https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

Generation of angular-normalized, cloud-filled, 0.01°-downscaled land surface temperature from 2018 to 2023 based on official FY-4A dataset

Qiang Na[1,2,3], Biao Cao[4], Boxiong Qin[5], Tian Hu[6], Hua Li[1,2,3], Lixin Dong[7], Huanyu Zhang[8,2], Wenfeng Zhan[9], Qinhuo Liu[1,2,3]

[1]National Engineering Research Center for Satellite Remote Sensing Applications, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China,
[2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
[3]State Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute of Chinese Academy of Sciences, Beijing, 100101, China
[4]State Key Laboratory of Earth Surface Processes and Disaster Risk Reduction, the Advanced Interdisciplinary Institute of Satellite Applications, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[5]Guangdong Provincial Key Laboratory of Applied Botany & Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China.
[6]Department of Environment Research and Innovation, Luxembourg Institute of Science and Technology, Belvaux 4362, Luxembourg
[7]Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing 100081, China
[8]State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[9]Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, International Institute for Earth System Science, Nanjing University, Nanjing, China

Correspondence to: Biao Cao (caobiao@bnu.edu.cn)

Abstract. Land surface temperature (LST) is an essential climate variable in geophysical, ecological, and environmental researches. Remote sensing provides a unique observation approach for obtaining large-scale LST products. However, current official LST datasets (such as FY-4A) are limited by the unaddressed thermal radiation directionality effect, and suffer the spatial discontinuities due to the pervasive presence of clouds. What's more, the geostationary LST products have relatively coarser resolution than those of polar-orbiting satellites due to trade-off between spatial and temporal resolutions. Based on the official hourly FY-4A LST dataset, this study proposes a novel framework for generating angular-normalized, cloud-filled, and 0.01°-downscaled LST (ANCFDS-LST) product, encompassing directional (T_{dir}), nadir (T_{nadir}), and hemispherical (T_{hemi}) LST layers. First, the angular-normalized T_{nadir} and T_{hemi} were generated using a time-evolving kernel driven model (TEKDM) with the inputs of multi-temporal FY-4A T_{dir}. Subsequently, hypothetical clear-sky LST were predicted using a CatBoost model optimized via Bayesian methods. The cloudy-sky LST values were then derived through a cloud radiation force (CRF) correction. Finally, the 0.05° all-weather T_{dir}, T_{nadir}, and T_{hemi} values were downscaled to 0.01° resolution using an improved hybrid downscaling algorithm (IHDA) combining fusion and kernel-based methods. Taking the daytime clear-sky near-nadir VNP21A1 LST as reference, the 0.05° T_{dir} before angular-normalization has a root mean

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

squared difference (RMSD) of 6.21 K and a mean bias difference (MBD) of -4.04 K, whereas the angularly normalized T_{nadir} has a much smaller RMSD of 3.48 K and a better MBD of -2.13 K. For the all-weather T_{hemi}, temperature-based validation over 15 sites in the Heihe River Basin and the Tibetan Plateau shows a root mean squared error (RMSE) and mean bias error (MBE) of 2.99 K and -0.77 K under clear-sky conditions, 4.56 K and -1.56 K under cloudy-sky conditions. After the spatial downscaling, the 0.01° all-weather T_{hemi} with abundant texture details exhibits an RMSE (MBE) of 3.99 K (-1.32 K) over 15 sites. The generated LST products from 2018 to 2023 over the FY-4A disk exhibit enhanced angular consistency, spatial continuity, and finer resolution, offering valuable support for subsequent LST-related applications. The ANCFDS-LST data is freely available at https://doi.org/10.11888/RemoteSen.tpdc.303249 (last access: 30 January 2026; Na et al., 2026).

1 Introduction

Land Surface Temperature (LST) quantifies the thermal properties of the Earth's land surface and serves as a driving force of climate change, radiation budgets, water cycle, and atmospheric processes (Wei et al., 2020, 2021). Compared to polar-orbiting satellite LST products, geostationary LSTs have the advantage of enhanced temporal resolution, which is more suitable to characterize the dynamic variations of land surface thermal conditions (Li et al., 2023c). To ensure both retrieval accuracy and computational efficiency, numerous easy-to-implement methods such as the split-window (SW) and temperature and emissivity separation (TES) methods have been developed for the geostationary satellites over the past decades (Li et al., 2013). These methods have been successfully applied in the generation of official LST products, including the Fengyun-4A (FY-4A) Advanced Geosynchronous Radiation Imager (AGRI) LST product (Dong et al., 2013, 2023a), Geostationary Operational Environmental Satellites R-Series (GOES-R) Advanced Baseline Imager (ABI) LST products (Yu et al., 2009), and the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) LST product (Freitas et al., 2013; Trigo et al., 2011).

