Dear Dr. Tian,

Thank you so much for overseeing the review of our manuscript. We would like to thank the anonymous reviewers for their positive and constructive comments. The manuscript has greatly benefited from their insightful suggestions. We have carefully revised our manuscript to account for the recommendations by reviewers. The key improvements are as follows.

1. To enhance the credibility of the method and resulting TRISM LST dataset, we significantly improved the manuscript, especially in the Methods, Results, and Discussion sections. Specifically, we improved the descriptions of the random forests (RF)-based modeling and corresponding core concepts.

2. In response to the reviewers' suggestions, we added an introduction explaining the basic situation of RTM methods. Moreover, we clarified that this work succeeds Zhang et al. (2021).

3. We have refined the validation and comparison sections of the manuscript. Our intention with these revisions was to improve the quality of the TRIMS LST dataset and enhance the overall readability of the paper.

We sincerely appreciate the time and effort you put into editing and reviewing the manuscript. Your valuable input has played a significant role in improving both the dataset and the manuscript. We are looking forward to collaborating with you to bring the manuscript closer to publication in Earth System Science Data.

Sincerely yours,

Wenbin Tang, Ji Zhou on behalf of all authors

School of Resources and Environment

University of Electronic Science and Technology of China

Chengdu, Sichuan Province, 611731, China

Emails: wenbint@std.uestc.edu.cn; jzhou233@uestc.edu.cn

General Comment:

The authors responded to certain inquiries I raised, yet the feedback also prompted the emergence of further concerns. Kindly review my comments below for clarification.

Response: Thank you very much for your interest in our study and for providing valuable feedback. We greatly appreciate your reminder regarding the methodology and comparison sections of the manuscript. We have carefully considered your suggestions to further improve these sections, aiming to enhance the quality of TRIMS LST and the readability of the paper. Please see our responses below. (The changes are highlighted in red color in our revised manuscript.)

Major issues:

(1) Comment 1

In the response letter, I highlighted a significant issue pertaining to the application of the Terra LST model to Aqua LST. The authors' explanation mentioned that “Firstly, the LST of the same pixel at different time points during a given day satisfies the standard conditions of similar pixels in the RTM method (Zhang et al., 2021). This forms the theoretical foundation for the conversion between Terra/MODIS LST and Aqua/MODIS LST (Li et al., 2018).” I'm struggling to grasp the meaning of "satisfied the standard condition of similar pixels in the RTM method." Could you provide a comprehensive explanation of this concept? Furthermore, it's important for the manuscript to explicitly address and rectify such conceptual flaws for future enhancements.

Response: Thank you very much for your comment. As a basis for this study, similar pixel determination criteria were provided by Zhang et al. (2019) and Zhang et al. (2021): (i) they are located within the sliding window; (ii) they have the same land cover as the target pixel (the correlation coefficient between the MODIS LST time series of target pixel and a similar pixel is over 0.8); and (iii) they are under the same

weather condition (clear sky or unclear sky) as the target pixel.

In the TSETR approach, when calculating the HFC, due to the lack of Aqua/MODIS LST, we used the information provided by Terra LST and the corresponding GLDAS LST of Terra/Aqua. Accordingly, similar pixels were determined according to the following conditions: **(i) it should have the same land cover type as M and (ii) the R of the Terra-MODIS LST time series corresponding to S and M need to be greater than 0.8.**

We have clarified the concept of "similar pixels" in the revised manuscript to avoid confusion for readers and users.

"*$f_{M\text{-}T2}$ is constructed as follows. Initially, the correlation image of the target pixel M is determined within the T2 period and the following two conditions need to be satisfied by the correlation image: (i) the mean bias deviation (MBD) of the DTC estimated from its corresponding GLDAS LST (10:00-14:00 and 21:00-3:00 local solar time) should be lower than 1 K, and (ii) the difference in the average observation time between the GLDAS pixels should not exceed 0.5 h. Using the correlation image, the similar image family S of the target pixel M is determined. Subsequently, in the correlation image, using similar land cover type criteria, the similar image family S of the target pixel M within the GLDAS pixels is identified. S needs to meet the following two conditions: (i) it should have the same land cover type as M and (ii) R of the Terra/MODIS LST time series corresponding to S and M need to be greater than 0.8.*"(Lines 273–280)

More details and explanations are shown as follows.

(1) In the original manuscript, as demonstrated in Fig. A, we thought that MODIS LSTs at different moments on the same day could satisfy the **similar pixel determination condition** (Zhang et al., 2021). HFC represents the LST change between t_{avg} and t_{ins} due to the Earth's rotation under clear conditions. So far, there is no available physical model to determine HFC from remote sensing observations due to the impossibility of satisfying the input requirement of the HFC-involved land surface models. However, there is a potential solution. As pointed out by Zhan et al. (2012; 2014), it is reasonable to describe HFC using its descriptors through empirical or implicit functions.

