

Response to Reviewer #3

We appreciate a lot for your efforts in providing detailed comments and recommendation. They are very helpful to improve the quality of the manuscript. We have revised the manuscript according to your comments. The comments from the reviewers are kept in regular font, our responses use blue highlighting, and the revised sentences or words in the revised manuscript are highlighted with red color.

Q1. Why was the MERRA-2 atmospheric dataset selected considering its coarse spatial resolution?

Response: Thank you for your comments. We would like to make some explanations on the selection of MERRA-2 atmospheric dataset as follows:

(1) To the best of our knowledge, there are only two global reanalysis datasets, the fifth-generation European Center for Medium-Range Weather Forecasts atmospheric reanalysis dataset (ERA5) and the Second Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) dataset, which could provide hourly atmospheric water vapor content (WVC) data from 1981 to 2005. The ERA5 and MERRA-2 provide hourly WVC at $0.25^{\circ} \times 0.25^{\circ}$ and $0.625^{\circ} \times 0.5^{\circ}$ spatial resolution, respectively. Huang et. al. (2021) systematically assessed the hourly WVC of ERA5 and MERRA-2 by a comparison with 33 Global Navigation Satellite System stations from 2017 to 2018. The results of the comparison are as follows: ① The accuracies of the ERA5- and MERRA-2-derived WVC are very high; ② The performance of ERA5 is slightly better than MERRA-2 due to its higher spatial resolution; ③ With the same grid spatial resolution, the mean root mean square difference between two reanalysis data sets is 0.1mm (0.01g/cm²).

(2) To match the GT-LST pixels, these datasets all need to be resampled to 0.05° spatial resolution. However, selecting ERA5 will consume twice as much local storage space and memory as MERRA-2.

However, the goal of this study is to develop a global historical twice-daily LST product from 1981 to 2005, where WVC is only one of intermediate variables that

obtain the nonlinear generalized split-window (GSW) algorithm coefficients corresponding to the subrange of WVC. Considering the tradeoff between accuracy, local storage space and memory and computational burden, we choose the MERRA-2-derived WVC to estimate LST.

Q2. In my opinion, the SURFRAD measurements are not the best option for LST validation, especially in the case of evaluating medium/coarse spatial resolution LST, considering the substantial spatial heterogeneity of the sites. Moreover, the measured longwave radiations by pyrgeometers are different from the directional radiance collected by satellites, which has been reported in different studies.

Response: Thanks a lot for your comments. We would like to make some explanations as follows:

(1) In order to maximize the usefulness of GT-LST for research it is necessary to assess the accuracy of GT-LST using as many methods as possible. Ground-based validation is the most traditional and well-used method, and it provides suitable validation results for well-defined and dedicated sites in most cases. However, there are the limited number of high-quality sites (i.e., KIT stations and NASA JPL stations) around the world that are dedicated to LST validation due to their high cost and logistical barriers (Guillevic et al., 2014). Moreover, they could only provide measurements after 2009 (Guillevic, et al., 2018). Although the SURFRAD network was not initially designed for LST validation, SURFRAD can provide high-quality radiance measurements from 1995 to present, which are useful for validating satellite LST products. SURFRAD measurements have already been used for evaluating ASTER, GOES, MODIS, VIIRS, AVHRR and AMSR-E LST products (Wang and Liang, 2009; Yu et al., 2012; Guillevic et al., 2012; Liu et al., 2019; Jiménez et al., 2017). All SURFRAD stations are selected in this study. Fig. R1 shows the surroundings of the sites on the AVHRR scale, where all of these sites except DRA are located in large flat agricultural areas covered by crops and grass. Due to not all stations are representative of spatially homogeneous areas at GT-LST scales, we used as much in situ measurements as possible in order to characterize them correctly. Considering

the limitation of the ground-based validation method, we compared GT-LST with a large number of well documented and validated LST products derived from satellites to characterize GT-LST performance.

(2) Indeed, as you mentioned, longwave radiations from satellite and ground measurements may differ according to their measurement methods. AVHRR measures directional measurements in the atmospheric window, while the measured longwave radiations by pyrgeometers are hemispheric, wider spectrum derivations. If the surface is black body, the LSTs derived from these two measurements are the same (Wang et al., 2005; Li et al., 2023). However, most natural objects are not black bodies. AVHRR view zenith angles were considered to be an important factor influencing the results when comparisons were made with in situ measurements. Fig. R2 shows the relationship between view zenith angles and bias in instantaneous LST at station pixels. Our result shows that a high view angle does not necessarily bring a high bias or a low view angle does not necessarily always bring a low bias. This means that view angle should not be a significant source for the bias of GT-LST at a 0.05° pixel size.