Despite the significant advancement of LST retrieval methods, current geostationary LST products still suffer from three major limitations: 1) Most existing LST retrieval methods assume that surface-emitted radiance is isotropic. However, the complex structure and heterogeneous sub-pixel temperature distribution lead to different LST values when observing a pixel from different directions at the same time, i.e., the thermal radiation directionality (TRD) effect (Cao et al., 2019a). As a result, current LST products are directionally dependent, requiring to be normalized into a reference direction. 2) Because thermal-infrared (TIR) signals cannot penetrate clouds, current LST products exhibit significant spatial discontinuities, with more than half of the land surface often obscured by cloud cover (Stubenrauch et al., 2013). Therefore, generating all-weather (including clear-sky and cloudy-sky conditions) LST has attracted considerable research interest in the TIR remote sensing community (Jia et al., 2024; Wu et al., 2021). 3) A trade-off between spatial and temporal resolution is inherent to LST products. Hourly LST products from geostationary satellites typically have a spatial resolution of 2~5 km. Improving sensor performance is one direct approach, while spatial downscaling methods offer a more practical and efficient alternative for enhancing the texture detail of LST products (Sun et al., 2024; Wu et al., 2021; Zhan et al., 2013).

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

The influence of TRD effect during summer is as large as 4.0 K in sparsely vegetated areas and 5.1 K in urban regions (Coll et al., 2019; Du et al., 2023, 2025; Zhan et al., 2025). The semi-physical kernel-driven model (KDM) is regarded as the most potential approach to reduce the LST angular dependence and achieve the aim of angular normalization (Cao et al., 2019b, 2021; Michel et al., 2023). It simulates the LST angular distribution through a linear combination of several kernel functions. The primary step is to calibrate the kernel coefficients, which then allows for correcting the directional LST (i.e., T_{dir}) to a nadir LST (i.e., T_{nadir}) or to a hemispherical LST (i.e., T_{hemi}) by integrating over the upper hemisphere. Solving for three or four unknown kernel coefficients typically requires at least three or four simultaneous multi-angle observations which cannot be satisfied by current satellite sensors. To address this problem, previous studies have either omitted one anisotropic kernel (Qin et al., 2023; Teng et al., 2023) or assumed that certain kernel coefficients remain constant over broad spatial or temporal scales (Chang et al., 2025; Ermida et al., 2017, 2018a, b; Vinnikov et al., 2012). Recently, Qin et al. (2025) proposed a time-evolving kernel-driven model (TEKDM) with seven parameters which captures the multi-temporal multi-angle LST patterns within a single day and enables the coefficient calibration in the overlapping region of two geostationary satellites. To broaden the application region of TEKDM, Na et al. (2024b) normalized the LST in the overlapping region of TERRA/AQUA Moderate Resolution Imaging Spectroradiometer (MODIS) and GOES-16 ABI LST products. Results showed the root mean squared error (RMSE) was reduced from 3.29 K to 2.34 K. Here, the FY-4A AGRI and TERRA/AQUA MODIS official LST products were jointly employed for solving the TEKDM and achieving the angular normalization of FY-4A LST (i.e., producing T_{nadir} and T_{hemi}).

Due to the lack of TIR information of the land surface under cloudy-sky conditions, all-weather LST estimation for geostationary satellites typically relies on various types of auxiliary data and can be categorized into three main approaches: interpolation-based methods, surface energy balance (SEB) based methods, and simulation-based methods. Interpolation-based methods utilize spatially or temporally adjacent information to estimate missing LST values (Hong et al., 2021, 2022; Quan et al., 2018; Wang et al., 2024b). These methods can effectively preserve fine spatial textures, but they are uncertain under extensive cloud cover (Jia et al., 2024). Furthermore, they do not account for the influence of cloud radiation force (CRF). The SEB-based methods are commonly used to calculate the CRF effects in all-weather LST estimation. For example, Jia et al. (2021) proposed an iterative CRF correction method, Liu et al. (2023) solved a quartic equation to perform CRF correction, and Zhang et al. (2024) developed an analytical CRF correction formula. LST values after CRF correction can represent the actual thermal properties of the land surface and are recommended for large-scale applications (Jia et al., 2024; Wu et al., 2021). Simulation-based methods also show excellent potential for estimating all-weather LST, but they suffer from coarse resolution and substantial biases in simulated LST (Ding et al., 2022; Dong et al., 2022). The rapid advancement of machine learning (ML) models offers promising opportunities to improve the all-weather LST estimation. Zhang et al. (2024) proposed a two-step gap-filling algorithm in recent. In the first step, hypothetical LST values are estimated using an ML model with the input of reanalysis data. Then, an SEB-based CRF correction is applied to generate the all-weather LST. Results showed stable accuracy with the maximum RMSE within 4 K. In this study, Zhang's method is employed as the

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

basic gap-filling framework. The estimation of hypothetical LST is routinely based on clear-sky directional LST (i.e., T_{dir}). Here, introducing the angular independent clear-sky nadir and hemispherical LST (i.e., T_{nadir}, T_{hemi}) have the potential to improve the reliability of hypothetical LST and further enhance the accuracy of generated cloudy-sky LST.