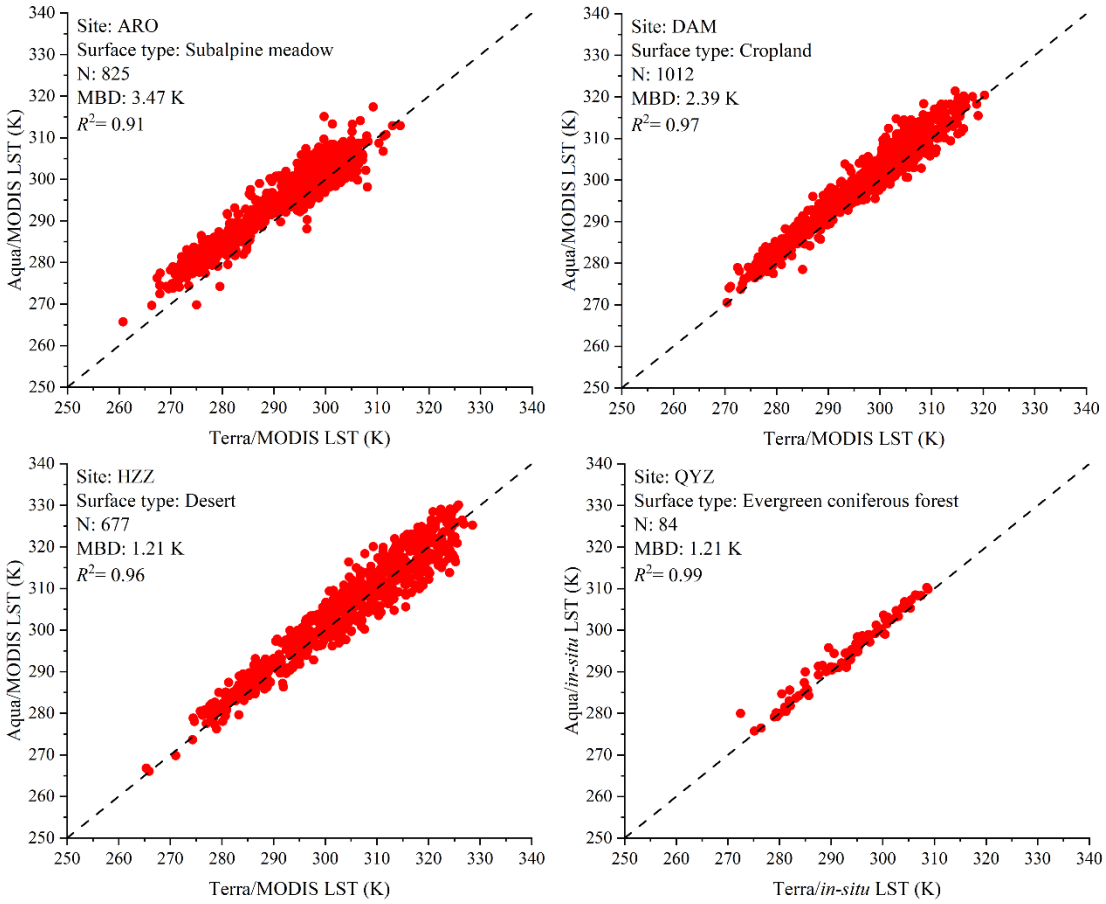

Fig. A. Scatterplot of Aqua/MODIS LST and Terra/MODIS LST during the daytime.

The Aqua/Terra MODIS LST results for the four distinct land surface types of sites are depicted in Fig. A. To compensate for the limited QYZ data, in-situ LSTs have been implemented instead in Fig. A. The findings demonstrate a correlation coefficient greater than 0.9 between the time series of both Aqua/MODIS LST and Terra/MODIS LST. This is the theoretical foundation for the prior study that converted Terra/MODIS LST and Aqua/MODIS LST (Li et al., 2018). Furthermore, the significance of HFC and the parameterization scheme anticipates a robust correlation between similar pixels. the factors employed to construct the mapping model effectively capture the impact of diurnal LST variation (Δt_M) and weather variation (v_M). Therefore, our original manuscript explored the possibility of estimating Aqua/HFC by utilizing $f_{M\text{-}Terra\text{-}T1}$:

$$HFC_{\text{M-Aqua-T1}}(t_d, t_{ins}) = f_{\text{M-Terra-T1}}(lat_M, lon_M, DEM_M, NDVI_M(t_d), slp_M(t_d), \alpha_M(t_d), \Delta t_M(t_d), v_M(t_d)), \quad (A)$$

where lat_M, lon_M, DEM_M, $NDVI_M$, slp_M, α_M, Δt_M, and v_M are the latitude, longitude, DEM, NDVI, slope, albedo, difference between t_{ins} and t_{avg} , and the atmospheric water

vapour content, respectively; and $f_{M\text{-}Terra\text{-}T1}$ is an RF mapping model based on Terra MODIS data and its corresponding descriptors in T1.

(2) However, we have failed to address a significant issue that you have raised due to the lack of HFC estimation research: the disparity in timing between Terra and Aqua observations that results in distinct fluctuations of LST. The observation window for Terra is typically from 10:00 a.m. to 12:00 p.m. local solar time. According to Duan et al. (2014a), LST changes linearly within that time. Aqua exhibits a non-linear change in LST over the observation period, as depicted by the DTC curves (Duan et al., 2014b; Jin and Dickinson, 1999).

In the original RTM method, the HFC of a similar pixel (S) can be expressed as:

$$HFC_{\mathrm{S}}(t_{\mathrm{d}},t_{\mathrm{ins}}) = \mathbf{RF}(lat_{\mathrm{S}},lon_{\mathrm{S}},DEM_{\mathrm{S}},NDVI_{\mathrm{S}},slp_{\mathrm{S}},\alpha_{\mathrm{S}},\theta_{\mathrm{s\text{-}S}},\theta_{\mathrm{m\text{-}S}},\Delta t_{\mathrm{S}},v_{\mathrm{S}})+\varepsilon \quad \text{(B)}$$
$$\text{where } \mathrm{S}_n \in \mathrm{S} \subseteq \mathrm{W1}$$

where *lat, lon, DEM, NDVI, slp,* α, θ_{s}, θ_{m}, Δt, and Δv are latitude, longitude, DEM, NDVI, slope, surface albedo, solar zenith angle, MODIS observation angle, difference between instantaneous and intraannual average observation time of MODIS, and atmospheric water vapor, respectively. If the descriptors in this equation accurately and adequately describe the situation, the approximation error ε can be minimized.

At this stage, the RF, which is built using similar pixels, can be readily employed to estimate the HFC of the target pixel M:

$$HFC_{\mathrm{M}}(t_{\mathrm{d}},t_{\mathrm{ins}}) = \mathbf{RF}(lat_{\mathrm{M}},lon_{\mathrm{M}},DEM_{\mathrm{M}},NDVI_{\mathrm{M}}(t_{\mathrm{d}}), slp_{\mathrm{M}}(t_{\mathrm{d}}),\alpha_{\mathrm{M}}(t_{\mathrm{d}}),\theta_{\mathrm{s\text{-}M}}(t_{\mathrm{d}}),\theta_{\mathrm{m\text{-}M}}(t_{\mathrm{d}}),\Delta t_{\mathrm{M}}(t_{\mathrm{d}}),v_{\mathrm{M}}(t_{\mathrm{d}})) \quad \text{（C）}$$

According to the original RTM method, obtaining HFC requires a defined number of Aqua/MODIS LSTs as training samples, which are unfortunately unavailable between DOY 1 of 2000 and DOY 184 of 2002. However, TRIMS-Terra LSTs obtained with Module I and Module II for the period T1 provide an opportunity to establish a transformation relation to obtain the Aqua/HFC. The precision of the HFC estimations remains unaffected by the exclusion of solar radiation as the primary contributor to warming, as indicated in Eq. (C). This outcome can be attributed to the utilization of Aqua MODIS LST during modeling, which directly conveys the highest LSTs that

occur during the day. In addition, the temporal gradient of DTC is mostly at its minimum during this period.

