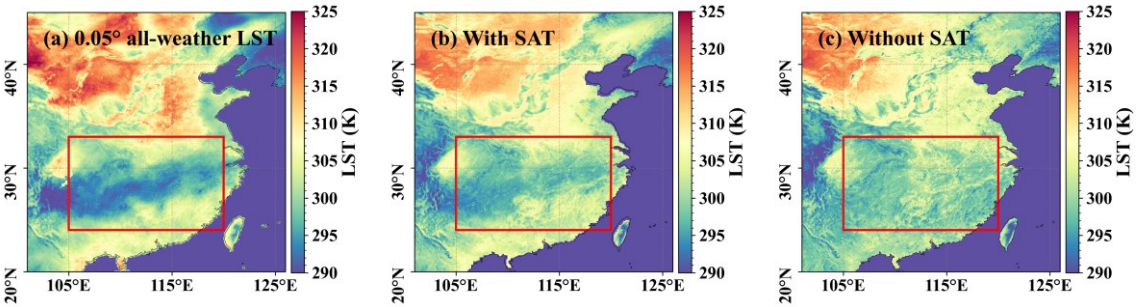

Fig. R5 Spatial distribution of (a) all-weather LST to be downscaled, (b) reconstructed LST via downscaling model with SAT as input, (c) the same as (b), but without SAT as input on 2020/06/24 at 12:00.

SAT can also capture the temporal evolution of near-surface thermal states under both clear-sky and cloud-sky conditions. We tested the fitting accuracy of the ML model with and without SAT as an input feature. As shown in Fig. R4, excluding SAT from the model increases the RMSD between ML–predicted LST and all-weather 0.05° LST from 2.28 K to 3.00 K (an increase of 0.72 K, 31.6%). This degradation would propagate into the final downscaled LST product, indicating that incorporating SAT substantially improves model performance.

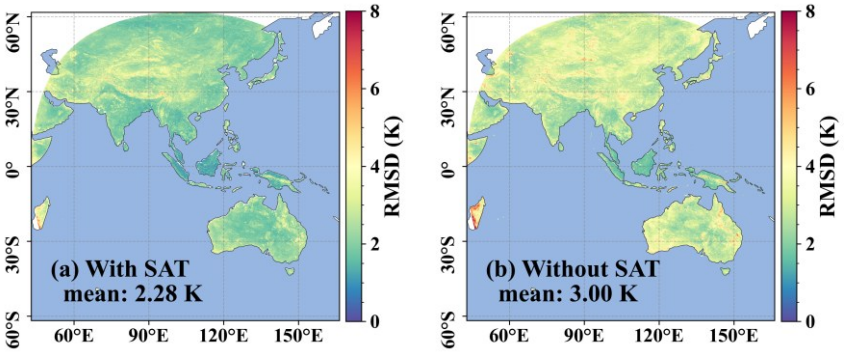

Fig. R4 Fitting accuracy of ML model (a) with and (b) without SAT as input.

(2) Without introducing significant systematic bias

Although ERA5 SAT differs systematically from LST, this discrepancy is mitigated within the statistical downscaling framework. First, the ML model is primarily constrained using all-weather LST as the target variable. Feature importance analysis indicates that SAT contributes 21.9% of the total importance, rather than dominating the model. Second, a residual redistribution step ensures consistency between the aggregated 0.01° downscaled LST and the original 0.05° LST, further reducing potential bias. As shown in Fig. R6, the RMSE between resampled 0.01° LST and original 0.05° all-weather LST is 0.79 K, with no evident systematic bias.

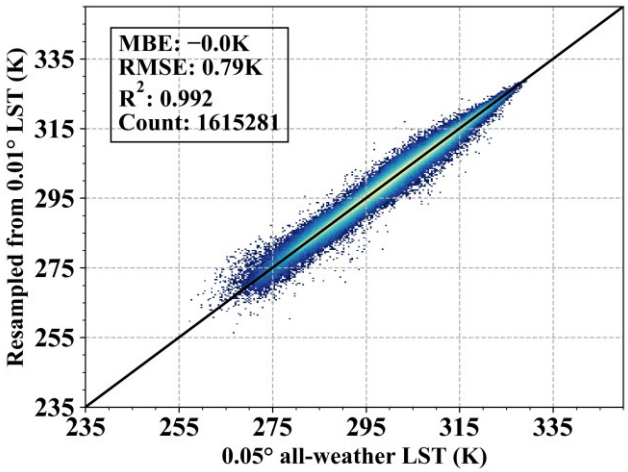

Fig. R6 Comparison between original 0.05° LST and resampled 0.01° LST on 2020/06/24 at 12:00.

Based on above considerations, we retain SAT as an input feature due to its few influence on systematic bias, contribution to more realistic spatial representation and improved fitting accuracy (Xia, 2025). Nevertheless, we acknowledge that the use of reanalysis data introduces additional uncertainty. This limitation has now been explicitly stated in the revised manuscript. Thank you again for your valuable comments, which have helped us substantially improve the robustness and clarity of our study.

Reference:

Xia, H., 2025. Geographically Constrained Machine Learning-Based Kernel-Driven Method for Downscaling of All-Weather Land Surface Temperature. Remote Sens. 17, 1413. https://doi.org/10.3390/rs17081413

In Subsection 4.3:

Downscaled LST is expected to exhibit enhanced spatial texture, which is quantitatively evaluated in this study. The local variance ratio (LVR) and the structural similarity index (SSIM) were employed as metrics to compare the spatial textures of bilinearly interpolated LST (derived from 0.05° all-weather LST) and the downscaled 0.01° nadir LST. The reference texture was reconstructed using the ATC model, as interpolation-based ATC approaches have been shown to better preserve spatial texture characteristics (Jia et al., 2024). Texture similarity improves as LVR and SSIM approach 1. As shown in Fig. 13, the median LVR increases from 0.41 to 0.97, while the median SSIM improves from 0.91 to 0.94. These quantitative results demonstrate that the downscaled LST provides a substantially improved representation of spatial texture.

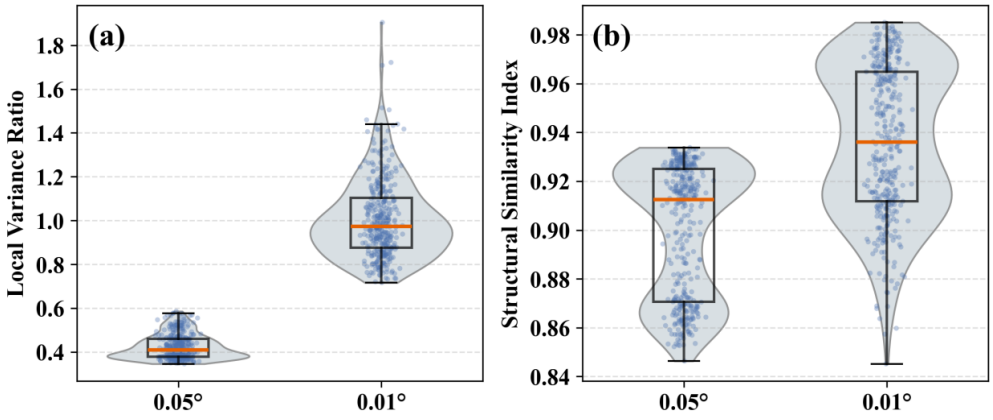

Figure 13: (a) Local variance ratio and (b) structural similarity index between bilinearly interpolated LST and the downscaled 0.01° nadir LST, with ATC-simulated LST used as the reference.