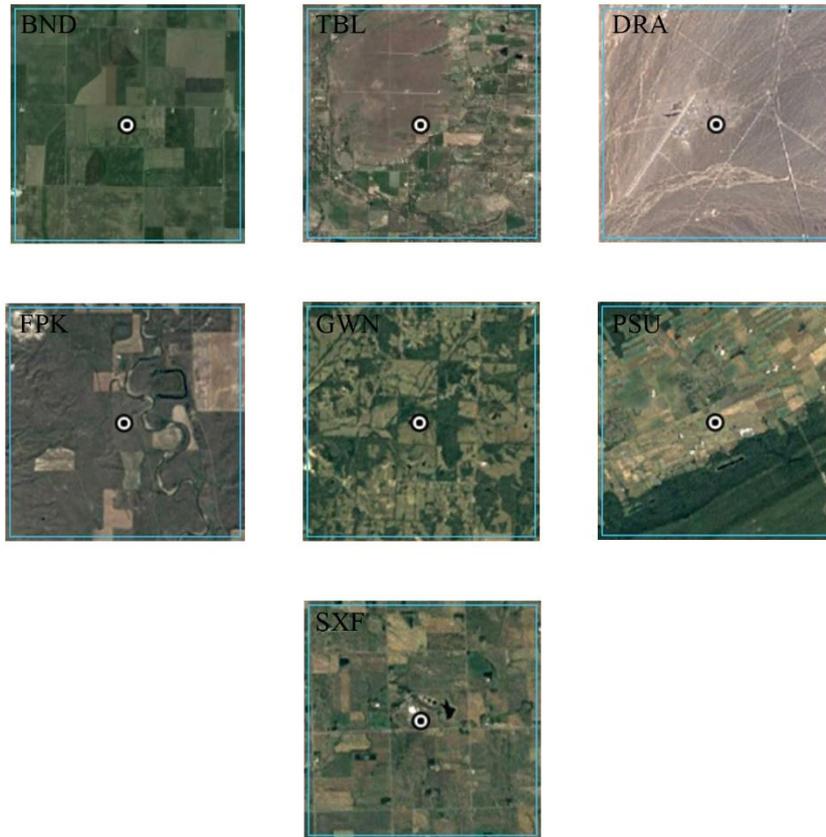


Figure R1: Aerial photos of the SURFRAD sites. The black dot marks the position of the site, and the blue square marks the size of an AVHRR pixel

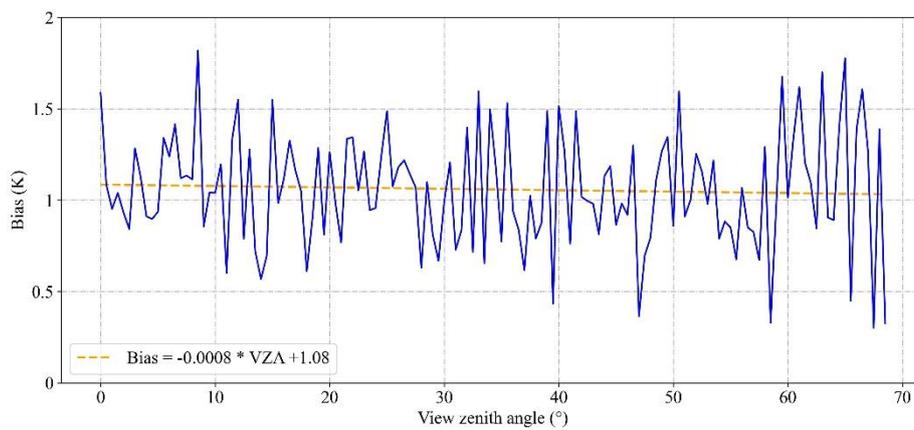


Figure R2: The relationship between bias of land surface temperature (GT-LST minus in situ observations) and view zenith angle.

Q3. Why was the MYD11 LST product selected instead of the MYD21 LST, or geostationary LST products that have closer spatial resolutions to the GT-LST product? In the inter-comparison, the MDY11A1 LST was spatially aggregated to the spatial resolution of the GT-LST product with a simpler arithmetic mean. I doubt the validity of the MYD11 LST after the simple aggregation.

Response: Thank you for your comments. The reason that we selected MYD11A1 LST product instead of the MYD21A1 is that the MYD11A1 LST products have been well validated by using the temperature-based method and radiance-based method methods with an accuracy of approximately 1 K (Wan, 2014; Li et al., 2023). To reduce the discrepancies induced by viewing location, time, geometry and quality control (QC), we used five criteria to guarantee the reliability of the intercomparison results. As for the criterion of collocation in space, besides that MYD11A1 LST was aggregated to the spatial resolution of the GT-LST product by averaging all MYD11A1 pixels, it also requires all MYD11A1 pixels with QC = 0 (i.e., the highest quality) within a coarse spatial resolution pixel ($0.05^{\circ} \times 0.05^{\circ}$).

In addition, according to your suggestion, we have compared GT-LST with MYD21A1 LST (Aqua/MODIS LST product using the TES algorithm, Collection 6.1). Spatially, this intercomparison was conducted at the global scale. Temporally, it was performed on 4 months in 2004 (January, April, July, and October) which cover different seasons. The result of the inter-comparison, in Fig. 9, is as follows: The daytime and nighttime RMSD values of 3.2 K and 2.5 K and that of bias of 0.1 K and 1.3 K. Compared to the result of MYD11A1, the significantly smaller bias was obtained for MYD21A1. The possible reason is attributed to the fact that the MYD21A1 LST uses the same observations with MYD11A1 but uses TES method to dynamically retrieve LSE. The following contents have been added in Line 184-187 and Line 405-412, respectively.

“In this study, Collection-6.1 MYD11A1 of 2004 was selected for sensor-to-sensor comparison. MYD21A1 LST product, which uses the same observations with MYD11A1

but uses temperature–emissivity separation method to dynamically retrieve LST and emissivity, was also selected to make an intercomparison with GT-LST in this study. This inter-comparison was conducted on 4 months in 2004 (January, April, July, and October) which cover different seasons.”

“As a result, the *dynamic* emissivity of GT-LST is typically lower than that of MYD11A1, which leads to overestimation of the LST (Hulley et al., 2016; Guillevic et al., 2014; Reiners et al., 2021; Ren et al., 2011). To further demonstrate this point, we compared GT-LST with MYD21A1 LST. Fig. 9 shows the daytime and nighttime RMSD values of 3.2 K and 2.5 K and that of bias of 0.1 K and 1.3 K between GT-LST and MYD21A1 LST for 4 months in 2004. Compared to the result of MYD11A1, the significantly smaller bias was obtained for MYD21A1. The possible reason is attributed to the fact that the MYD21A1 LST uses the same observations with MYD11A1 but uses a physics-based method to dynamically retrieve emissivity.”