When there is no valid Aqua LST available, we can enhance Eq. (A) as follows：

$$\begin{cases} HFC_{\text{M-Aq-T1}}(t_{\text{d}}, t_{\text{ins-T1}}) = HFC_{\text{M-Te-T1}}(t_{\text{d}}, t_{\text{ins-T1}}) + \Delta HFC_{\text{M-Te-Aq-T1}}(t_{\text{d}}, t_{\text{ins-T1}}) \\ \Delta HFC_{\text{M-Te-Aq-T1}}(t_{\text{d}}, t_{\text{ins-T1}}) = \text{f}_{\text{M-T2}}(g_{\text{M}}, DEM_{\text{M}}, NDVI_{\text{M}}(t_{\text{d}}), slp_{\text{M}}(t_{\text{d}}), \alpha_{\text{M}}(t_{\text{d}}), v_{\text{M}}(t_{\text{d}}), \Delta LFC_{\text{M}}, \Delta DTC_{\text{M}}) \\ \Delta DTC_{\text{M}} = \Delta DTC_{\text{M-Aq}}(t_{\text{d}}, t_{\text{ins-T1}}, t_{\text{avg-T1}}) - \Delta DTC_{\text{M-Te}}(t_{\text{d}}, t_{\text{ins-T1}}, t_{\text{avg-T1}}) \end{cases} \quad \text{(D)}$$

where ΔLFC characterizes the systematic deviation of the steady state component, ΔDTC characterizes the warming effect of solar radiation, and the weather effect can be characterized by the atmospheric water vapor content. According to Zhang et al. (2021), the HFC characterizes the change in LFC with ΔDTC and WTC superimposed under ideal clear-sky conditions. The detailed calculation of ΔDTC can be found in Zhang et al. (2019).

(3) Finally, we can analyze how the similar pixel concept is used in all-weather LST research. For example, Long et al. (2020) utilized the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) (Zhu et al., 2010), which is primarily applied to albedo and NDVI products, to generate all-weather LST. According to Zhu et al. (2018), ESTARFM is categorized as a weighting model because it assists in calculating the value of the target pixel by assigning weights to neighbouring similar pixels. First, the target pixel corresponding to the similar pixel is determined, typically within a moving window of a particular size. Then, the weight of each similar pixel is computed.

In the spatiotemporal fusion of reflectance and NDVI products, **pixels within the window with the same land cover type as the target pixel (center pixel) are defined as similar pixels**. There are two primary methods for determining similar pixels in finer-resolution images. The first involves applying an unsupervised clustering algorithm to the image, identifying neighboring pixels within the same cluster as the central pixel. The second method calculates the reflectance difference between neighboring pixels and the central pixel in the fine-resolution image, using specified thresholds to identify those that are similar. The thresholds can be determined by the standard deviation of a population of pixels from the high-resolution image and the

estimated number of land-covering categories present in the image. (Gao et al., 2006; Zhu et al., 2010; Long et al., 2020).

Since LST exhibits different characteristics from reflectance and NDVI products, the determination conditions for similar pixels in all-weather LST estimation need to be targeted and modified.

Reference:

Duan, S.-B., Li, Z.-L., Tang, B.-H., Wu, H., and Tang, R.: Generation of a time-consistent land surface temperature product from MODIS data, Remote Sens. Environ., 140, 339–349, https://doi.org/10.1016/j.rse.2013.09.003, 2014a.

Duan, S.-B., Li, Z.-L., Tang, B.-H., Wu, H., Tang, R., Bi, Y., and Zhou, G.: Estimation of Diurnal Cycle of Land Surface Temperature at High Temporal and Spatial Resolution from Clear-Sky MODIS Data, Remote Sens., 6, 3247–3262, https://doi.org/10.3390/rs6043247, 2014b.

Gao, F., Masek, J., Schwaller, M., and Hall, F.: On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance, IEEE Transactions on Geoscience and Remote Sensing, 44, 2207–2218, https://doi.org/10.1109/TGRS.2006.872081, 2006.

Jin, M. and Dickinson, R. E.: Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 1. Without clouds, Journal of Geophysical Research: Atmospheres, 104, 2105–2116, https://doi.org/10.1029/1998JD200005, 1999.

Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., Yang, H., Tian, F., Zeng, C., Meng, X., and Shi, C.: Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach, Remote Sensing of Environment, 246, 111863, https://doi.org/10.1016/j.rse.2020.111863, 2020.

Zhan, W., Chen, Y., Voogt, J. A., Zhou, J., Wang, J., Ma, W., and Liu, W.: Assessment of thermal anisotropy on remote estimation of urban thermal inertia, Remote Sensing of Environment, 123, 12–24, https://doi.org/10.1016/j.rse.2012.03.001, 2012.

Zhan, W., Zhou, J., Ju, W., Li, M., Sandholt, I., Voogt, J., and Yu, C.: Remotely sensed soil temperatures beneath snow-free skin-surface using thermal observations from tandem polar-orbiting satellites: An analytical three-time-scale model, Remote Sensing of Environment, 143, 1–14, https://doi.org/10.1016/j.rse.2013.12.004, 2014.

Zhang, X., Zhou, J., Göttsche, F.-M., Zhan, W., Liu, S., and Cao, R.: A Method Based on Temporal Component Decomposition for Estimating 1-km All-Weather Land Surface Temperature by Merging Satellite Thermal Infrared and Passive Microwave Observations, IEEE Transactions on Geoscience and Remote Sensing, 57, 4670–4691, https://doi.org/10.1109/TGRS.2019.2892417, 2019.

Zhang, X., Zhou, J., Liang, S., and Wang, D.: A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature, Remote Sensing of Environment, 260, 112437, https://doi.org/10.1016/j.rse.2021.112437, 2021.

Zhu, X., Cai, F., Tian, J., and Williams, T. K.-A.: Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions, Remote Sensing, 10, 527, https://doi.org/10.3390/rs10040527, 2018.

Zhu, X., Chen, J., Gao, F., Chen, X., and Masek, J. G.: An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions, Remote Sensing of Environment, 114, 2610–2623, https://doi.org/10.1016/j.rse.2010.05.032, 2010.

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

(2) Comment 2

At line 227 of the revised manuscript, I interpreted this sentence that RF models in

module 2 appear to have been built individually for each pixel. This approach raises concerns about potential inefficiencies. How did you manage to ensure that all these models do not fall victim to overfitting? Additionally, the reliance on soil moisture data as input sparks concerns regarding the true significance of deep soil moisture contributions to models (even those from 40 cm deep). To bolster the validity of the input data selection, it would be beneficial to include a quantitative analysis in this regard.

Response: Thanks very much for your comment.

(1) To address concerns about inefficiency, it should be noted that although training a per-pixel RF model is a relatively quick process, implementing it by looping through the entire study area (4255*5213 pixels) in code would consume significant time. Therefore, a multi-threading technique is essential for processing. Using Matlab as an example, pixel positions in the image matrix are linearly indexed to enable the conditions for the operation of 'parfor' to be met. On a PC platform configured in this manner (as displayed in Table R1), the processing time for a year's worth of data is approximately 16 hours when implementing ‘parfor’ as opposed to 'for'. It is essential to note that, for this specific application of pixel-by-pixel metamodelling, the constraints imposed by the computer's memory size are of greater significance. This is because the data for a specific image pixel is relatively minimal, and consequently does not require extensive computational resources.