In Subsection 5.5:

This study extensively employs machine learning methods driven by reanalysis products. Although the primary observational constraints are derived from remote sensing products, uncertainties in the inputs may still propagate through the cloudy-sky LST estimation and downscaling processes. In future studies, developing robust error quantification techniques would aid in tracing how uncertainties in input parameters influence the final 0.01° all-weather LST, thereby further refining the ANCFDS-LST product (Li et al., 2024b).

9. **Limited and geographically biased validation**

The validation relies heavily on stations concentrated in the Heihe River Basin. Such spatially clustered sites cannot represent the accuracy of a product intended to cover the full FY-4 disk. In addition, the spatial representativeness of point measurements at the 4-km scale remains uncertain.

Response: Thank you for your constructive comments regarding the in-situ validation of LST. We agree that the spatial distribution and representativeness of in situ sites are of critical importance for LST evaluation. In response, we supplemented the dataset with sites from the Terrestrial Ecosystem Research Network (TERN) OzFlux network to enrich the available in situ observations (Li et al., 2020). Then, 11 HiWATER sites, 11 OzFlux sites, and 6 sites over the Tibetan Plateau (a total of 28 sites) were obtained. We conducted a comprehensive representativeness assessment following the approach of Zhang et al. (2024) to identify high-quality sites. Specifically, the nighttime RMSE and the standard deviation (STD) of Landsat-8 LST within a 5 km window were used as evaluation metrics. Sites that deviated from the median values of these indicators were excluded, as illustrated in Fig. R7. As a result, eight sites were retained, all of which meet the criteria of nighttime RMSE < 3.57 K and STD < 2.17 K.

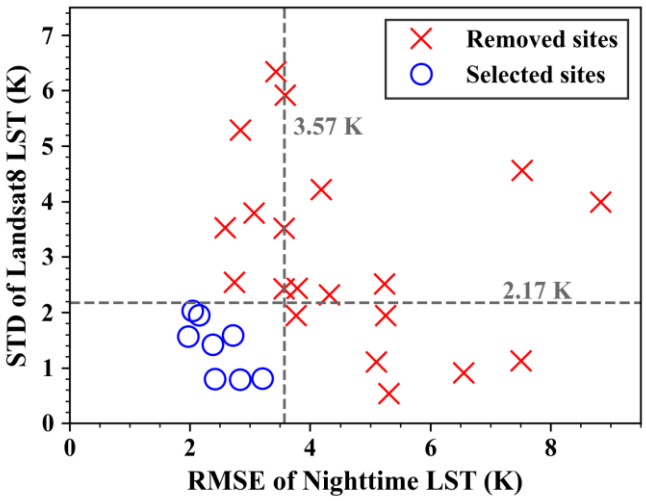

Fig. R7 The representativeness evaluation of employed in-situ sites.

The spatial distribution of high-quality sites is shown in Fig. R8. Five land cover types are represented across the 8 validation sites: 1 cropland (CRO) site, 3 grassland (GRA) sites, 1 barren/sparse vegetation (BSV) site, 2 savanna (SAV) sites, and 1 evergreen broadleaf forest (EBF) site.

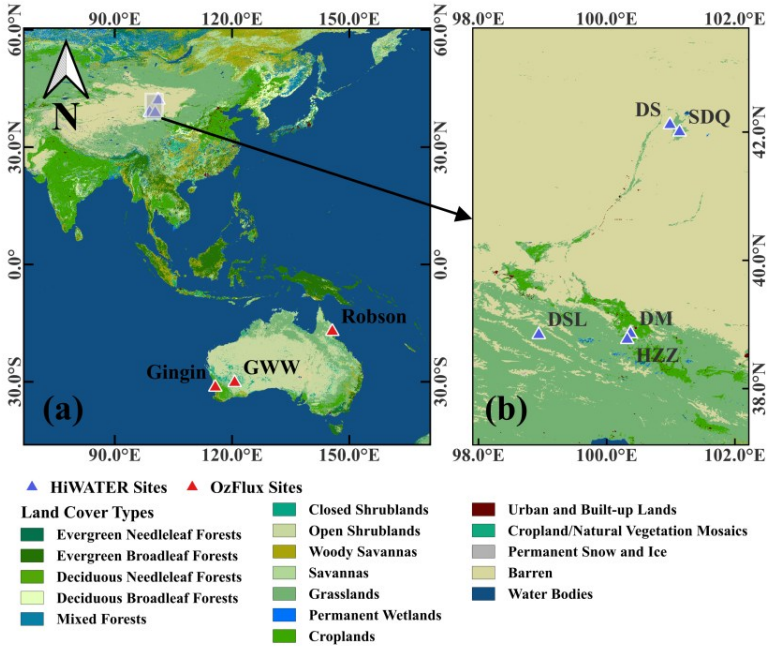

Fig. R8 The spatial distribution of employed in-situ sites.

After replacing the sites with high-quality ones, we re-evaluated the accuracy of the LST products, as shown in Fig. R9. The RMSE (MBE) of the hemispherical LST is 2.48 K (−0.83 K) under clear-sky conditions and 3.65 K (−1.43 K) under cloudy-sky conditions, which is better than the accuracy in original manuscript (which is 2.99 K (-0.77 K) under clear-sky conditions, and 4.56 K (-1.56 K) under cloudy-sky conditions). This level of accuracy is consistent with that reported for all-weather LST products in recent studies (see Table R3).

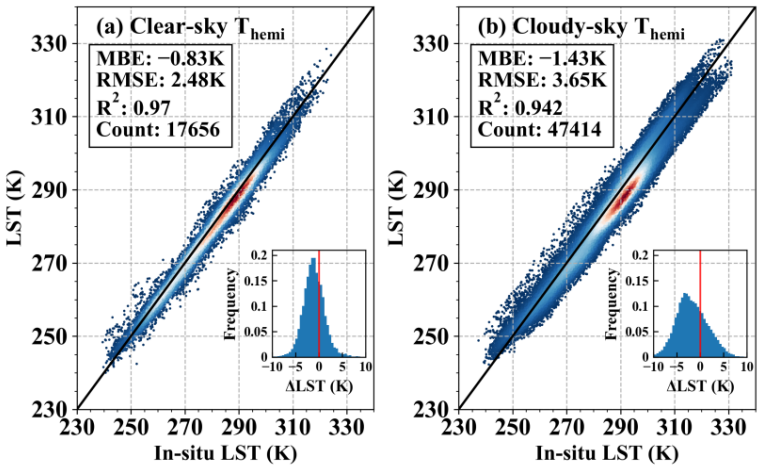

Fig. R9 The accuracy of hemispherical LST under (a) clear-sky and (b) cloudy-sky conditions.