Downscaling methods are widely used to produce LST products with high spatial resolution. These methods could be generally classified into three categories: kernel-based approaches (Dong et al., 2020; Zhan et al., 2013; Zhang et al., 2020; Zheng et al., 2024), fusion-based approaches (Tang et al., 2024; Wang et al., 2024b), and hybrid downscaling method combining both kernel and fusion-based approaches (Dong et al., 2023b; Li et al., 2023b; Xia et al., 2019). Kernel-based methods typically establish a relationship between LST and regression kernels such as the normalized vegetation index (NDVI) at a coarse resolution. This relationship is then applied at a finer scale to generate high-resolution LST. These methods have evolved from simple linear regressions using single variables to ML-based regressions incorporating multiple kernels (Agam et al., 2007; Ebrahimy and Azadbakht, 2019; Xu et al., 2024; Zheng et al., 2024). Fusion-based methods aim to estimate fine-scale LST variability using multi-resolution LST data as inputs (Tang et al., 2024; Wang et al., 2024b; Wu et al., 2015). In these approaches, high-resolution LST temporal variation is typically estimated by weighting the coarse-resolution LST temporal variation of neighbouring similar pixels. The estimated temporal variation is then added to the high-resolution LST at the initial time to derive the final LST at the target time. Hybrid downscaling method combining both kernel and fusion-based approaches offers improved accuracy and computational efficiency than single method (Dong et al., 2023b), which was adopted in this study. However, this method requires gap-free, high-resolution LST at the initial time as input, which is difficult to obtain at the full-disk scale due to widespread presence of clouds over large areas. The annual temperature cycle (ATC) model has the potential to provide the necessary gap-free high-resolution LST texture information (Quan et al., 2018; Zhan et al., 2016) and thereby ensure the generation of 0.01° LST at the full-disk scale.

The aim of this study is generating hourly, angular-normalized, cloud-filled, and 0.01°-downscaled LST (ANCFDS-LST) product based on the FY-4A official LST dataset. The generation of ANCFDS-LST consists of the following key steps: First, the TEKDM was calibrated using FY-4A and MODIS official LST products to generate daytime nadir and hemispherical LST at a 0.05° resolution. Second, an ML-based model was trained to generate hypothetical clear-sky LST, using angular-normalized LST as labels. Then, the all-weather LST was produced through an analytical CRF correction process. 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 for producing the 0.01° all-weather directional, nadir, and hemispherical LST products. The hemispherical LST is validated using 15 sites in the Heihe River Basin and the Tibetan Plateau, which measure the hemispherical longwave radiation via in-situ pyrgeometer. The nadir LST is cross-validated using Visible Infrared Imaging Radiometer Suite (VIIRS) near nadir LST. The structure of this study is as follows: Section 2 describes the remote sensing and reanalysis data, cross-validation data and in-situ validation data. Section 3 presents the TEKDM-based angular normalization method, the all-weather LST estimation method, and the IHDA downscaling approach.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Section 4 provides the results of the generated LST products. Section 5 gives the discussion and limitation of this study. Section 6 and 7 introduce the data availability and the main conclusions, respectively.

2 Data

2.1 Input remote sensing and reanalysis data

Table 1 lists the information of the required 15 datasets for the three main steps in the generation of ANCFDS-LST (namely, LST angular normalization, all-weather LST estimation, and spatial downscaling). First, FY-4A and MODIS directional LSTs (i.e., 2 datasets) were used to generate angular-normalized nadir LST (T_{nadir}) and hemispherical LST (T_{hemi}). Next, 10 remote sensing and reanalysis products were employed to drive the generation of hypothetical clear-sky LST and the application of CRF correction, for estimating all-weather T_{dir}, T_{nadir}, and T_{hemi} at a 0.05° spatial resolution. Finally, these three 0.05° all-weather LST products were downscaled using 0.01° regression kernels obtained from 3 datasets, including MYD11A1, GTOPO30 DEM, and ERA5 Land.

Table 1. The employed remote sensing and reanalysis dataset.