Table R1: Hardware information of the computer

CPU	Inter (R) Xeon(R) Platinum 8179M CPU@2.40GHz (2

	processors) Kernel: 52 Logical Processors: 52 L1 Cache: 3.2 M L2 Cache: 52.0 M L3 Cache: 71.5 M
RAM	96.0 G (+105G virtual memory)

Any machine learning algorithm is susceptible to overfitting. TRIMS LST products were developed by training numerous RF models, which have regularly demonstrated indications of overfitting. To provide a clearer response, I found the subsequent paragraph on the personal website of Leo Breiman, RF's creator:

'Random forests does not overfit. You can run as many trees as you want' (https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#remarks).

RF is not susceptible to overfitting and mainly depends on three random procedures. Firstly, the decision tree samples are chosen at random. Secondly, the decision tree's eigenvalues are randomly selected. Finally, fission is determined by selecting a random one from the N best directions during the tree generation process. As the number of trees generated by a Random Forest approaches infinity, the large number theorem theoretically proves that the training and testing errors converge. However, the generation of an infinite number of decision trees is impractical, and the presence of noise in actual data cannot be overlooked. Therefore, the overfitting of results will be affected to different degrees due to the challenge of setting model parameters within the same runtime.

(2) To mitigate overfitting, as previously noted, we implemented a hyperparameter approach to determine the optimum parameters of our RF model. The two-fold objective of this method is to prevent overfitting and enhance operational efficiency.

To determine the most effective parameters, we undertake a two-step process. Initially, we specify a comprehensive range for every hyperparameter requiring optimization. As

an illustration, we might stipulate a range extending from 10 to 5000 for the "n estimators" hyperparameter. Subsequently, we utilize an optimization algorithm to explore the ideal value of each hyperparameter inside this range. This requires conducting numerous cross-validations for every parameter combination in the given range and choosing the combination with the greatest mean score as the optimal parameter.

(3) To answer your question about the contribution of soil moisture (particularly 40cm below the surface). We use the Mean Decrease Accuracy (MDA) technique to determine the significance of a variable. This is achieved by randomly swapping the value of a feature in permuted out-of-bag (OOB) data and then recalculating the prediction to determine the importance of the feature. This is done by measuring the level to which the regression accuracy decreases (Biau and Scornet, 2016).

For the regression process, MDA is expressed as:

$$MDA_r(j)=\frac{1}{T}\sum_{t}^{T}\left[\frac{1}{|D_t|}\left(\sum_{X_i^j\in D_i^j}\sum_{k}\left(R_k\left(X_i^j\right)-y_i^k\right)^2-\sum_{X_i\in D_t}\sum_{k}\left(R_k\left(X_i\right)-y_i^k\right)^2\right)\right] \quad (E)$$

Where T represents the number of randomly generated trees, X_i denotes the input data, y_i corresponds to the output regression result, X_t^j refers to the input data after the j-th random feature exchange, D_t signifies the OOB sample set of the random tree t, D_t^j denotes the sample set after the j-th random feature exchange, and $R(X_i)$ predicts the output corresponding to X_i.

Eq. (E) indicates the degree to which the precision of the model regression reduces following the haphazard substitution of data on the j-th dimension (feature) of the sample, applying OOB data for each tree. If the accuracy reduces more after this random substitution exercise, the feature is deemed more significant. $MDA_r(j)$ represents the increase in the OOB sample mean square error (MSE) when feature j is randomly replaced.

Compared to the mean decrease in impurity (MDI) (Louppe et al., 2013; Li et al., 2021), MDA is a relatively slow method. The time taken to run the MDA is much greater than the time taken for random forest training. To achieve our objective and reduce the time

required, we adopted the following strategy to implement the MDA method: 100,000 land pixels were randomly selected and their corresponding RF models were constructed. The MDA results of the RF model corresponding to all pixels are shown in Fig. B.

Based on the experimental findings, it could be concluded that soil moisture has a minor impact and can be disregarded when reducing data. Conversely, soil temperature has a significant effect. Fig. B reveals that the highest seven components of the MDA score contribute over 97%. Consequently, if the reader or user wishes to replicate our methodology in the same zone, selecting solely the top seven ranked factors for modeling is practicable.

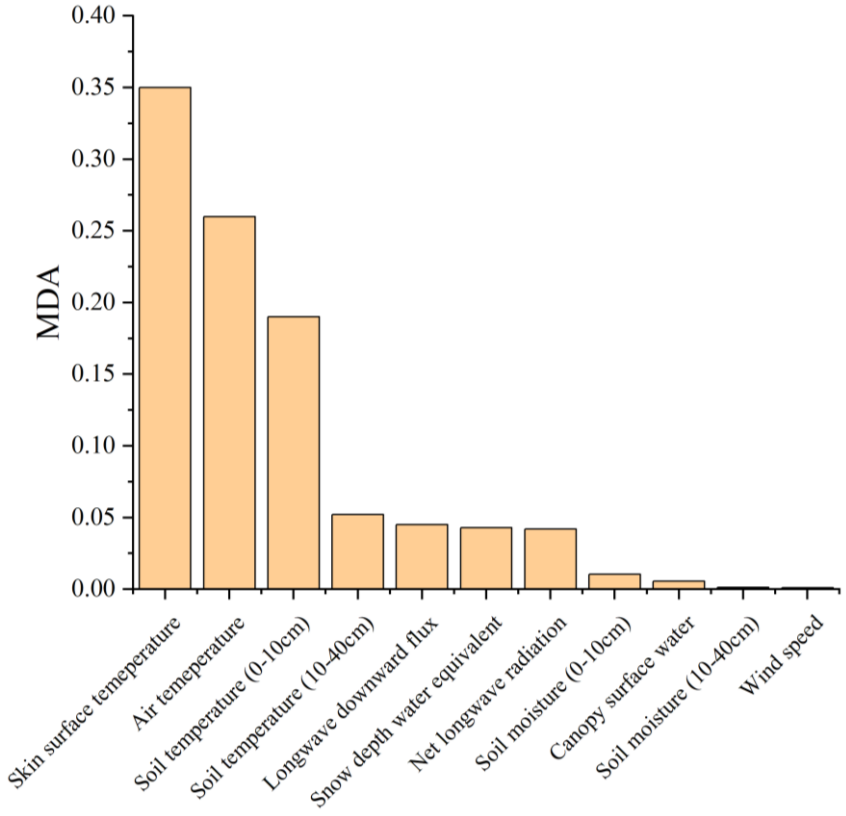

Fig. B. MDA results of RF models.

Reference:

Biau, G. and Scornet, E.: A random forest guided tour, TEST, 25, 197–227,

https://doi.org/10.1007/s11749-016-0481-7, 2016.

Li, B., Liang, S., Liu, X., Ma, H., Chen, Y., Liang, T., and He, T.: Estimation of all-sky 1 km land surface temperature over the conterminous United States, Remote Sensing of Environment, 266, 112707, https://doi.org/10.1016/j.rse.2021.112707, 2021.