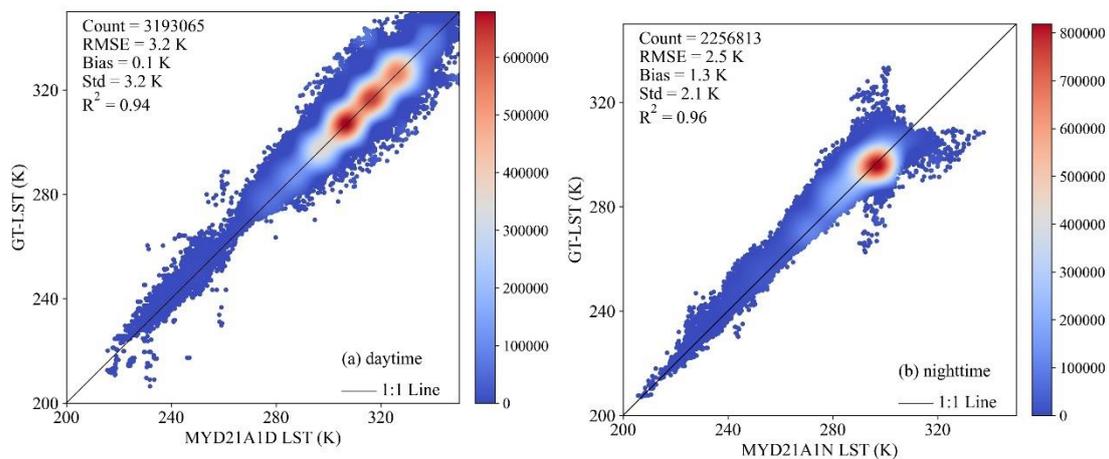


Figure 9: Intercomparison of GT-LST and MYD21A1 LST in January, April, July, and October 2004: (a) daytime; (b) nighttime.

Q4. Typo in Eq. 5. It should be ‘AVH’ rather than ‘AST’ on the left side of the equation.

Response: Thank you for your careful reading. We have corrected this issue.

Q5. The description of the emissivity retrieval process is unclear. How were the soil

type data used?

Response: Sorry for the unclear expression. To make it clear, the paragraph in Line 254-258 has been revised to as follows:

“...where $\varepsilon_{i,s}^{AST}$ is the bare soil emissivity in ASTER channel i ($i=10, \dots, 14$), and $\varepsilon_{i,v}^{AST}$ is the emissivity of dense vegetation in ASTER channel i . Because the emissivity spectra of dense vegetation are similar and vary slightly in the TIR region, we used the dense vegetation emissivity of ASTER channel i provided by Meng et al. (2016). ε_i^{AST} is the emissivity of the ASTER GED product in channel i . P_v is calculated from the NDVI of the ASTER GED product according to Eq. (3). For long-term cloud cover pixels and dense vegetation pixels ($P_v = 1$), the bare soil emissivity of these ASTER pixels are null values. To generate a global gap-free bare soil emissivity map of ASTER, we used the average emissivity of the same soil type within 5×5 neighborhood pixels to fill these null values. Because of some pixels with no valid neighbor pixel for averaging we needed to enlarge the neighborhood until all null values are filled. Soil-type data are described in Section 2.3.”

Q6. Why were the RMSEs over savannas and croplands the largest amongst different land surface types?

Response: Fig. R3 shows the intercomparison results between GT-LST and MYD11A1 LST over savannas (i.e., woody savannas and savannas) and croplands (i.e., cropland/natural vegetation mosaics and croplands). We would like to make some explanations on relatively large disparities over savannas and croplands between these two products as follows:

According to NDVI threshold method, the daily emissivity of an AVHRR pixel can be derived using the following formula:

$$\varepsilon = \varepsilon_{veg} * FVC + \varepsilon_{soil} * (1 - FVC) \quad (R1)$$

Here, ε is the emissivity, ε_{veg} is the vegetation emissivity, ε_{soil} is the bare soil emissivity, and FVC is the fraction of vegetation cover.

For a vegetation pixel, its FVC varies greatly due to the influence of natural and human factors, which leads to the underestimation of emissivity comparing with fixed emissivity, resulting in an overestimation of LST. The situation is particularly evident

over croplands and savannas. Specially, natural disasters (e.g., drought and pests) and agricultural activities (e.g., harvest, cropland lies fallow) can significantly decrease cropland density and result in higher exposure of the soil. It leads to a decrease in cropland emissivity, resulting in an overestimation of LST. The emissivity for savannas decreases because of the increasing proportion of soil by grazing, fire and annually a long period in which moisture inadequate, resulting in an overestimation of LST. We have added the following descriptions in Line 391-393:

“Savannas and croplands, including woody savannas and savannas, croplands and cropland/natural vegetation mosaics, respectively, had the largest RMSD. The possible reason is that the fraction of vegetation cover of savannas and croplands vary greatly due to the influence of natural and human factors, which leads to the underestimation of emissivity comparing with fixed emissivity of MYD11A1, resulting in an overestimation of LST. Snow and ice and water bodies had the smallest RMSD.”