Table R3 The reported RMSE of all-weather estimation methods in recent years

Research	Reported RMSE
Zhang et al. (2024, RSE)	2.6 K
Yu et al. (2025, IEEE TGRS)	3.1 K
Jia et al. (2023, ESSD)	3.3 K
Tang et al. (2024, ESSD)	3.7 K
Ma et al. (2024, RSE)	3.7 K
Li et al. (2025, RSE)	5.7 K

As shown in Table R3, the RMSE of current all-weather LST estimation methods ranges from 2.6 K to 5.7 K, which is comparable to the RMSE of our ANCFDS-LST. In the revised manuscript, the accuracy evaluation based on in-situ sites has been updated. Thank you again for your constructive suggestions. They indeed help to improve the quality of our manuscript.

Reference:

Jia, A., Liang, S., Wang, D., Ma, L., Wang, Z., Xu, S., 2023. Global hourly, 5 km, all-sky land surface temperature data from 2011 to 2021 based on integrating geostationary and polar-orbiting satellite data. Earth Syst. Sci. Data 15, 869–895. https://doi.org/10.5194/essd-15-869-2023

Li, C., Wu, P., Duan, S.-B., Jia, Y., Sun, S., Shi, C., Yin, Z., Li, H., Shen, H., 2025. LFSR: Low-resolution Filling then Super-resolution Reconstruction framework for gapless all-weather MODIS-like land surface temperature generation. Remote Sens. Environ. 319, 114637. https://doi.org/10.1016/j.rse.2025.114637

Li, R., Li, H., Sun, L., Yang, Y., Hu, T., Bian, Z., Cao, B., Du, Y., Liu, Q., 2020. An Operational Split-Window Algorithm for Retrieving Land Surface Temperature from Geostationary Satellite Data: A Case Study on Himawari-8 AHI Data. Remote Sens. 12, 2613. https://doi.org/10.3390/rs12162613

Ma, J., Shen, H., Jiang, M., Lin, L., Meng, C., Zeng, C., Li, H., Wu, P., 2024. A mechanism-guided machine learning method for mapping gapless land surface temperature. Remote Sens. Environ. 303, 114001. https://doi.org/10.1016/j.rse.2024.114001

Tang, W., Zhou, J., Ma, J., Wang, Z., Ding, L., Zhang, Xiaodong, Zhang, Xu, 2024. TRIMS LST: a daily 1 km

all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). Earth Syst. Sci. Data 16, 387–419. https://doi.org/10.5194/essd-16-387-2024

Yu, W., Deng, X., Xiao, Y., Huang, Y., Zhou, W., Liu, X., 2025. Estimating All-Weather Land Surface Temperature: A Method Considering Cloud Fraction and Energy Balance. IEEE Trans. Geosci. Remote Sens. 63, 1–15. https://doi.org/10.1109/TGRS.2025.3576661

Zhang, H., Tang, B.-H., Li, Z.-L., 2024. A practical two-step framework for all-sky land surface temperature estimation. Remote Sens. Environ. 303, 113991. https://doi.org/10.1016/j.rse.2024.113991

In Subsection 2.3:

As shown in Fig. 2, 5 sites from the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) experiment within the Heihe River Basin (HRB) and 3 sites from the Terrestrial Ecosystem Research Network (TERN) OzFlux network in 2020 were selected, which had been used in LST evaluation studies (Beringer et al., 2016; Che et al., 2019; Li et al., 2025, 2020). More detailed information on the spatial representativeness of the in situ sites is provided in Subsection 5.3.

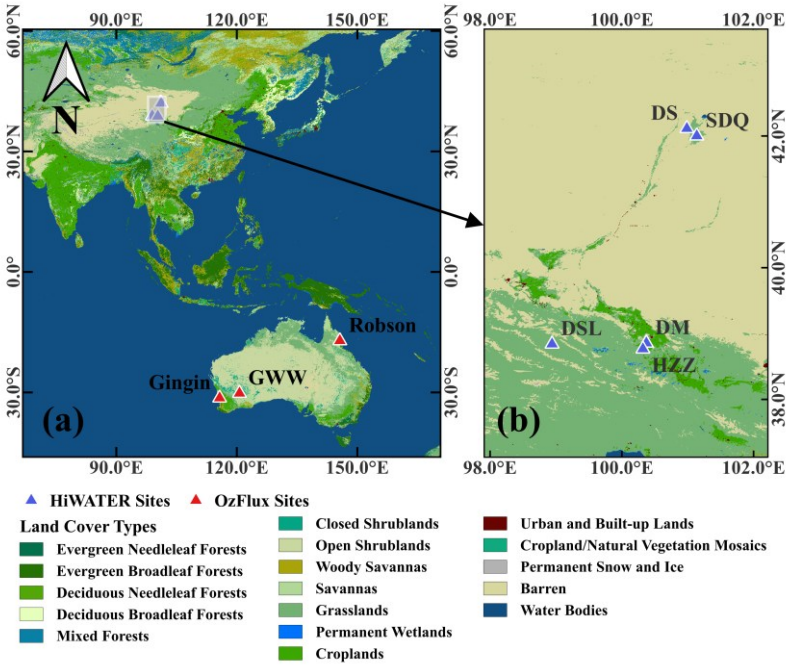

Figure 2: Spatial distribution of selected in-situ sites.

Table 2. Information of the in-situ sites

Network	Site name	Longitude (° E)	Latitude (° N)	Elevation (m)	Temporal resolution (minute)	Land cover type
HiWATER	DM	100.37	38.86	1560	10	CRO
	DSL	98.94	38.84	3787	10	GRA
	DS	100.99	42.11	927	10	BSV
	HZZ	100.32	38.77	1732	10	GRA
	SDQ	101.14	42.00	935	10	GRA
OzFlux	Gingin	115.71	-31.38	99	30	SAV
	GWW	120.65	-30.19	505	30	SAV
	Robson	145.63	-17.12	667	30	EBF

As listed in Table 2, five land cover types are represented across the 8 validation sites: 1 cropland (CRO) site, 3 grassland (GRA) sites, 1 barren/sparse vegetation (BSV) site, 2 savanna (SAV) sites, and 1 evergreen broadleaf forest (EBF) site. The temporal resolution of the in-situ observations varies by network: 10 minutes for the HiWATER sites, and 30 minutes for the OzFlux sites.

In Subsection 4.1:

The evaluation of 0.05° FY-4A T_{dir} and T_{hemi} during daytime over 8 stations is shown in Fig. 7a-b. Results show that the RMSE decreased from 2.22 K to 2.00 K, with an improvement of 0.22 K. The MBE was reduced from 0.19 K to -0.54 K. For the temperatures between 280 K to 300 K with most densely distribution, the scatter points become closer to the 1:1 line after the angular normalization. Fig. 7c-d shows the RMSE and MBE of FY-4A T_{dir} and T_{hemi} at different local times. T_{dir} exhibits a pronounced overestimation with higher RMSE in the morning, and an underestimation with smaller RMSE in the afternoon. After the angular normalization, the T_{hemi} shows significantly improved performance in the morning, with an RMSE reduced by approximately 0.4-0.8 K (Fig. 7c). As shown in Fig. 7d, the MBEs of T_{hemi} is much smaller than those of T_{dir} in the morning, with a maximum reduction of 1.7 K at 11:00.