Step	Product	Variable	Resolution	Date range	Usage	Access Link
Step 1: LST angular normalization	1) FY-4A	LST	4 km, hourly	2018 - 2023	LST angular normalization	http://data.nsmc.org.cn
	2) MxD11A1	LST	1 km, daily	2018 - 2023	LST angular normalization	https://lpdaac.usgs.gov/product_search/
Step 2: All-weather LST estimation	3) MCD12Q1	Land cover type (LCT)	500 m, yearly	2018 - 2023	Hypothetical clear-sky LST estimation	https://lpdaac.usgs.gov/product_search/
	4) Köppen-Geiger maps	Climate type (CT)	1 km	-	Hypothetical clear-sky LST estimation	http://www.gloh2o.org/koppen
	5) GLASS	Fractional vegetation cover (FVC)	0.05°, 8 days	2018 - 2023	Hypothetical clear-sky LST estimation	https://www.glass.hku.hk/download.html
	6) GTOPO30	Digital elevation model (DEM)	1 km	-	Hypothetical clear-sky LST estimation	https://www.usgs.gov/search
	7) ERA5 Land	Air temperature and dew-point temperature	0.1°, hourly	2018 - 2023	Hypothetical clear-sky LST estimation	https://cds.climate.copernicus.eu
	8) MxD11A1	LST	1 km, daily	2018 - 2023	Hypothetical clear-sky LST estimation	https://lpdaac.usgs.gov/product_search/
	9) ERA5	Cloudless and cloudy radiation	0.25°, hourly	2018 - 2023	Hypothetical clear-sky LST estimation;	https://cds.climate.copernicus.eu

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

					CRF correction	
	10) GLASS	Albedo	0.05°, daily	2018 - 2022	Hypothetical clear-sky LST estimation; CRF correction	https://www.glass.hku.hk/download.html
	11) MCD43C3	Albedo	0.05°, 16 days	2023	Hypothetical clear-sky LST estimation; CRF correction	https://lpdaac.usgs.gov/product_search/
	12) FY-3B MuSyQ	Broad band emissivity (BBE)	1 km, daily	2018 - 2023	Hypothetical clear-sky LST estimation; CRF correction	-
Step 3: Spatial downscaling	13) MYD11A1	LST	1 km, daily	2018 - 2023	Simulating LST at reference time to be fused	https://lpdaac.usgs.gov/product_search/
	14) GTOPO30	DEM	1 km	-	Regression kernel	https://www.usgs.gov/search
	15) ERA5 Land	Air temperature	0.1°, hourly	2018 - 2023	Regression kernel	https://cds.climate.copernicus.eu

For the first step (i.e., LST angular normalization), the FY-4A directional LST (LST_{FY}) and MODIS LST (LST_{MxD}) products (i.e., dataset 1-2 in Table. 1) were jointly utilized to calibrate the TEKDM model, then the clear-sky T_{nadir} and T_{hemi} could be generated with the inputs of kernels and calibrated kernel coefficients. The official FY-4A LST was retrieved using a SW algorithm with land surface emissivity (LSE) estimated via the NDVI-based threshold method (Dong et al., 2013, 2023a). The MODIS MxD11A1 LST was retrieved using a generalized split-window (GSW) algorithm, incorporating land-cover-based LSE as input (Wan and Dozier, 1996). Both FY-4A and MODIS LST products were resampled to a same spatial resolution (i.e., 0.05°) using plain average before the jointly estimation. The pixels with the view zenith angle (VZA) greater than 70° were masked (Freitas et al., 2013). To reduce systematic discrepancies between these two datasets, the 0.05° FY-4A LST was linearly adjusted to match the 0.05° MODIS MxD11A1 LST using a linear transformation (i.e., $LST_{MxD} = a \times LST_{FY} + b$). The parameters a and b were determined based on nighttime matchups with the conditions of VZA< 50°, VZA difference< 5°, and LST difference< 5 K, referenced to Ermida et al. (2017, 2018a). A total of 132,972,698 matchup pairs were collected from 2018/1/1 to 2023/12/31 in the disk of FY-4A, resulting in a slope of a = 1.0240 and an intercept of b = -5.7378.

The second step (i.e., all-weather LST estimation) includes the estimation of hypothetical clear-sky LST and the CRF correction. In the hypothetical clear-sky LST estimation, the dataset 3-7 in Table 1 (including the MCD12Q1 IGBP land cover type (LCT) product (dataset 3), the Köppen-Geiger climate type (CT) product (dataset 4), the Global LAnd Surface Satellite (GLASS) fractional vegetation cover (FVC) product (dataset 5), the GTOPO30 DEM data (dataset 6), the ERA5-Land 2-m air temperature (T_{2m}), and dew-point temperature (D_{2m}) data (dataset 7)) were employed to depict surface property (Wei et al., 2019). Moreover, the dataset 8-12 in Table 1 were further used to calculate three additional variables to depict surface thermal conditions (including the ATC-simulated LST, the surface incoming radiation (R_{in}), and hypothetical clear-sky skin temperature ($T_{ERA5\text{-}skin}$)).