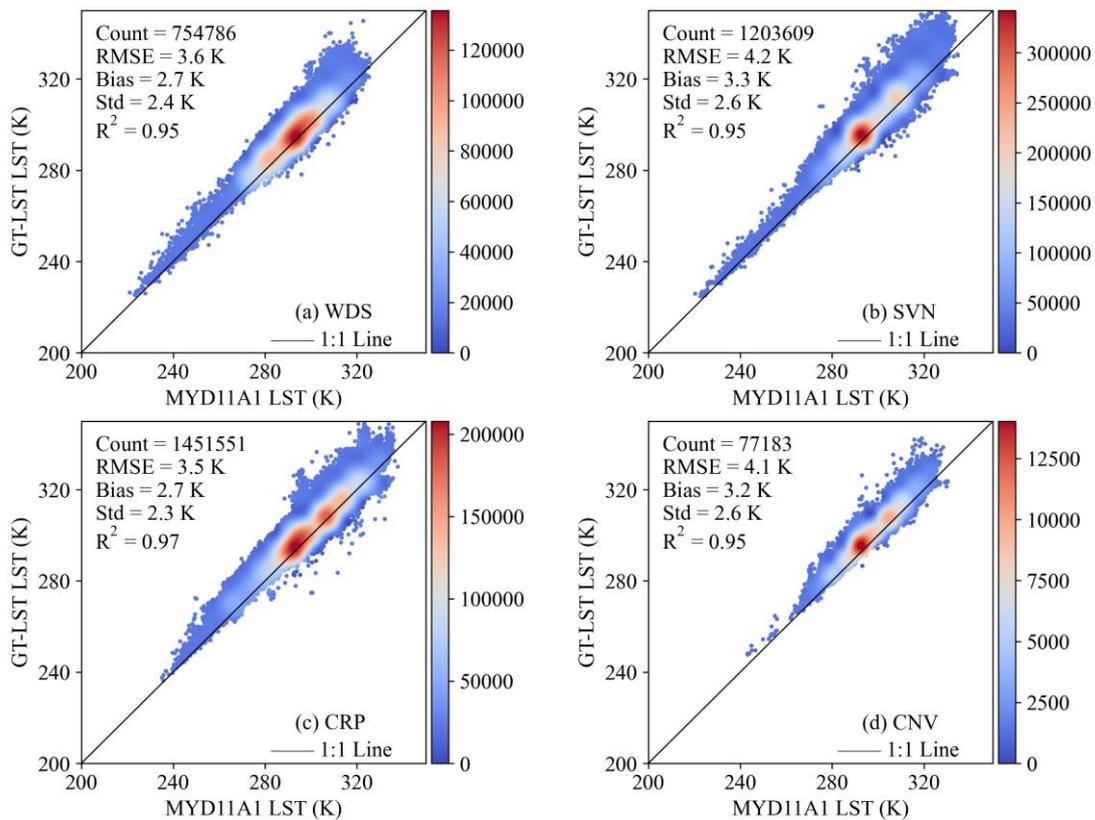


Figure R3. Scatterplots of GT-LST versus MYD11A1 LST during 2004 over WDS (a), SVN (b), CRP (c), and CNV (d). WDS: woody savannas, SVN: savannas, CRP:

croplands, and CNV: cropland/natural vegetation mosaics.

Q7. Any explanations for the higher uncertainties in spring and summer?

Response: Thanks a lot for your comments. A reasonable explanation could be that differences seasons associated with different atmospheric conditions: cool and dry in autumn and winter, hot and wet in spring and summer. Generally, large differences between different LST products were found under high temperatures and high atmospheric water vapor content conditions. Some literatures got similar result. For example, Reiners et al. (2021) compared AVHRR LST product of the TIMELINE project with MYD11_L2 LST product from 2003 to 2014. The result shows the seasonal pattern, with lower accordance and higher bias in summer and higher accordance and lower bias in winter.

Q8. Line 370, why is the ASTER GED-based emissivity retrieval used in the GT-LST product lower than the classification-based emissivity used in the MYD11 product?

Response: Thanks a lot for your valuable comments. We would like to make some explanations on the lower emissivity of GT-LST product as follows:

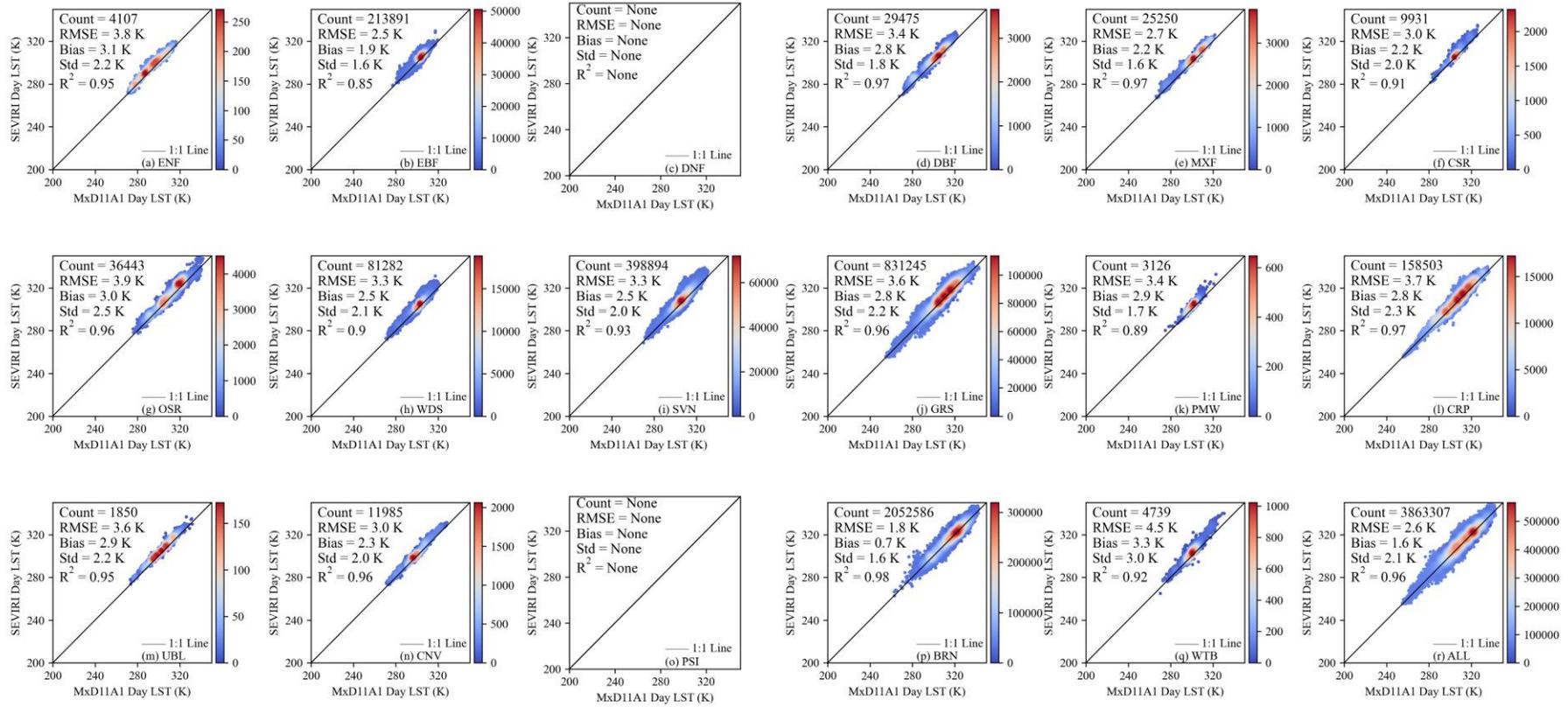
(1) The MYD11A1 LST product is generated by the split-window (SW) algorithm. Land surface emissivities of this product are assigned according to classification-based method that produces emissivities with fixed values for a limited number of land cover types. This method works well over densely vegetated areas and water where emissivities are relatively stable. However, cold biases of 3-5 K are often found over semi-arid and arid regions because these regions have much higher emissivity variability, and only one fixed overestimated emissivity inferred from land cover types is assigned to these regions (Coll et al., 2009; Hulley and Hook 2009; Wan et al., 2002). In order to represent the natural variation in emissivity, we used an improved NDVI threshold method to dynamically retrieve daily emissivity. Based on the analysis above, emissivity derived from dynamic methods is lower than emissivity according to classification-based method, which makes the proposed GT-LST is higher than

MYD11A1 LST (i.e., positive bias). We note that earlier researches on this issue had similar results. Reiners et al. (2021) compared AVHRR LST product of the TIMELINE project with MYD11_L2 LST product from 2003 to 2014, the result shows that the TIMELINE dynamic emissivity is lower than the MYD11_L2 fixed emissivity and a general positive bias (i.e., bias=2.2 K) of TIMELINE LST towards MYD11_L2 LST. Martins et al. (2019) compared MSG LST and GOES-16 LST and revealed that a positive bias (MSG > GOES) of around 1.6 K persists due to the overestimation of the fixed emissivity of GOES. Mao et al. (2007) analyzed the retrieval result by radiative transfer model with neural network algorithm and MODIS product algorithm, indicating that MOD11_L2 LST product overestimates the emissivity, resulting in an underestimation of LST.

(2) To further illustrate the MYD11A1 LST product overestimates the emissivity, we present an intercomparison exercise between MxD11A1 LST products (Terra and Aqua/MODIS using SW algorithm, Collection 6) with fixed emissivity and MSG LST products (MSG/SEVIRI using SW algorithm) with dynamic emissivity for 4 days (January 15, April, 15, July 15, and October 15, 2020). The criteria in Sec 3.2 were used to guarantee the reliability of the intercomparison results. The result is shown in Fig. R4, indicating that a general positive bias (daytime ranges from 0.7 K to 3.3 K, nighttime ranges from 0.2 K to 1.4 K) of MSG LST towards MxD11A1 LST for each land cover types.

Based on the analysis above, it is reasonable to conclude that the GT-LST dynamic emissivity is lower than the MYD11A1 classification-based emissivity.