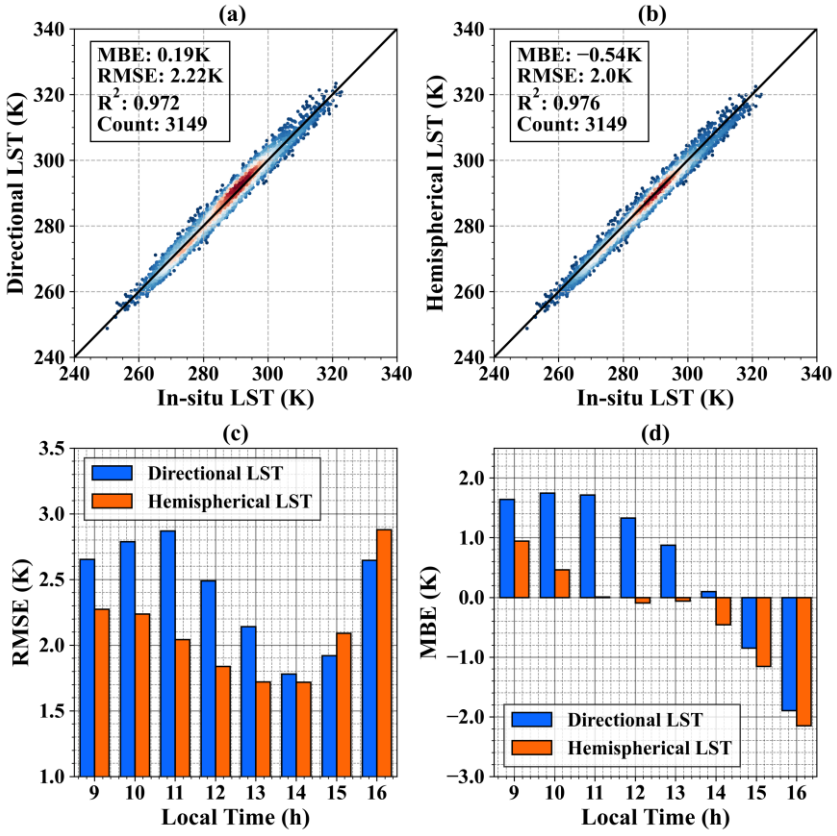

Figure 7: The accuracy of FY-4A (a) directional LST, (b) hemispherical LST, and the temporal variation of (c) RMSE and (d) MBE for directional and hemispherical LST.

In Subsection 4.2:

The estimated 0.05° all-weather T_{hemi} results were further evaluated by the in-situ hemispherical LSTs from 8 in-situ stations (Fig. 9). The samples are divided into clear-sky (Fig. 9a) and cloudy-sky (Fig. 9b) conditions. The "3σ-Hampel identifier" method was applied to minimize the influence of outliers at each site. Both daytime and nighttime T_{hemi} values were considered here. Under clear-sky conditions, 0.05° T_{hemi} shows an

RMSE of 2.48 K and an MBE of -0.83 K. The distribution of residuals ($\Delta LST = T_{hemi} - T_{insitu}$) approximately follows a normal distribution. Under cloudy-sky conditions, T_{hemi} shows a higher RMSE of 3.65 K and a more significant negative MBE of -1.43 K. It is challenging to maintain the same level of accuracy under cloudy-sky conditions, as uncertainties in both the CatBoost-predicted hypothetical LST and the estimated ΔT_{CRF} are introduced.

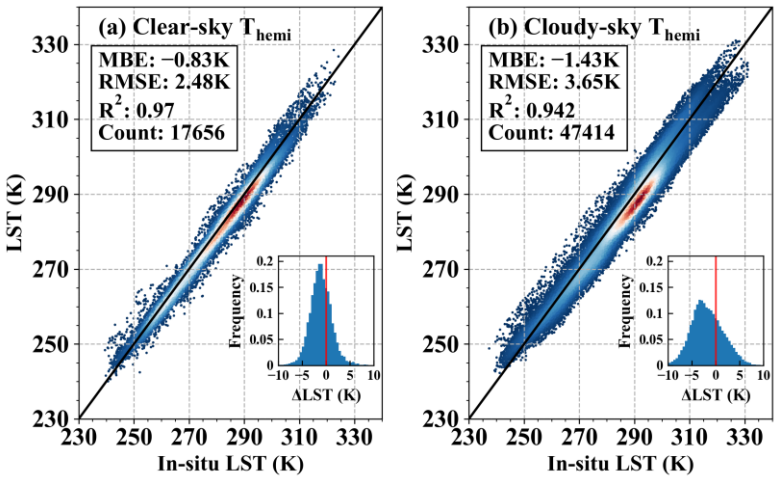

Figure 9: The accuracy of generated all-weather hemispherical LST under (a) clear-sky and (b) cloudy-sky condition.

In Subsection 5.3:

T_{nadir} was evaluated against near-nadir LST from VNP21A1, whereas the 0.05° T_{hemi} product was validated using in-situ measurements. The T-based validation of T_{hemi} requires that in-situ sites exhibit high spatial representativeness. To this end, the spatial heterogeneity of 11 HiWATER sites, 11 OzFlux sites, and 6 sites over the Tibetan Plateau in 2020 was assessed (Beringer et al., 2016; Che et al., 2019; Y. Ma et al., 2024). Following Zhang et al. (2024), two metrics were adopted: (1) the median standard deviation (STD) of Landsat 8 LST within a 0.05° window, and (2) the RMSE between nighttime FY-4A LST and in situ LST. Sites that deviated from the median values of these two indicators were excluded, as illustrated in Fig. 17 (a). Ultimately, 8 sites were retained, all satisfying the criteria of nighttime RMSE< 3.57 K and STD< 2.17 K.

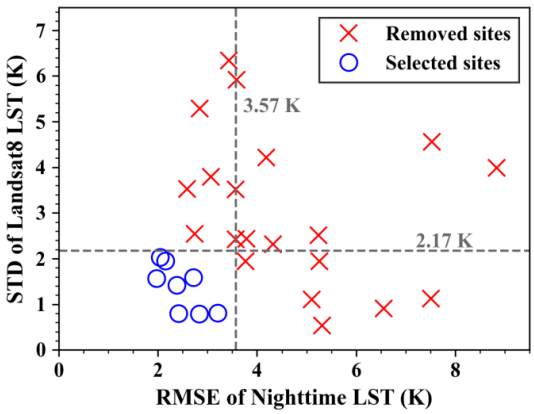

Figure 17: The spatial representativeness evaluation of the employed in situ sites.

10. Overall accuracy improvement is not convincing

The reported errors remain relatively large, and in some cases are lower than those reported for other products. Given the complexity of the workflow and the number of assumptions introduced, the

demonstrated improvement in accuracy does not convincingly justify the methodological complexity.