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

The ATC model has three parameters as shown in Eq. 1, which was calibrated with the input of MxD11A1 LST product (dataset 8). In total, four ATC models were obtained since TERRA and AQUA MODIS sensors provide observations during both day and night. A 3 × 3 spatial median filter was applied to fill invalid ATC parameters caused by persistent cloud coverage. Then, the LST values simulated using the calibrated ATC model are used in the estimation of hypothetical clear-sky LST. These ATC-simulated LSTs are denoted as $T_{ATC,d1}$, $T_{ATC,d2}$, $T_{ATC,n1}$, and $T_{ATC,n2}$, where subscript "*d*" and "*n*" refer to daytime and nighttime data, respectively, and "1" and "2" indicate TERRA and AQUA platforms.

$$T_{ATC}(DOY) = MAST + YAST \times \cos\left[\frac{2\pi}{365}\left(DOY - dx\right)\right] \quad (1)$$

where $T_{ATC}(DOY)$ is the ATC-simulated LST in the day-of-year (DOY). $MAST$ is the annual mean surface temperature, $YAST$ is the yearly temperature amplitude, and dx is the phase shift. The QC band was filtered to ensure high-quality inputs for the ATC model calibration.

The R_{in} is the sum of absorbed shortwave and longwave radiation as calculated in Eq. 2, with the input of surface downward shortwave radiation ($SDSR_{clr}$), land surface albedo ($Albedo$), surface downward longwave radiation ($SDLR_{clr}$) and broad band emissivity (BBE, ε_{bb}). Both $SDSR_{clr}$ and $SDLR_{clr}$ under hypothetical clear-sky condition are extracted from ERA5 products (dataset 9). The $Albedo$ is from the GLASS product from 2018 to 2022 (dataset 10) and supplemented by MCD43C3 for 2023 (dataset 11). A Harmonic ANalysis of Time Series (HANTS) method was employed to smooth and fill the gaps in MCD43C3 product in this study (Zhou et al., 2023, 2022). The BBE is generated from MUlti-source data SYnergized Quantitative (MuSyQ) remote sensing system (dataset 12) (Li et al., 2019). The datasets 9-12 were resampled to 0.05° resolution to match the LST_{FY}.

$$R_{in} = (1 - Albedo) \cdot SDSR_{clr} + \varepsilon_{bb} \cdot SDLR_{clr} \quad (2)$$

Based on the Stefan–Boltzmann's law, the ERA5 hypothetical clear-sky surface thermal radiation (e.g., surface downward longwave radiation ($SDLR_{clr}$) and surface net thermal radiation STR_{clr}) could be used to calculate the hypothetical clear-sky skin temperature ($T_{ERA5\text{-}skin}$) in the scale of 0.25° to depict temporal dynamics of LST. It was then physically downscaled to 0.05° resolution using DEM data, assuming a temperature lapse rate (K) of 6.5 K/km (Minder et al., 2010). Therefore, the $T_{ERA5\text{-}skin}$ was calculated by the following equation:

$$T_{ERA5-skin} = \sqrt[4]{\frac{SDLR_{clr} - STR_{clr} - \left(1 - \varepsilon_{bb}\right) SDLR_{clr}}{\varepsilon_{bb}\sigma_s}} + K \cdot \left(DEM_i - DEM_m\right) \quad (3)$$

where σ_s is Stefan–Boltzmann's constant 5.67×10^{-8}. The DEM_i and DEM_m refer to the elevation of the 0.05° pixel and the mean elevation of the corresponding 0.25° pixel, respectively.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

After generating the hypothetical clear-sky LST, a CRF correction is required to convert it into cloudy-sky LST. This correction is physically based on the surface energy balance differences between clear-sky and cloudy conditions. Specifically, cloud-free and cloudy SDSR from ERA5 reanalysis data (dataset 9), together with albedo data (dataset 10-11), are used to describe the shortwave radiation budget. For the longwave radiation budget, ERA5 cloud-free and cloudy-sky SDLR data (dataset 9), BBE data (dataset 12), the hypothetical clear-sky LST and cloudy-sky LST are required. Then, the only unknown cloudy-sky LST (expressed by the sum of hypothetical clear-sky LST and CRF correction value) can be retrieved by solving the SEB equation. More details will be given in section 3.2.