a



b

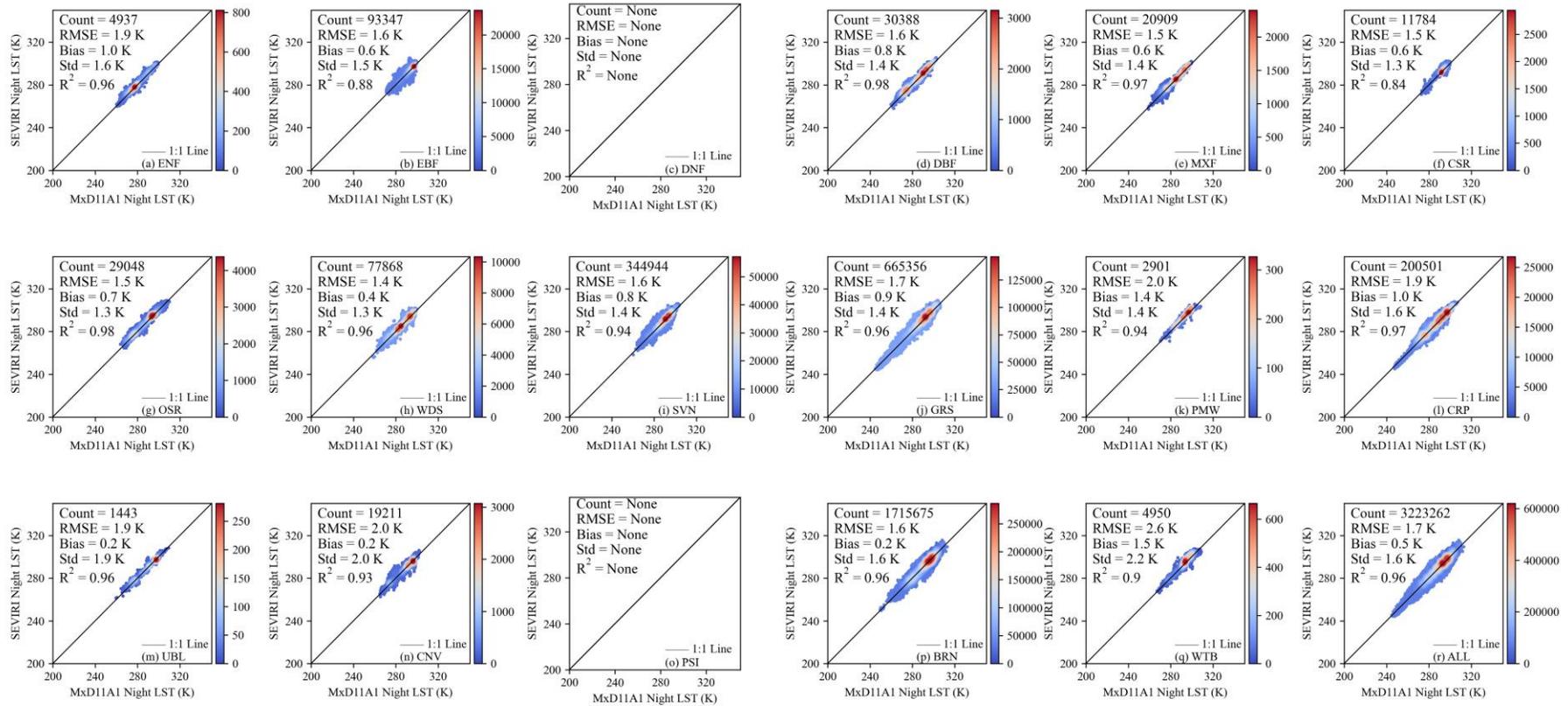


Figure R4: The bias (SEVIRI LST minus MxD11A1 LST) and RMSD between the MxD11A1 product and SEVIRI during daytime (a) and nighttime (b) over various land cover types.

Q9. Line 400, why is the emissivity of GT-LST lower than that of RT-LST? The analysis is too simple to understand the intercomparison between different LST products. More in-depth investigations are needed for the comparison with the existing AVHRR LST data.

Response: RT-LST product provided only a rough estimation of emissivity using a land cover classification map, the FAO soil map of Africa and additional maps of tree, herbaceous, and bare soil percent cover. All of these data sets are static and therefore authors do not account for local phenological, environmental changes, and human factors in time. Based on the analysis above, emissivity derived from dynamic methods is lower than emissivity according to static method.

Q10. There are quite some minor grammatical errors, e.g., Line 411, ‘an improved method that consider annual changes’. Please check them carefully.

Response: Thank you for your careful reading. Following your suggestion, we have checked the whole manuscript and corrected these issues.

Q11. The authors mentioned the increased uncertainties of AVHRR LST with time due to the orbital drift. It would be useful to add some analyses of the variation in the accuracy of the GT-LST product in the time series.

Response: Thanks a lot for your comments. Indeed, one of the intentions of GT-LST is providing effective supplementary data for global long-term time series analysis. The analysis requires daily, monthly or annual mean LST (i.e., DMLST, MMLST, and AMLST) more than instantaneous LST as these mean LSTs are key indicators when monitoring global LSTs over a long time series (Li et al., 2023; Liu et al., 2023; Xing et al., 2021). It is possible to derive an estimate of the global accurate DMLST, MMLST and AMLST based on twice-daily LST product. However, impacting of the NOAA satellite orbital drift, daytime and nighttime observations of NOAA afternoon satellites cannot represent maximum and minimum temperatures well. Therefore, calculating the daily and monthly mean LST by averaging daytime and nighttime LSTs derived from GT-LST has a significantly lower accuracy than other studies (Fig. A4). Inspired by the

work of Xing et al. (2021), we use simple linear combinations of daytime and nighttime LST values that were observed at observation times for NOAA to estimate DMLST and MMLST. In order to validate the accuracy of DMLST and MMLST according to the simple linear regression method, we compared DMLST and MMLST derived from GT-LST with that of in situ LST observations from SURFRAD sites, and reported RMSE values of approximately 2.4 K and 2.7 K, respectively. These results are similar to that of Xing et al. (2021) and Chen et al. (2017). In this way, we still obtain accurate DMLST and MMLST without satellite orbit drift correction. Then, we rephrase the paragraph in Line 429-436 as follows:

*“...To estimate MMLST, first obtain the mean instantaneous clear-sky LST at daytime and nighttime, and then use these mean values to estimate MMLST according to the **simple linear regression method (see Appendix B)**. In order to validate the accuracy of MMLST results, we compared MMLST based on GT-LST with that of in situ LST observations from SURFRAD sites for **1994–2005**. All in situ LST measurements are all-sky and complete on a certain month, which means that the in situ MMLST is true MMLST. Fig. 15 showed that MMLST derived from GT-LST are related to the true MMLST, with an R^2 value of **0.94** and an RMSE value of **2.7 K**. This result is **similar to that of Chen et al. (2017)**, who compared MMLST from MODIS day and night instantaneous clear-sky LST with actual MMLST from 156 flux tower stations, and reported RMSE bias values of approximately 2.7 K.”*

We have redrawn Fig. 13 according to the simple linear regression method. For your convenience, we listed it below.