Response: Thank you for your constructive comments regarding the overall accuracy of our ANCFDS-LST. The relatively high methodological complexity of our approach primarily stems from the need to meet the stringent requirements of the Global Climate Observing System (GCOS), particularly in terms of achieving consistent all-weather, hourly, and 1 km LST products. To this end, the framework is composed of three sequential steps. Among them, the core innovation lies in the angular normalization component, which enables the simultaneous retrieval of nadir and hemispherically integrated LST. This is an aspect that is not addressed in most existing cloud-filling studies. The remaining components (i.e., cloud-gap reconstruction and spatial downscaling) are designed to support this objective and ensure compliance with GCOS requirements, rather than representing independent major innovations.

Following your suggestion, we have substantially strengthened the validation in the revised manuscript by implementing three key improvements: (1) evaluation of site representativeness in T-based validation; (2) implementation of bias correction in cross-validation; and (3) inclusion of quantitative metrics for assessing downscaling performance. Based on these updates, we obtained a new set of quantitative results, which demonstrate that the overall accuracy of our product is comparable to the state-of-the-art studies reported in the literature.

1) Site representativeness evaluation

As described above, we conducted a strict representativeness assessment using the nighttime RMSE and the standard deviation (STD) of Landsat-8 LST within a 5 km window as indicators. In the original manuscript, the clear-sky and cloudy-sky RMSE values were 2.99 K and 4.56 K, respectively. After excluding low-quality sites, these values improved to 2.48 K and 3.65 K, respectively, as shown in Fig. R9. The updated accuracy is comparable to that reported in recent all-weather LST studies (see Table R3).

2) Bias correction in cross validation

In the angular normalization process, bias correction between MODIS and FY-4A official LST products was performed following Ermida et al. (2017). However, the bias between VNP21A1 LST and the generated FY-4A LST was not corrected in the original manuscript, which may have introduced uncertainties due to systematic bias in the cross-validation results. To address this issue, we corrected VNP21A1 LST against FY-4A LST using nighttime matchups under the following conditions: VZA < 50°, VZA difference < 5°, and LST difference < 5 K (consistent with the bias correction applied between MOD11A1 and FY-4A LST products). We then re-evaluated the normalized nadir LST using VIIRS near-nadir LST after bias correction. The updated cross-validation results are presented in Fig. R10. The RMSD (MBD) is 4.53 K (−1.85 K) for directional LST and 2.56 K (0.06 K) for normalized nadir LST, corresponding to a 43.5% reduction in RMSD and an almost unbiased MBD (see Fig. R10a–b). This accuracy is better than that in original manuscript, which is 6.21 K (-4.04 K) before angular normalization, and 3.48 K (-2.13 K) after angular normalization. The most pronounced reductions in RMSD and MBD occur at VZA > 40° (see Fig. R10c-d), highlighting the non-negligible impact of thermal radiation directionality (TRD) correction on LST products.

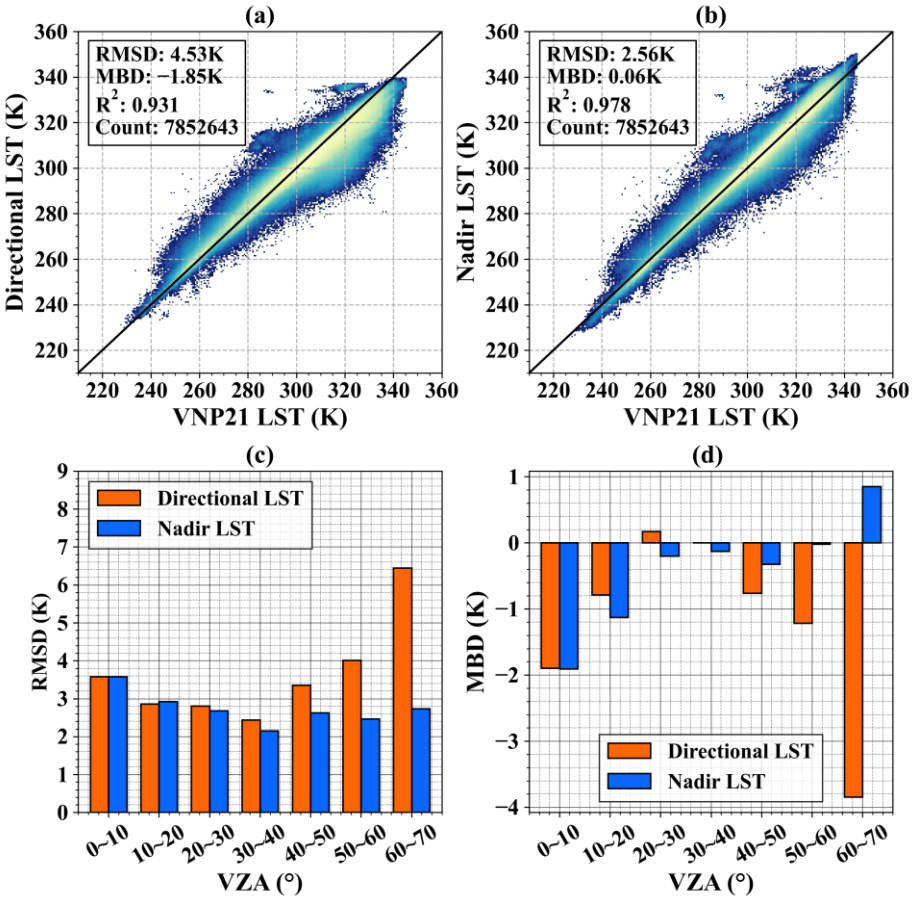

Fig. R10 The cross-validation of (a) directional and (b) nadir LST against VNP21 near-nadir LST at 0.05° resolution. The angular dependence of (c) RMSD and (d) MBD for directional and nadir LST at 0.05° resolution.

3) Quantitative metrics for evaluating spatial textures

In spatial downscaling studies, the quantitative assessment of fine-resolution spatial textures is of critical importance, yet it was lacking in the original manuscript. To address this, we employed the local variance ratio (LVR) and the structural similarity index (SSIM) as evaluation metrics to compare the spatial textures of bilinearly interpolated LST (from 0.05° data) and the downscaled 0.01° nadir LST. The reference texture was reconstructed using the ATC model, as interpolation-based ATC approaches have been shown to better preserve spatial texture characteristics (Jia et al., 2024). The evaluation results show that both LVR and SSIM values are closer to 1 (Fig. R11) for downscaled LST, indicating that the downscaled LST exhibits superior spatial texture representation.