Louppe, G., Wehenkel, L., Sutera, A., and Geurts, P.: Understanding variable importances in Forests of randomized trees, in: Advances in Neural Information Processing Systems, 2013.

(3) Comment 3

Figure 10 presents the temporal continuity analysis, but the inclusion of the annual cycle seems to obscure the clarity of the data series. I propose incorporating figures that focuses solely on anomaly variations after removing the annual cycles. More critically, it is advisable to specify the connection period (around March 2000 for Terra and July 2022 for Aqua) for meticulous analysis, particularly to demonstrate the absence of series disruptions at the conclusion of the filled temporal span.

Response:Thanks very much for your comments.

This will significantly enhance the quality of the article and manuscript. After thorough consideration, we decided to retain Fig. 10. Additionally, to better illustrate the atypical changes in the data and establish that there is no gap in the sequence at the end of the temporal gap (around March 2000 for Terra and about July 2002 for Aqua), we have included the subsequent findings and analyses.

(1) we determined percentage of valid pixels in TRIMS LST and MODIS LST, respectively (Fig.C). The findings reveal that TRIMS LST is spatio-temporally continuous during the temporal gaps.

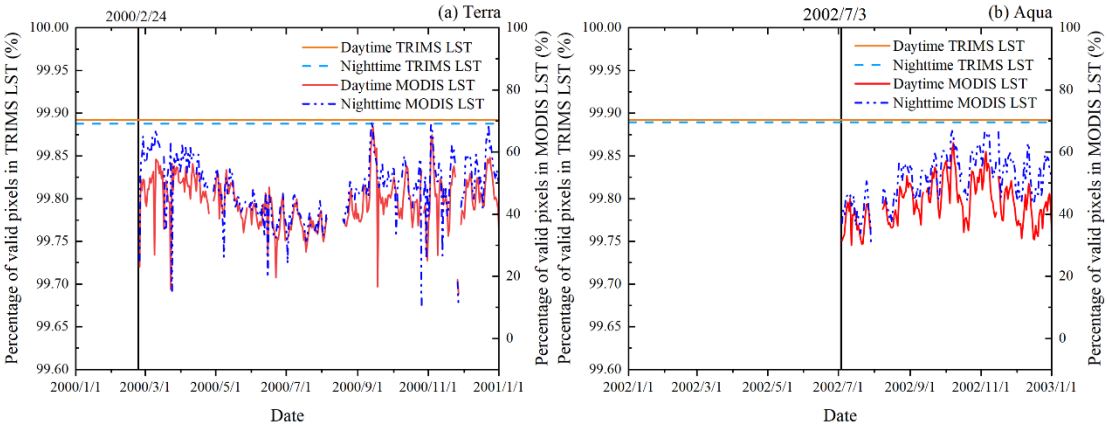

Fig. C. Percentage of valid pixels in MODIS LST and TRIMS LST in 2000 (a) and 2002 (b).

The percentage of valid pixels in MODIS LST ranges from approximately 10% to 70%, exhibiting substantial seasonal fluctuations. The rise in water vapour, heightened convection, and increased cloud cover during summers could account for the reduced number of effective pixels. This condition also clarifies why in most circumstances, fewer valid pixels are evident throughout the daytime than at nighttime. By comparison, the number of valid pixels in TRIMS LST changes moderately over time. Approximately 1‰ pixels were left unoccupied for a few days, likely due to the unavailability of reference groups and corresponding pixels within the search window during the determination of HFC_{cld} for these pixels. However, the percentage of valid pixels almost reached 100% following the combination of the RFSTM and TSETR approach. This phenomenon is a quantitative demonstration of the success of the E-RTM method in recovering unspecified LSTs during the temporal gaps.

(2) Our analysis focused on examining the temporal variations in LST during the connectivity period (February and March 2000 for Terra; June and July 2002 for Aqua). The outcomes reveal that there is no interruption in the sequence at the conclusion of the filled duration (Fig.D).

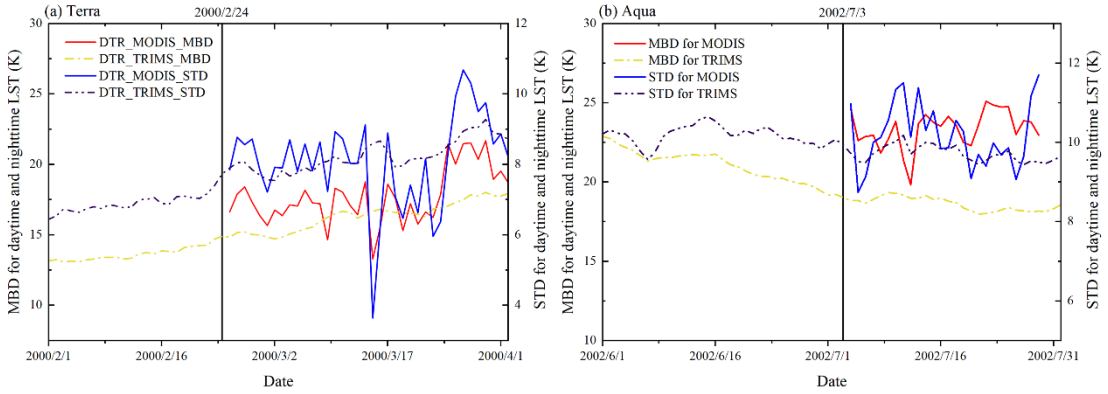

Fig. D. MBD and STD for daytime LST compared to nighttime.

During the connectivity period, we calculated MBD and STD for daytime LST compared to nighttime. The MBD and STD for MODIS LST exhibited large fluctuations between dates, whereas the MBD/STD of TRISM LST showed smoother overall trends with less fluctuation between dates. This is attributed to the consideration of LFC as the primary base element in the E-RTM method, particularly during the temporal gap when valid MODIS LST data is lacking. The trends in MBD/STD for TRIMS LST and MODIS LST are generally consistent outside of the temporal gaps. Specifically, between 24 February and 31 March 2000, MBD/STD demonstrated a general upward trend, while between 3 July and 31 July 2002, it showed an overall downward trend. Importantly, the trend of MBD/STD changes before and after the connecting dates is continuous without abrupt changes or breaks, as depicted in Fig. D, indicating uninterrupted LST time series during the temporal gaps.

Tech issues:

(4) Comment4

Will this dataset keep updating, or it will stop at 2022?

Response: Thanks very much for your comment. We will annually update the TRIMS LST dataset to the public, as long as relevant inputs of our model are available. We have added relevant descriptions in the revised manuscript.

"*The TRIMS LST dataset will be made publicly available on an annual basis, contingent on availability of pertinent input data for the model*." (Lines 619–620)

It is noteworthy that the Terra and Aqua satellites are anticipated to be decommissioned

in 2026 as they age, resulting in a reduction in orbital altitudes and significant deviations between equator crossing times and those established during launch (Fig. E). As a result, it will be necessary to consider alternative satellite data, such as VIIRS, to replace MODIS in 2025 and beyond.