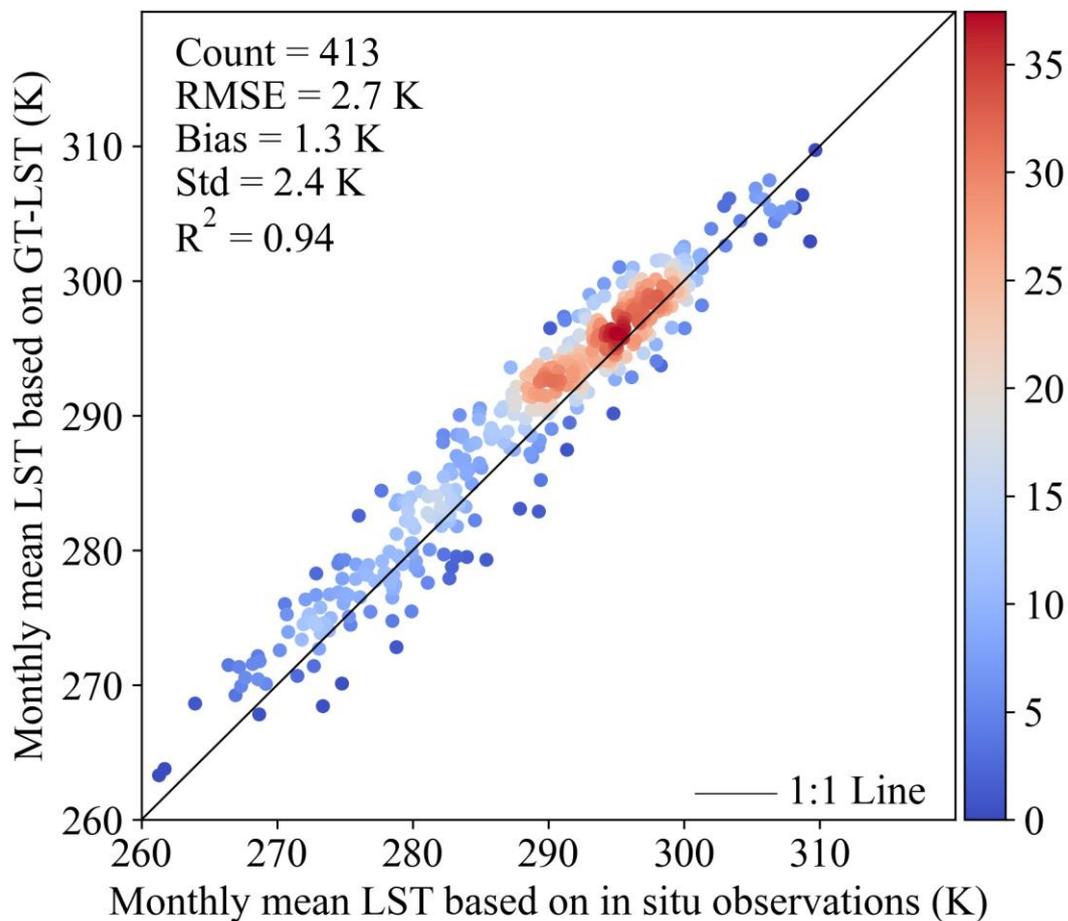


Figure 13: Monthly mean LST based on GT-LST versus monthly mean LST based on in situ LST from 1994 to 2005.

In addition, as for some details of the simple linear regression method, we have added the following descriptions in Appendix B.

“Impacting of the NOAA satellite orbital drift, daytime and nighttime observations of NOAA afternoon satellites cannot represent maximum and minimum temperatures well. Therefore, the MMLST according to the simple average method has a significantly lower accuracy than other studies (Fig. A4). Xing et al. (2021) proposed to use 9 combinations of two to four MODIS instantaneous retrievals of which at least one daytime LST and one nighttime LST to estimate mean LSTs, and determined the weight for every moment. Inspired by the work of Xing et al. (2021), we determined to use simple linear combinations of monthly mean daytime and nighttime LST values that

were observed at observation times for NOAA to estimate MMLST with ground-based measurement. For the combinations of two valid monthly mean LSTs (one daytime and one nighttime LST), the regression models can be written as follows:

$$MMLST = a_1 * MMLST_{day} + a_2 * MMLST_{night} + b \quad (B1)$$

where MMLST is the ground-based monthly mean LST, a_1 , a_2 and b are the fitting coefficients, $MMLST_{day}$ is the monthly mean in situ LST at the NOAA daytime observation, $MMLST_{night}$ is the monthly mean in situ LST at the NOAA nighttime observation.

Taking into account the observed times of NOAA satellites with orbital drift effect since 1981, combinations of two observations from these satellites contain eight cases: 13:30–17:00/01:30–05:00 local solar time in 0.5-hour interval. Based on the in situ LST measurements during the period 2003 to 2018 at 227 flux stations operating in globally diverse regions, we obtained the fitting coefficients (Table A1). Then, we calculated the MMLST of GT-LST using GT-LST monthly mean daytime and nighttime LSTs, Eq. (B1), and the fitting coefficients listed in Table A1.”

Table A1. Statistics for the relationship between the regressions of the eight combinations and actual monthly mean LST.

Case	Time	a_1	a_2	b	RMSE	R^2	Number
1	13:30/01:30	0.3844	0.5783	10.3446	2.0	0.97	12095
2	14:00/02:00	0.4010	0.5621	10.2042	1.9	0.98	12241
3	14:30/02:30	0.4235	0.5451	8.6172	1.9	0.98	12381
4	15:00/03:00	0.4490	0.5211	8.2652	1.8	0.98	12303
5	15:30/03:30	0.4816	0.4840	9.5710	1.8	0.98	12165
6	16:00/04:00	0.5250	0.4349	11.2284	2.0	0.97	11818
7	16:30/04:30	0.5663	0.3884	12.8572	2.2	0.96	10992
8	17:00/05:00	0.6040	0.3621	9.7302	2.4	0.96	9765

References for the above responses are listed below:

Chen, X., Su, Z., Ma, Y., Cleverly, J., Liddell, M.: An accurate estimate of monthly mean land surface temperatures from MODIS clear-sky retrievals, *J. Hydrometeorol.*, 18, 2827-2847, <https://doi.org/10.1175/JHM-D-17-0009.1>, 2017.