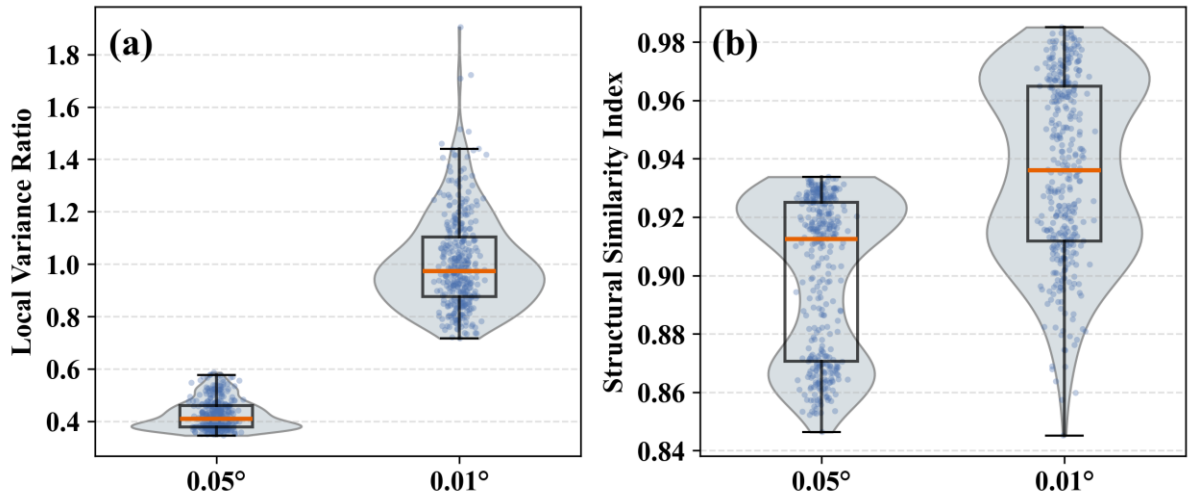

Fig. R11 The comparison of (a) local variance ratio and (b) structural similarity index for linear interpolated from 0.05° LST and downscaled 0.01° LST.

Figure R12 presents the nadir and hemispherical LST before and after downscaling. Two local zoom-in regions are highlighted, including the western Tibetan Plateau and the Greater Khingan Mountains. As shown in Fig. R12a–d, the overall spatial patterns of LST remain largely consistent before and after downscaling. However, the spatial textures are significantly enhanced after downscaling (Fig. R12e–l), particularly in terms of the contrast between the upper and lower parts of mountainous areas.

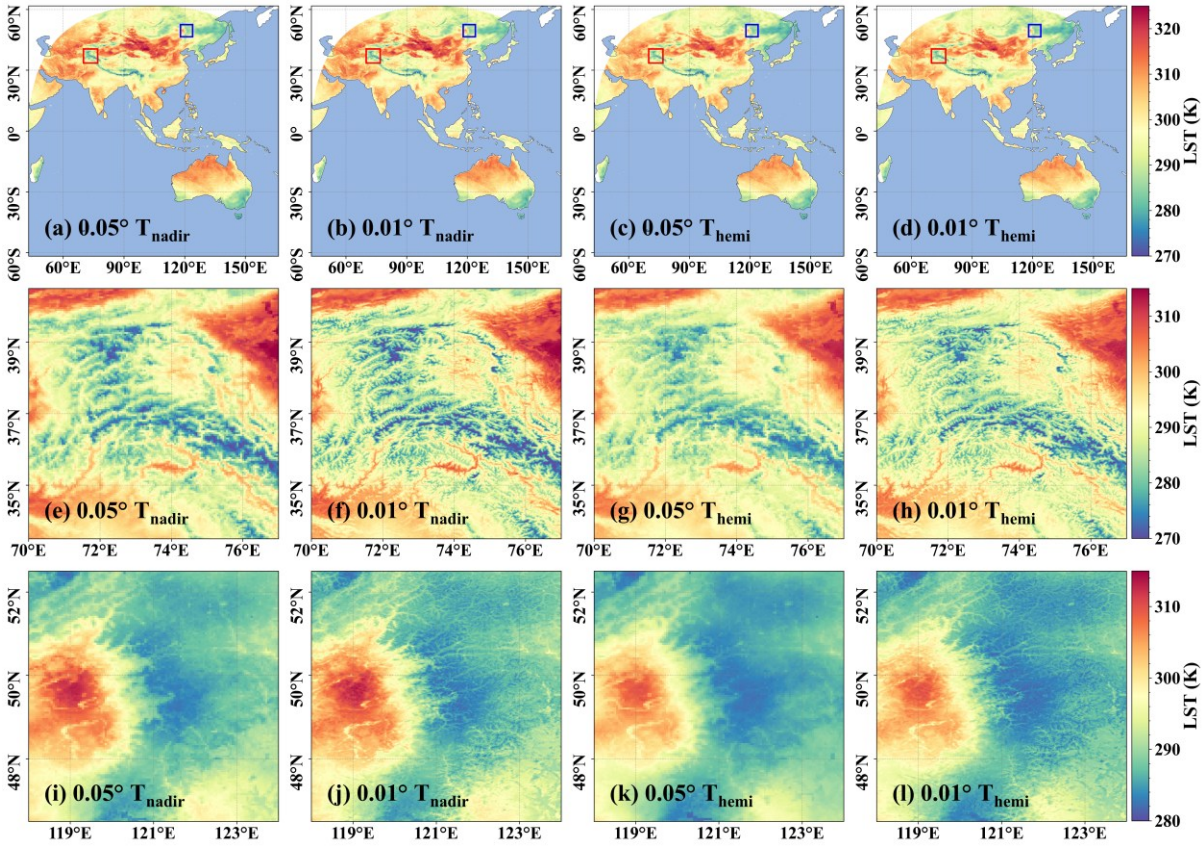

Fig. R12 The spatial distribution of 0.05° and downscaled 0.01° nadir LST/hemispherical LST in FY-4A disk (a-d), western part of Tibet Plateau (e-h), and the Greater Khingan Mountains (j-l).

Finally, we have added more detailed descriptions of the in-situ validation, cross-validation, and spatial texture evaluation procedures. The resulting accuracy is consistent with that reported in relevant studies. We also restructured the revised manuscript by exhibiting the results in three steps (i.e., angular normalization, cloudy-sky LST estimation, and spatial downscaling). Thank you again for your insightful and constructive suggestions, which have significantly improved the quality and rigor of our manuscript.

Reference:

Ermida, S.L., DaCamara, C.C., Trigo, I.F., Pires, A.C., Ghent, D., Remedios, J., 2017. Modelling directional effects on remotely sensed land surface temperature. Remote Sens. Environ. 190, 56–69. https://doi.org/10.1016/j.rse.2016.12.008

Du, H., Zhan, W., Liu, Z., Wang, C., Huang, F., 2025. A universal yet easy-to-use data-driven method for angular normalization of directional land surface temperatures acquired from polar orbiters across global cities. Remote Sens. Environ. 328, 114840. https://doi.org/10.1016/j.rse.2025.114840

Jia, A., Liang, S., Wang, D., Mallick, K., Zhou, S., Hu, T., Xu, S., 2024. Advances in Methodology and Generation of All-Weather Land Surface Temperature Products From Polar-Orbiting and Geostationary Satellites: A comprehensive review. IEEE Geosci. Remote Sens. Mag. 2–43. https://doi.org/10.1109/MGRS.2024.3421268

In Subsection 2.2:

Here, the VNP21 LST values with the VZA< 5° were employed for cross evaluation of T_{nadir} as did by Wei et al. (2025). **The resulting near-nadir VNP21 LST was first resampled to a spatial resolution of 0.05°, and then bias-corrected to FY-4A LST using nighttime matchups (consistent with the bias correction applied between MODIS and FY-4A LST products).** The normalized hourly FY-4A T_{nadir} product was temporally interpolated to match the exact overpass time of S-NPP VIIRS sensor. Finally, the root mean square difference (RMSD), mean bias difference (MBD), and coefficient of determination (R^2) were used as three evaluation indicators.