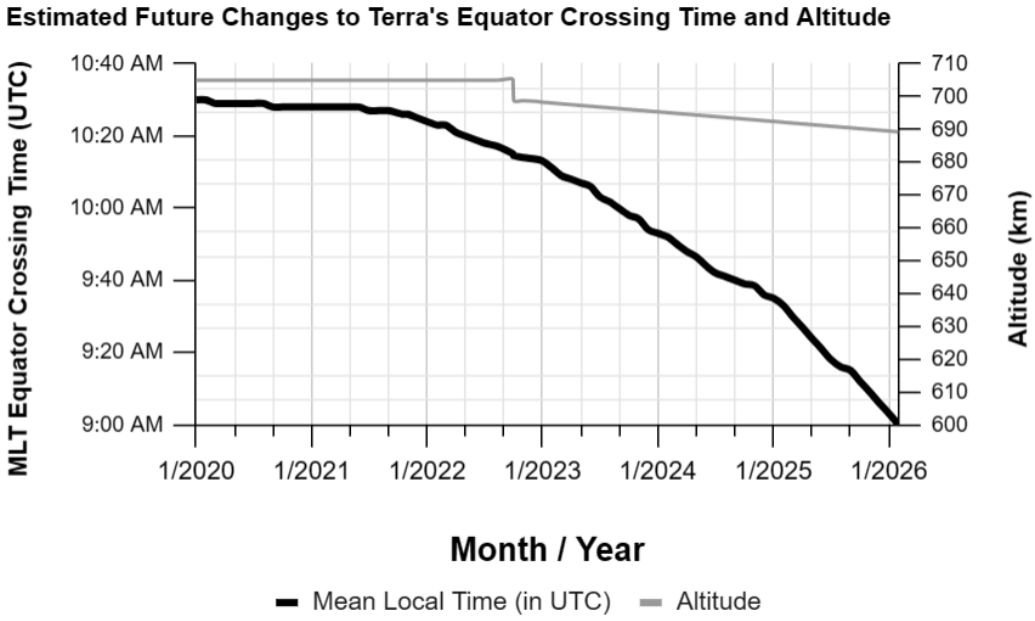

Fig. E. Terra's equator crossing time (in UTC Mean Local Time) (https://terra.nasa.gov/).

(5) Comment 5

Line 18: remove 'i.e.,'

Response: Thanks very much for your comment. We have corrected it in the revised manuscript (Page 1, Line 18).

(6) Comment 6

Line 25: typo 'and/AATSR'

Response:

Thanks very much for your comment. We have corrected it in the revised manuscript (Page 1, Line 25).

(7) Comment 7

Line 38: suggest replacing the citation 'Jiang and Liu 2014' to an actual land surface modeling study that used LST products.

Response: Thanks very much for your comment. We have corrected it in the revised manuscript (Page 2, Lines 38-39).

(8) Comment 8

Line 105: The introduction section should clarify this work is the succession of Zhang et al 2021 and basic introduction of the original RTM paper should be given.

Response: Thanks very much for your comment. The RTM method merged reanalysis data and TIR data to produce all-weather LST data, corresponding with a temporal component decomposition process. We have added relevant descriptions in the revised manuscript.

"*The theoretical foundation of the RTM method lies in the temporal component decomposition model of LST (Zhan et al., 2014; Zhang et al., 2019b). Upon comparing with independent TIR LST and validating in-situ LST, significant agreement between RTM LST and TIR LST was observed, demonstrating the effectiveness of the RTM method in all weather conditions. RTM method fully utilizes reanalysis data and TIR data to produce prospective, high-resolution, and reliable LST records on regional, continental, and global scales for the long term.*" (Lines 93–97)

"*In this study, we proposed the enhanced RTM (E-RTM) method to produce a daily (four observations per day) 1-km all-weather LST dataset for the Chinese landmass and its surrounding areas (19°N–55°N, 72°E–135°E), which was named as the Thermal and Reanalysis Integrating Moderate-resolution Spatial- seamless LST (TRIMS LST), a successor to the work of Zhang et al. (2021).*" (Lines 103–110)

(9) Comment 9

Line 123: what is daily synthesis variable.

Response: Thanks very much for your comment. The daily synthesis variable means daily Maximum Value Composite Synthesis of NDVI. We have improved this description in the revised manuscript.

"(2) *global 1-km daily Maximum Value Composite Synthesis of 'Satellite Pour l'Observation de la Terre' (SPOT) VEGETATION (VGT) Images (VGT-S1) (January 2000 to February 2000) (https://spot-vegetation.com/en) (Toté et al., 2017)*" (Lines 122-124)

More details on the relevant dataset can be found from the data product's websites (https://services.terrascope.be/collectioncatalogue/srv/eng/catalog.search#/metadata/urn:ogc:def:EOP:VITO:VGT_S1)

'*GT-S1 products (daily synthesis) are composed of the 'Best available' ground reflectance measurements of all segments received during one day for the entire surface of the Earth. This is done for each of the images covering the same geographical area. The areas distant from the equator have more overlapping parts so the choice for the best pixel will be out of more data. These products provide data from all spectral bands, the NDVI and auxiliary data on image acquisition parameters*.'

'***VGT-S1 products are MVC or Maximum Value Composite Syntheses.*** *The pixels selected for the syntheses are based on the selection of the maximum NDVI value, to ensure coverage of all landmasses worldwide with a minimum effect of cloud cover. The pixel brightness count is the ground area's reflectance (corrected for atmospheric effects); pixels in the sea area are set to 0. A map of computed normalized difference vegetation index values (NDVI image plane) is also supplied with the product. The products provide data from all spectral bands (SWIR, NIR, RED, BLUE), the NDVI, auxiliary data on image acquisition parameters and spectral band quality information*.'

(https://docs.terrascope.be/#/DataProducts/SPOT-VGT/Level3/Level3):

'*The SPOT-VGT Level 3 data (VGT-S products) are available as 1-day (S1) and 10-day (S10) TOC reflectance and NDVI syntheses. VGT S1 products are composed of the 'best available' surface reflectance measurements of all segments received during one day over nearly the entire Earth's surface. This is done for each image covering the same*

geographical area. High-latitude areas are more frequently observed and thus have more overlapping parts, so for these areas, the best observation is selected from multiple observations. The S1 product provides data from all spectral bands (B0, B2, B3, and SWIR), the NDVI, and auxiliary data on image acquisition parameters. VGT-S1 products are MVC or Maximum Value Composite Syntheses. The pixels selected for the syntheses are based on the selection of the maximum NDVI value, to ensure coverage of all landmasses worldwide with a minimum effect of cloud cover. The pixel brightness count is the ground area's reflectance (corrected for atmospheric effects); pixels in the sea area are set to 0. A map of computed normalized difference vegetation index values (NDVI image plane) is also supplied with the product. The products provide data from all spectral bands (SWIR, NIR, RED, BLUE), the NDVI, auxiliary data on image acquisition parameters, and spectral band quality information.'