Coll, C., Wan, Z., Galve, J. M.: Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, *J Geophys. Res-Atmos.*, 114, 1-15, <https://doi.org/10.1029/2009jd012038>, 2009.

Guillevic, P. C., Biard, J. C., Hulley, G. C., Privette, J. L., Hook, S. J., Olioso, A., Göttsche F. M., Radocinski, R., Román, M. O., Yu, Y., and Csiszar, I.: Validation of Land Surface Temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements, *Remote Sens. Environ.*, 154, 19-37, <https://doi.org/10.1016/j.rse.2014.08.013>, 2014.

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J. A., Remedios, J., Román, M., and Camacho, F.: Land surface temperature product validation best practice protocol, https://lpvs.gsfc.nasa.gov/PDF/CEOS_LST_PROTOCOL_Feb2018_v1.1.0_light.pdf, 2018.

Guillevic, P. C., Privette, J. L., Coudert, B., Palecki, M. A., Demarty, J., Ottlé, C., and Augustine, J. A.: Land Surface Temperature product validation using NOAA's surface climate observation networks—Scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS), *Remote Sens. Environ.*, 124, 282-298, <https://doi.org/10.1016/j.rse.2012.05.004>, 2012.

Huang, L., Mo, Z., Liu, L., Zeng, Z., Chen, J., Xiong, S., and He, H.: Evaluation of hourly PWV products derived from ERA5 and MERRA-2 over the Tibetan Plateau using ground-based GNSS observations by two enhanced models, *Earth Space Sci.*, 8, e2020EA001516, <https://doi.org/10.1029/2020EA001516>, 2021.

Hulley, G. C., Hook, S. J.: The North American ASTER land surface emissivity database (NAALSED) version 2.0, *Remote Sens. Environ.*, 113, 1967-1975, <https://doi.org/10.1016/j.rse.2009.05.005>, 2009.

Li, Z.-L., Wu, H., Duan, S.-B., Zhao, W., Ren, H., Liu, X., Leng, P., Tang R., Ye, X., Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B.-H., Yan, G., and Zhou, C.: Satellite remote sensing of global land surface temperature: Definition, methods, products, and applications, *Rev. Geophys.*, 61, e2022RG000777, <https://doi.org/10.1029/2022RG000777>, 2023.

Liu, X., Li, Z.-L., Li, J.-H., Leng, P., Liu, M., and Gao, M.: Temporal upscaling of MODIS 1-km instantaneous land surface temperature to monthly mean value: Method evaluation and product generation, *IEEE Trans. Geosci. Remote Sens.*, <https://doi.org/10.1109/TGRS.2023.3247428>, 2023.

Liu, X., Tang, B.-H., Yan, G., Li, Z.-L., Liang, S.: Retrieval of Global Orbit Drift Corrected Land Surface Temperature from Long-term AVHRR Data, *Remote Sens.*, 11, 2843, <https://doi.org/10.3390/rs11232843>, 2019.

Jiménez, C., Prigent, C., Ermida, S. L., and Moncet, J. L.: Inversion of AMSR-E observations for land surface temperature estimation: 1. Methodology and evaluation with station temperature, *J. Geophys. Res-Atmos.*, 122, 3330-3347, <https://doi.org/10.1002/2016JD026144>, 2017.

Mao, K., Shi, J., Li, Z. L., Tang, H.: An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data, *J Geophys. Res-Atmos.*, 112, 1-17, <https://doi.org/10.1029/2007JD008428>, 2009.

Martins, J. P., Coelho e Freitas, S., Trigo, I. F., Barroso, C., Macedo, J.: Copernicus

Global Land Operations-Lot I “Vegetation and Energy” Algorithm Theoretical Basis Document. Land Surface Temperature—LST, 1, 2019.

Reiners, P., Asam, S., Frey, C., Holzwarth, S., Bachmann, M., Sobrino, J., Göttsche, F., Bendix, J. Kuenzer, C.: Validation of AVHRR Land Surface Temperature with MODIS and In Situ LST—A TIMELINE Thematic Processor, *Remote Sens.*, 13, 3473, <https://doi.org/10.3390/rs13173473>, 2021.

Wan, Z., Zhang, Y., Zhang, Q., Li, Z. L.: Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, *Remote Sens. Environ.*, 83, 163-180, [https://doi.org/10.1016/S0034-4257\(02\)00093-7](https://doi.org/10.1016/S0034-4257(02)00093-7), 2002.

Wan, Z.: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, *Remote Sens. Environ.*, 140, 36–45, <https://doi.org/10.1016/j.rse.2013.08.027>, 2014.

Wang, K., and Liang, S.: Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites, *Remote Sens. Environ.*, 113, 1556–1565, <https://doi.org/10.1016/j.rse.2009.03.009>, 2009.

Xing, Z., Li, Z. L., Duan, S. B., Liu, X., Zheng, X., Leng, P., Gao, M., Zhang, X., Shang, G.: Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime MODIS instantaneous observations, *ISPRS J. Photogramm.*, 178, 51-67, <https://doi.org/10.1016/j.isprsjprs.2021.05.017>, 2021.

Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., Vinnikov, K. Y., Sun, D., and Tian, Y.: Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties,

IEEE Trans. Geosci. Remote Sens., 50, 704-713,
<https://doi.org/10.1109/TGRS.2011.2162338>, 2012.