In the third step (i.e., spatial downscaling), a new algorithm combining existing fusion and kernel-based methods was proposed to downscale LST to the target time, consisting of two main processes. First, a 0.05° gap-free high-resolution LST at the initial time is required. In this study, $T_{ATC,d2}$ simulated by the ATC model (dataset 13) is resampled as the initial LST. Second, the 0.05° LST difference between the initial and target time is predicted using an ML method. The input features include $T_{ATC,d2}$; AQUA MODIS ATC parameters for both day ($MAST_{d2}$, $YAST_{d2}$, and dx_{d2}) and night ($MAST_{n2}$, $YAST_{n2}$, and dx_{n2}), derived from dataset 13; *DEM* data (dataset 14); and ERA5-Land T_{2m} (dataset 15). Aimed to apply this spatial downscaling model at the 0.01° scale, the T_{2m} data were further downscaled to 0.01° resolution using high-resolution DEM and the temperature lapse rate (K) of 6.5 K/km. Finally, the gap-free high-resolution LST at the target time is obtained by adding the estimated 0.01° LST difference to 0.01° $T_{ATC,d2}$.

2.2 Cross-validation dataset for T_{nadir}

For cross-validation of the generated T_{nadir}, the VIIRS VNP21A1 LST data (https://lpdaac.usgs.gov/product_search/) in 2020 was adopted as reference. The VIIRS onboarded Suomi National Polar-orbiting Partnership (S-NPP) provides clear-sky TIR observations at the local time of 1:30 p.m. The VNP21 LST was retrieved using a TES algorithm taking observations from M14 – M16 bands as input. Recent temperature-based (T-based) evaluation showed its RMSE is 1.79 K during nighttime and 2.79 K during the day (Na et al., 2024a). Here, the VNP21 LST values with the VZA< 5° was employed for cross evaluation of T_{nadir} as did by Wei et al. (2025). Firstly, the 1-km VNP21 LST product was resampled to the 0.05° resolution and the normalized FY-4A hourly T_{nadir} product was temporally interpolated to match the exact overpass time of S-NPP VIIRS. Then, the root mean square difference (RMSD), mean bias difference (MBD), and coefficient of determination (R^2) were used as three evaluation indicators. The extracted VNP21 LST results in four typical days are shown in Fig. 1. Only several narrow strips were extracted due to the limitation of VIIRS field-of-view (VZA< 5°). VNP21 LST shows significant seasonal variation. 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 southern hemisphere.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