In Subsection 4.1:

There will be significant inconsistencies in directional LST between FY-4A and MODIS due to the TRD effect. Figure 5a quantifies this discrepancy by presenting the RMSD between FY-4A and MOD11A1 directional LST. After correcting both two LSTs to nadir values, the corresponding RMSD is shown in Fig. 5b. The average RMSD across the full FY-4A disk decreases from 3.26 K to 2.12 K, representing a reduction of 1.14 K (35.0%). The extremely high values near the western edge of the disk are substantially reduced after correction. However, relatively high values over the western Qinghai–Tibetan Plateau persist, which may be partly attributed to the limitation of the employed LSF-Chen model, as it is derived under the assumption of flat terrain. Overall, the TEKDM demonstrates strong performance in reducing angular inconsistencies between LST products as expected.

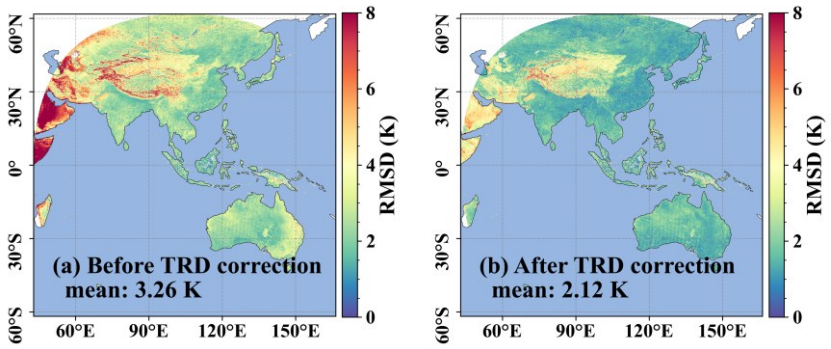

Figure 5: The spatial distribution of RMSD between FY-4A and MODIS (a) before and (f) after TRD correction.

The cross-validation between FY-4A T_{dir}, T_{nadir}, and the VNP21A1 near-nadir LST (i.e., the VZA< 5°) in 2020 is shown in Fig. 6. The root mean squared difference (RMSD) and mean bias difference (MBD) are 4.53 K and -1.85 K for T_{dir} as shown in Fig. 6a. After angular normalization, the RMSD and MBD for T_{nadir} is 2.56 K and 0.06 K as shown in Fig. 6b, with a 1.97 K (i.e., 43.5%) reduction in RMSD and almost unbiased for MBD. Fig. 6c shows the RMSD values of T_{dir} and T_{nadir} across different VZA intervals with a step of 10°. The RMSD difference between them remains within 0.2 K when VZA is less than 40°. However, when VZA exceeds 40°, the RMSD for T_{dir} increases substantially (from 4.0 K to 6.5 K), while the RMSD for normalized T_{nadir} remains lower than 3.0 K for large VZA. The higher RMSD of T_{dir} at large VZA can be partially attributed to its significant underestimation as shown in Fig. 6d. The MBD of T_{dir} exceeds -3.8 K when the VZA is between 60° and 70°, whereas it is 0.9 K for the normalized T_{nadir}. The gap fraction effect could explain this angular dependence of MBD, i.e., a larger VZA leads to a smaller T_{dir} compared with nadir LST.

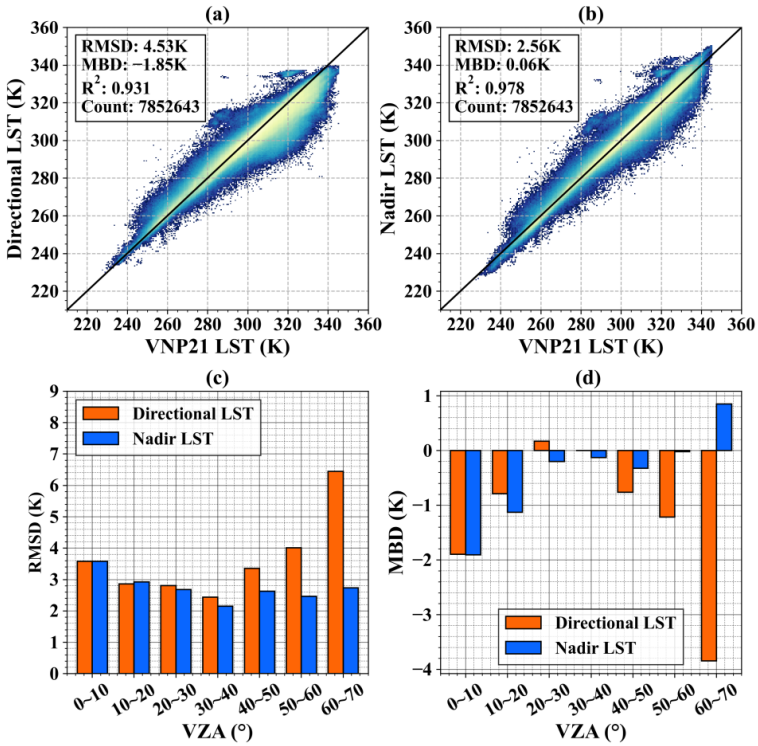

Figure 6: The cross-validation of (a) directional and (b) nadir LST against VNP21 near-nadir LST at 0.05° resolution. The angular dependence of (c) RMSD and (d) MBD for directional and nadir LST at 0.05° resolution.

In Subsection 4.3:

The spatial distributions of all-weather T_{nadir} and T_{hemi} at 0.05° and 0.01° resolutions on 2020/6/24 at 04:00 UTC (12:00 Beijing time) are presented in Fig. 14. Two representative regions are highlighted for detailed comparison: the western Tibetan Plateau (red rectangles in Fig. 14a–d) and the Greater Khingan Mountains (blue rectangles in Fig. 14a–d). As shown in Fig. 14a–d, the large-scale spatial patterns of LST are well preserved after downscaling, with both T_{nadir} and T_{hemi} exhibiting clear latitudinal and elevational gradients (i.e., lower temperatures are consistently observed in high-latitude and high-altitude regions). While T_{nadir} represents the directional LST observed at nadir viewing geometry, T_{hemi} corresponds to hemispherical LST, integrating directional effects over all viewing angles and thus generally showing smoother spatial variability. At finer scales (Fig. 14e–l), the downscaled results (0.01°) reveal substantially enhanced spatial details compared to the coarse-resolution counterparts. In particular, complex terrain features become more distinguishable, such as ridge–valley structures over the Tibetan Plateau and the heterogeneous forest–grassland patterns in the Greater Khingan Mountains. The contrast between sunlit and shaded slopes, as well as between higher and lower elevations, is more pronounced in the downscaled LST fields. These improvements in spatial texture and terrain-related variability demonstrate the effectiveness of the IHDA downscaling method in capturing fine-scale thermal heterogeneity.