(10) Comment 10

Line 124: why GLASS albedo is only used for prediction period, if so, the difference of albedo products for the training and prediction period will affect the results as there are clear bias in different albedo products.

Response: Thanks very much for your comment. We used GLASS albedo data as a substitution in the prediction step to fill the temporal gaps in MODIS LST because these temporal gaps are also effective in other MODIS products (including the MODIS albedo). Specifically, the GLASS albedo data is the best substitution for MODIS albedo data that we can find since they are strongly correlated with each other and have close accuracies according to existing studies.

We have added corresponding descriptions in the revised manuscript to clarify it.

"*We used GLASS albedo data as a substitution in the prediction step to fill the temporal gaps in MODIS LST because these temporal gaps are also effective in other MODIS products (including the MODIS albedo). Specifically, the GLASS albedo data is the best substitution for MODIS albedo data that we can find since they are strongly*

correlated with each other and have close accuracies according to existing studies (He et al., 2014; Wang et al., 2014; Chen et al., 2017; Lu et al., 2021)." (Lines 224–228)

More details and explanations are shown as follows.

(1) We acknowledge that the GLASS albedo and MODIS albedo datasets are different in retrieval algorithms and sensors onboard satellites (Qu et al., 2014). However, the GLASS albedo product was produced since 1981 from AVHRR and MODIS data, which is consistent with MODIS albedo data in spatotemporal resolutions and spatial coverage (Liang et al., 2013). Global albedo maps derived from AVHRR and MODIS are available at a resolution of 0.05° (~5 km) every 8 days. Meanwhile, albedo data at 1 km resolution in sinusoidal projection derived from MODIS observations are provided at the same temporal resolution.

(2) The GLASS albedo product showed comparable accuracy with that of MODIS albedo products when evaluated using ground measurements and MODIS albedo products (He et al., 2013; Liu et al., 2013b; He et al., 2014). Specifically, GLASS albedo products display acceptable precision within the terrestrial region of Chinese landmass and the surrounding areas.

For example, Wang et al. (2014) aggregated the Landsat TM albedo to a resolution of 1 km and compared it with the GLASS albedo at 1 km. The GLASS albedo showed an absolute error of less than 0.0163, thus demonstrating its dependability for most applications. Chen et al. (2017) compared and analyzed the spatial distribution continuity and ratio of the high-quality retrievals over the Tibetan Plateau of MODIS and GLASS. They found that (i) the GLASS land surface albedo product was superior to that of the MODIS land surface albedo in terms of both spatial distribution continuity and the ratio of the high-quality retrievals; (ii) these two products have high consistency with land surface albedo ground measurements during most of the retrieval periods and rather accurately reflect the abnormal changing processes of land surface albedo; (iii) the patchy snow is an important factor affecting the accuracies of these two products when performing a comparison with the ground measurements; (iii) the algorithm of the GLASS land surface albedo has a great advantage over that of the MODIS land

surface albedo under snow conditions. Lu et al. (2021) validated and analyzed the errors of GLASS, the GlobAlbedo, the Quality Assurance for Essential Climate Variables (QA4ECV) project, the MCD43GF, and the CM SAF Albedo dataset from the AVHRR data (CLARA-SAL) against the Chinese Ecosystem Research Network (CERN) measurements at different spatiotemporal scales over China from 2005 to 2015. The results showed that LSA estimated by GLASS agrees well with the CERN measurements on a continental scale. The GLASS product was characterized by an R^2 of 0.80, an RMSE of 0.09, and an MAE of 0.06.

(3) As GLASS Albedo is solely utilized in the RFSTM method, we employed the RFSTM method to merge the GLDAS and Terra/MODIS LST from 2003 to generate 1-km TRIMS-Terra LST. Results illustrated that TRIMS-Terra LST generated by the RTM method and RFSTM method were highly correlated with each other (Figure R1) and showed close accuracies (Table R2). MBEs have a difference of less than 0.50 K, and RMSEs have a difference of less than 1.2 K. Overall, the RFSTM method is slightly less accurate than the TRIMS-Terra LST generated by the RTM method. It is to be observed that the RFSTM method was only used to generate LST for 54 days, which has a relatively smaller impact on the overall accuracy of TRIMS LST.

Table R2: MBE, and RMSE from validation results of TRIMS-Terra LST with the in-situ LST

Site	Condition	TRIMS-Terra LST (RTM)				TRIMS-Terra LST (RFSTM)			
		Daytime		Nighttime		Daytime		Nighttime	
		MBE (K)	RMSE (K)	MBE (K)	RMSE (K)	MBE (K)	RMSE (K)	MBE (K)	RMSE (K)
D105	All	1.63	3.15	-1.05	1.94	1.75	3.3	-1.55	2.20
	clear	1.78	2.17	-1.17	2.04	1.85	3.34	-2.37	2.66
	cloudy	1.54	3.25	-0.88	1.78	1.04	3.44	-0.40	1.29
GZA	All	0.93	2.61	-0.78	1.76	1.26	3.10	-1.95	2.26
	clear	0.79	2.51	-0.68	1.70	0.94	2.72	-1.26	2.10
	cloudy	1.11	3.71	-0.94	1.85	1.61	4.20	-1.47	2.35

References:

Chen, A., Meng, W., Hu, S., and Bian, A.: Comparative analysis on land surface albedo from MODIS and GLASS over the Tibetan Plateau, Trans Atmos Sci, 43, 932-942., https://doi:10.13878/j.cnki.dqkxxb.20171030001, 2020.

He, T., Liang, S., Yu, Y., Wang, D., Gao, F., and Liu, Q.: Greenland surface albedo changes in July 1981–2012 from satellite observations, Environ. Res. Lett., 8, 044043, https://doi.org/10.1088/1748-9326/8/4/044043, 2013.

He, T., Liang, S., and Song, D.-X.: Analysis of global land surface albedo climatology and spatial-temporal variation during 1981–2010 from multiple satellite products, Journal of Geophysical Research: Atmospheres, 119, 10,281-10,298, https://doi.org/10.1002/2014JD021667, 2014.

Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., Tang, H., Qu, Y., Bo, Y., Qu, Y., Ren, H., Yu, K., and Townshend, J.: A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies, International Journal of Digital Earth, 6, 5–33, https://doi.org/10.1080/17538947.2013.805262, 2013.