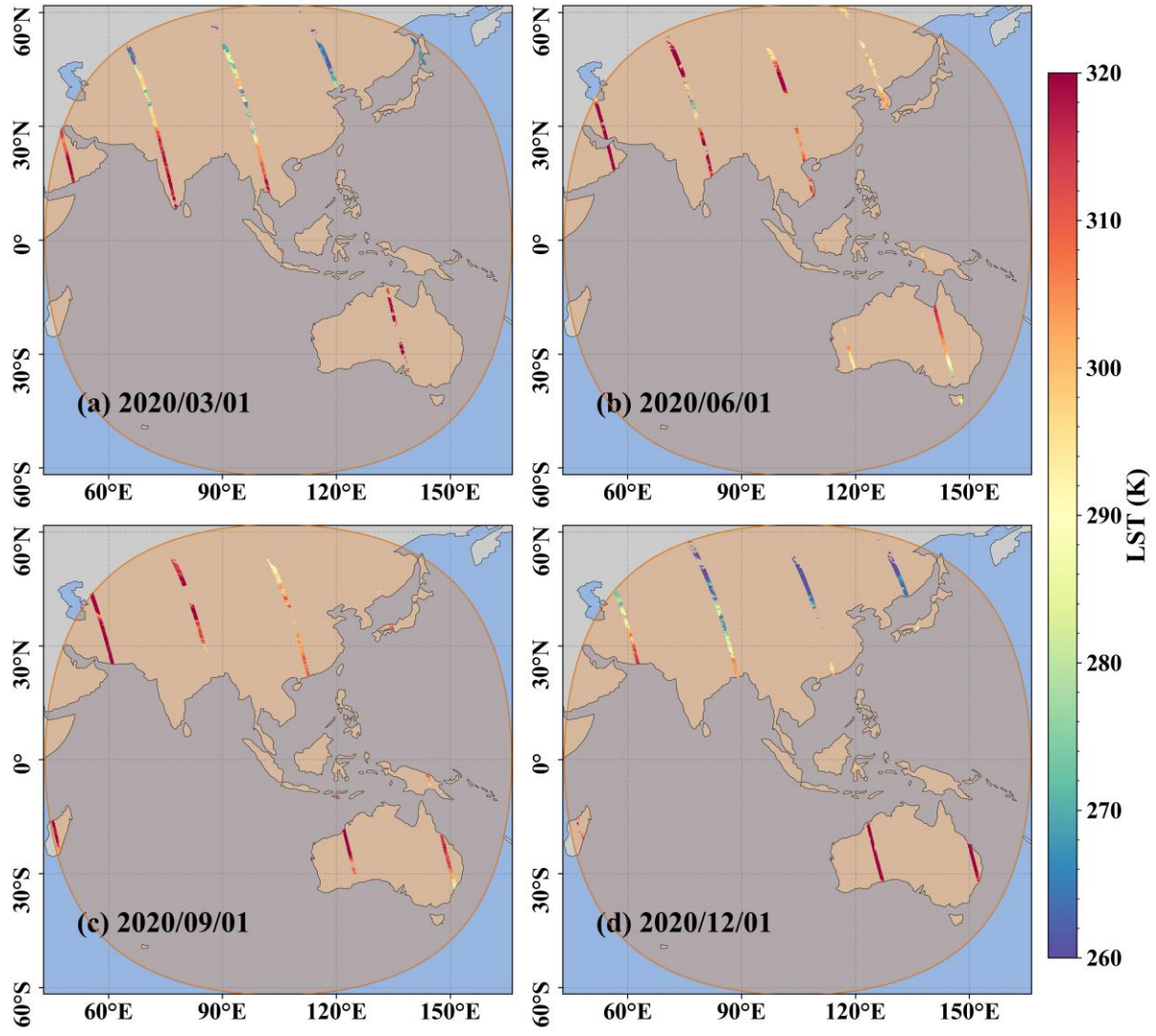


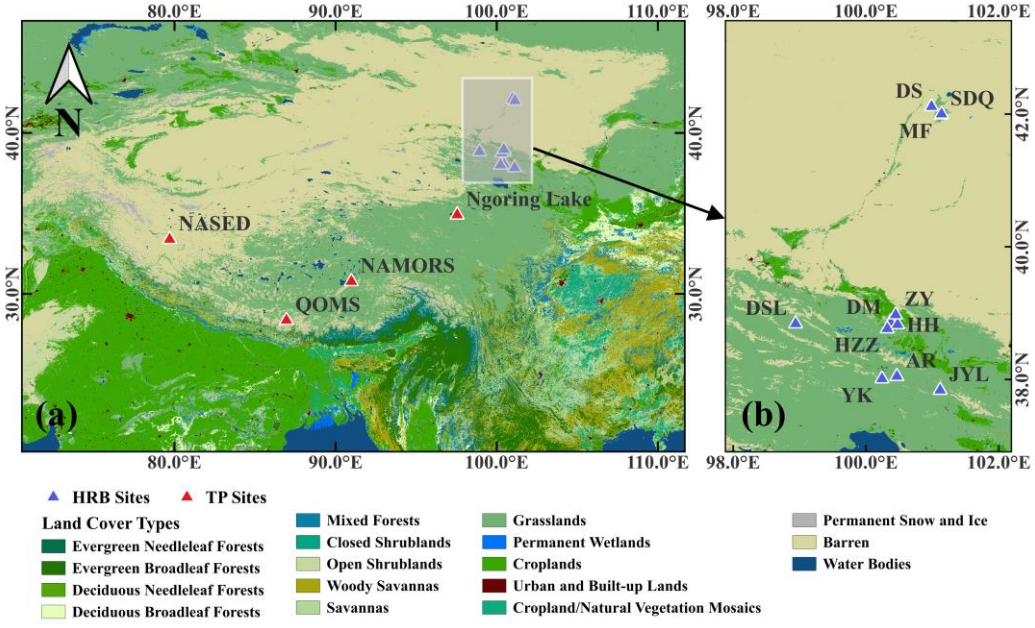
Figure 1: Spatial distribution of the extracted VNP21 LST strips in four typical days (a) 2020/3/1, (b) 2020/6/1, (c) 2020/9/1, (d) 2020/12/1. The FY-4A observed area is marked in orange region.

2.3 In-situ validation dataset for T_{hemi}

T-based evaluation (i.e., directly comparing satellite-derived LST with in-situ LST) is the most widely used validation method and should be conducted whenever possible (Guillevic et al., 2018; Li et al., 2014, 2023a; Na et al., 2024a). The in-situ pyrgeometer measures both upward and downward hemispherical longwave radiation, which can be converted into hemispherical LST to perform the T-based validation for the normalized T_{hemi} products. As shown in Fig. 2, 11 sites from the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) experiment within the Heihe River Basin (HRB)

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

and 4 sites from the land-atmosphere interactions dataset over the Tibetan Plateau in 2020 were selected, which had been used in the studies of (Che et al., 2019; Li et al., 2025; Ma et al., 2024b).