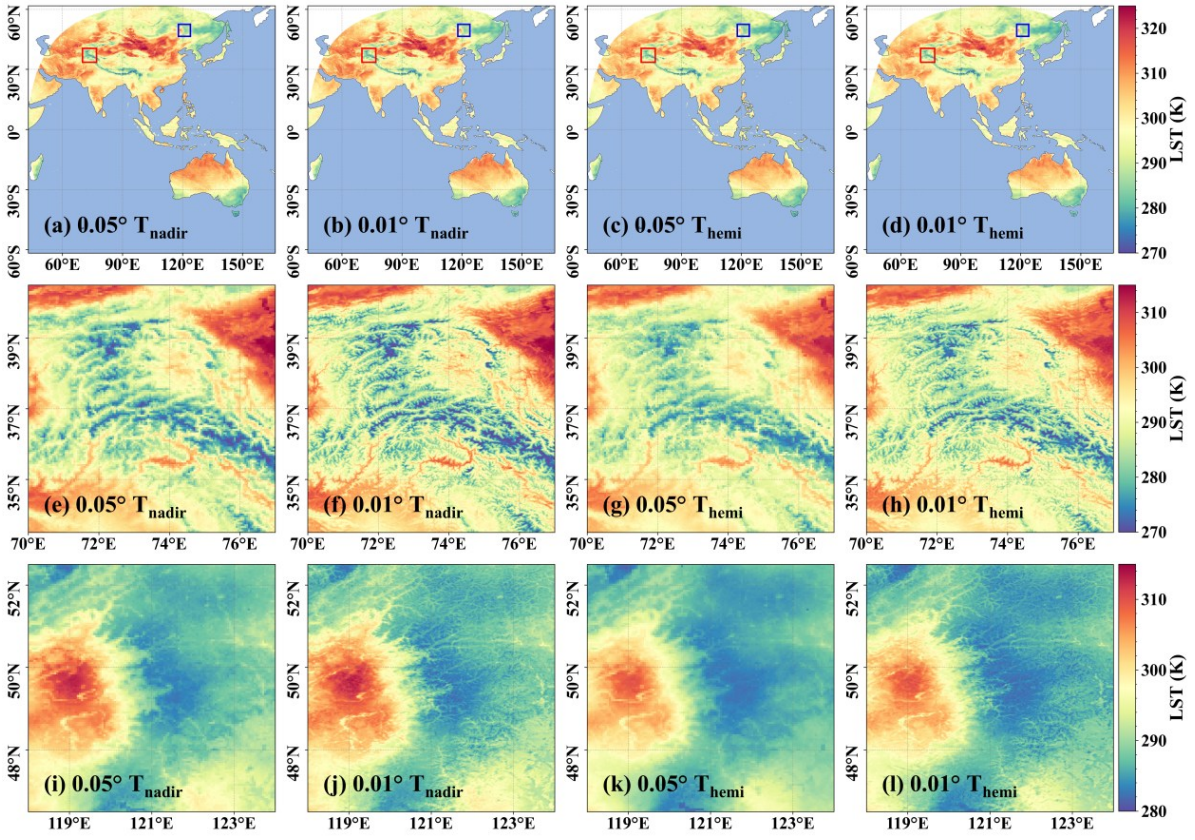

Figure 14: Spatial distributions of nadir LST at 0.05° (first column) and 0.01° (second column), and hemispherical LST at 0.05° (third column) and 0.01° (fourth column) at the UTC time of 4:00 on 2020/6/24. The first row shows LST over the FY-4A full disk, the second row over the western Qinghai–Tibetan Plateau, and the third row over the Greater Khingan Mountains.

Minor comments

1. **researches -> research**

We revised accordingly:

In line 25:

Land surface temperature (LST) is an essential climate variable in geophysical, ecological, and environmental **research**.

2. **suffer the spatial discontinuities -> suffer from**

We revised accordingly:

In line 26:

…, and **suffer from** the spatial discontinuities due to the pervasive presence of clouds

3. **"What's more" is informal in academic English**

We revised accordingly:

In line 27:

Moreover, the geostationary LST products have relatively

4. **hypothetical clear-sky LST were -> "was"**

We revised accordingly:

In line 32:

Subsequently, hypothetical clear-sky LST **was** predicted using a …

5. **combining fusion and kernel-based -> "combining fusion- and kernel-based"**

We revised accordingly:

In line 35:

an improved hybrid downscaling algorithm (IHDA) combining **fusion-** and kernel-based methods.

In line 139:

Finally, an improved hybrid downscaling algorithm (IHDA) combining **fusion-** and kernel-based methods was developed and carried out with the input of ATC-simulated gap-free 0.01° LST

6. **"jointly estimation" is wrong**

We revised accordingly:

In line 166:

Both FY-4A and MODIS LST products were resampled to the same spatial resolution (i.e., 0.05°) using plain average before the **joint estimation**.

7. **"resampled to a same spatial resolution" -> resampled to the same spatial resolution**

We revised accordingly:

In line 165:

Both FY-4A and MODIS LST products were resampled to **the** same spatial resolution (i.e., 0.05°) using plain average before the joint estimation.

8. **"with the conditions of" -> "under the condition"**

We revised accordingly:

In line 169:

The parameters a and b were determined based on nighttime matchups **under the condition** of VZA< 50°, VZA difference< 5°, and LST difference< 5 K, referenced to Ermida et al. (2018, 2017).

9. **"The VNP21 LST values … was employed" -> "were"**

We revised accordingly:

In line 229:

Here, the VNP21 LST values with the VZA< 5° **were** employed for cross-evaluation of Tnadir as done by Wei et al. (2025).

10. **"as did by" -> done**

We revised accordingly:

In line 230:

Here, the VNP21 LST values with the VZA< 5° were employed for cross-evaluation of Tnadir as **done** by Wei et al. (2025).

11. **"cross evaluation" -> "cross- evaluation"**

We revised accordingly:

In line 230:

Here, the VNP21 LST values with the VZA< 5° were employed for **cross-evaluation** of Tnadir as done by Wei et al. (2025).

12. **"in southern hemisphere" -> "the southern hemisphere"**

We revised accordingly:

In line 237:

It is relatively lower in 2020/3/1 and 2020/12/1 than that in 2020/6/1 and 2020/9/1 in the northern hemisphere and the situation is reversed in **the** southern hemisphere.

13. **"To abundant the angular information" ?**

We revised accordingly:

In line 309:

To enhance the angular information content of the input data, two daytime clear-sky MODIS observations (i.e., TERRA and AQUA) were utilized.

14. **"ML-based LST reconstructing have been widely used" wrong sentence**

We revised accordingly:

In line 330:

ML-based LST reconstruction methods have been widely used because it could accurately simulate the non-linear correlation

15. **"over fitting" -> "overfitting"**

We revised accordingly:

In line 335:

2) Ordered boosting, which minimizes gradient bias by employing permutation-driven training to reduce

overfitting;

16. data-sets -> datasets

We revised accordingly:

In line 337:

These improvements enhance the model's robustness and efficiency, especially when handling large-scale **datasets** containing categorical variables, and have proven effective in LST reconstruction studies.

17. in the right of -> "on the right-hand side of"

We revised accordingly:

In line 411:

The second term **on the right-hand side of** Eq. 20 is the CatBoost modeling residual to be redistributed.

18. widely-used -> widely used

We revised accordingly:

In line 414:

was resampled to 0.01° using the **widely used** bilinear interpolation.