Liu, N. F., Liu, Q., Wang, L. Z., Liang, S. L., Wen, J. G., Qu, Y., and Liu, S. H.: A statistics-based temporal filter algorithm to map spatiotemporally continuous shortwave albedo from MODIS data, Hydrology and Earth System Sciences, 17, 2121–2129, https://doi.org/10.5194/hess-17-2121-2013, 2013.

Liu, Q., Wang, L., Qu, Y., Liu, N., Liu, S., Tang, H., and Liang, S.: Preliminary evaluation of the long-term GLASS albedo product, International Journal of Digital Earth, 6, 69–95, https://doi.org/10.1080/17538947.2013.804601, 2013.

Lu, Y., Wang, L., Hu, B., Zhang, M., Qin, W., Zhou, J., and Tao, M.: Evaluation of satellite land surface albedo products over China using ground-measurements, International Journal of Digital Earth, 14, 1493–1513, https://doi.org/10.1080/17538947.2021.1946179, 2021.

Qu, Y., Liu, Q., Liang, S., Wang, L., Liu, N., and Liu, S.: Direct-Estimation Algorithm for Mapping Daily Land-Surface Broadband Albedo From MODIS Data, IEEE Transactions on Geoscience and Remote Sensing, 52, 907–919,

https://doi.org/10.1109/TGRS.2013.2245670, 2014.

Wang, L.，Zheng, X.，Sun, L.，Liu, Q., and Liu, S.: Validation of GLASS albedo product through Landsat TM data and ground measurements, Journal of Remote Sensing, 18, 547–558, https://doi.org/10.11834/jrs.20143130, 2014.

(11) Comment 11

Section 2.2: any resources and information for accessing the ground measurement networks?

Response: Thanks very much for your comment. The links to download ground measurements have been provided in the manuscript (Page 7, Line 165–170).

(12) Comment 12

Line 211 and 212: duplicated statements, suggest double the manuscript for typos.

Response: Thanks very much for your comment. It has been corrected in the revised manuscript (Page 10, Lines 215-216).

(13) Comment 13

It is good to summarize the available all-weather LST products; however, if the authors try to include such summary, a complete literature review should be done. Some products, e.g., high resolution LST over Europe (Rains et al. 2022), 2-km hourly product over the US (Jia et al. 2022), and hourly LST over China (https://data.tpdc.ac.cn/zh-hans/data/06414391-abd4-4d28-a844-bd036a0b8c55/)

References:

Dong, Shengyue, et al. "A Data Fusion Method for Generating Hourly Seamless Land Surface Temperature from Himawari-8 AHI Data." Remote Sensing 14.20 (2022): 5170.

Jia, Aolin, Shunlin Liang, and Dongdong Wang. "Generating a 2-km, all-sky, hourly land surface temperature product from Advanced Baseline Imager data." Remote Sensing of Environment 278 (2022): 113105.

Rains, Dominik, et al. "High-resolution all-sky land surface temperature and net radiation over Europe." Earth System Science Data Discussions (2022): 1-22.

Zhang, Xiaodong, et al. "A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature." Remote Sensing of Environment 260 (2021): 112437.

Response: Thanks very much for your comment. It is not our focus to summarize current all-weather LST offerings in this article. Instead, we have provided a table in Appendix C with all the currently available all-weather/all-sky/gap-free LST products that we can find for comparison. We have also emphasized the advantages of our TRIMS LST dataset compared to existing all-weather/gap-free LST datasets in Section 4.5.

'*Recently, several all-weather LST datasets have been released by the scientific communities (see Appendix C). An all-weather LST product series with a temporal resolution from 15 min (Martins et al., 2019) to monthly (Metz et al., 2017; Zhao et al., 2020a; Hong et al., 2022; Yao et al., 2023), a spatial resolution from 1 km to 0.5°, and a spatial coverage from a specific region (Qinghai-Tibet Plateau, Asia, Europe and Africa) to globe has been preliminarily formed. All-weather LST products based on MODIS LST interpolation (Zhang et al., 2022) or fusion with other multi-source data (Xu and Cheng. 2021; Zhang and Cheng. 2020b; Zhang et al. 2020c; Zhang et al., 2021; Yu et al. 2022) dominate the field. TRIMS LST similarly belongs to this group. Overall, the uniqueness or advantages of TRIMS LST are in three main areas*:'(Lines 592–598)

'*First, the TRIMS LST demonstrates comparable or better accuracy than existing publicly released all-weather/spatially seamless LST datasets. A thorough comparison with satellite TIR LST products has indicated the effectiveness of TRIMS LST, with MBD ranging from -1.5 K to 1 K and STD ranging from 1 K to 3 K, thus confirming its accuracy and consistency (Fig. 8, Fig. 9, and Fig. 10). Furthermore, in-situ LST evaluations show MBE ranging from -1.64 K to 2.88 K and RMSE ranging from 1.82 K to 3.48 K (Table II and Table III). Interestingly, no significant difference is observed between clear-sky and cloudy conditions, indicating the robustness of TRIMS LST across various situations. Furthermore, the RTM technique was utilized at four top-*

quality stations and the nearby region (11x11 km): Evora, Gobabeb, KIT-Forest, and Lake Constance (Meng et al., 2023). The TRIMS LST has performed favorably in validating results across different land cover types, including barren land, savannas, and forests, with an RMSE range of 1.90 K to 3.10 K. Additionally, over water site, TRIMS LST has an RMSE of 1.60 K. Thus, based on the results of this study, TRIMS LST can be considered a reliable source of LST.'(Line 599–608).

'*Second, the method employed in this study effectively overcomes the issue of boundary effects in reconstructing the all-weather process due to the large differences in spatial resolution between different data sources (Zhang et al., 2021; Quan et al., 2023). This is achieved through the utilization of the E-RTM method, which is based on a temporal decomposition model of LST. With this model, the LFC and HFC components can be directly determined from high-resolution MODIS and ancillary remote sensing data (Eq. 1). Consequently, only spatial downscaling of HFCcld is required, eliminating the need for direct downscaling of the GLDAS LST. This method reduces the possibility of insufficient spatial downscaling. Additionally, the E-RTM method considers the relationship between LSTs of neighboring pixels, resulting in decreased errors during spatial downscaling (Fig. 4)*.'(Line 609–616)

'*Third, TRIMS LST offers advantages in effectively recovering LST information and preserving temporal integrity under cloudy conditions. With a spatial resolution of 1 km, TRIMS LST covers both daytime and nighttime LST from 2000 to 2022, which is comparable in spatio-temporal resolution to other published seamless LST datasets (Appendix C). The TRIMS LST dataset will be made publicly available on an annual basis, contingent on availability of pertinent input data for the model. The E-RTM method effectively recovers temperature information under clouds, ensuring clear physical meaning and high accuracy and image quality of TRIMS LST. Moreover, TRIMS LST extends the all-weather LST coverage of the MODIS temporal gap. This enhances the completeness of long-time series LST datasets, creating a unique and valuable collection*.'(Lines 617–623)