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

Table 2. Information of the in-situ sites

Region	Site name	Longitude (° E)	Latitude (° N)	Elevation (m)	Temporal resolution (minute)	Land cover type
HRB	AR	100.46	38.05	3038	10	GRA
	DM	100.37	38.86	1562	10	CRO
	DSL	98.94	38.84	3787	10	GRA
	HH	100.48	38.83	1526	10	CRO
	HZZ	100.32	38.77	1722	10	GRA
	DS	100.99	42.11	925	10	BSV
	MF	101.13	41.99	937	10	GRA
	JYL	101.12	37.84	3725	10	GRA
	SDQ	101.14	42.00	932	10	GRA
	YK	100.24	38.01	4070	10	GRA
	ZY	100.45	38.98	1461	10	CRO
TP	QOMS	86.95	28.36	4297	60	GRA
	NAMORS	90.98	30.77	4738	60	GRA
	NASED	79.70	33.39	4252	60	BSV
	Ngoring Lake	97.55	34.91	4314	30	GRA

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

As listed in Table 2, three land cover types are represented across the 15 validation sites: 10 grassland (GRA) sites, 3 cropland (CRO) sites, and 2 barren/sparse vegetation (BSV) sites. The temporal resolution of the in-situ observations varies by site: 10 minutes for the HRB sites, 30 minutes for the Ngoring Lake site, and 60 minutes for the remaining TP sites. Linear interpolation was applied to align in-situ measurements with the exact satellite observation time. The in-situ LST (T_{insitu}) is calculated using the Stefan–Boltzmann's law, as shown in Eq. 4:

$$T_{insitu} = \sqrt[4]{\frac{SULR_{insitu} - (1-\varepsilon_{bb})\,SDLR_{insitu}}{\varepsilon_{bb}\sigma_s}} \tag{4}$$

where $SULR_{insitu}$ and $SDLR_{insitu}$ are the surface upward and downward longwave radiation measured by the in-situ pyrgeometer, respectively. ε_{bb} is the broadband emissivity derived from the FY-3B MuSyQ product. Outliers were identified and removed using the "3σ-Hampel identifier" method for each site (Davies and Gather, 1993; Pearson, 2002). The RMSE, mean bias error (MBE), and R^2 were used as evaluation metrics to validate the LST products.

3 Method

The generation of 0.01° ANCFDS-LST involves three main steps: 1) angular normalization of daytime LST; 2) generation of 0.05° all-weather LST using clear-sky LST as training labels; 3) downscaling of all-weather LST to 0.01° resolution using an IHDA method. The overall flowchart is shown in Fig. 3. First, 0.05° clear-sky FY-4A and resampled MxD11A1 directional LST products were matched and bias corrected for solving the unknown parameters of TEKDM using auxiliary information of viewing geometry and local time. For generating T_{nadir} and T_{hemi} with the inputs of FY-4A T_{dir} and calibrated KDM coefficients, the calibrated coefficients were temporally aggregated within a window of 17 days to ensure a complete spatial coverage of the FY-4A disk. Second, clear-sky T_{dir}, T_{nadir}, and T_{hemi} during the day, as well as nighttime FY-4A LST (where $T_{dir} = T_{nadir} = T_{hemi}$ due to the negligible angular effect at night) were used as training labels to predict the hypothetical clear-sky LST under cloudy conditions. A Bayesian optimization-based categorical boosting (CatBoost) model was trained with these variables described in Subsection 2.1. The training samples were spatially determined using a widely-used conditioned Latin hypercube sampling (cLHS) approach. The predicted hypothetical clear-sky LST requires a CRF correction using the cloudless and cloudy radiation budget variables, yielding the final cloudy-sky LST of three viewing angles. In the third step, the 0.05° all-weather LST was downscaled to 0.01° resolution using a proposed IHDA method. This approach downscales the LST difference (ΔLST) between the 0.05° all-weather LST and the 0.05° $T_{ATC,d2}$ at the reference time (i.e., 13:30) using another CatBoost model. After predicting three ΔLSTs at 0.01° resolution and applying a residual redistribution procedure, the final downscaled 0.01° ΔLSTs was obtained. The final 0.01° all-weather T_{dir}, T_{nadir}, and T_{hemi} were derived by summing their own downscaled 0.01° ΔLST with the corresponding 0.01° ATC-simulated LST.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

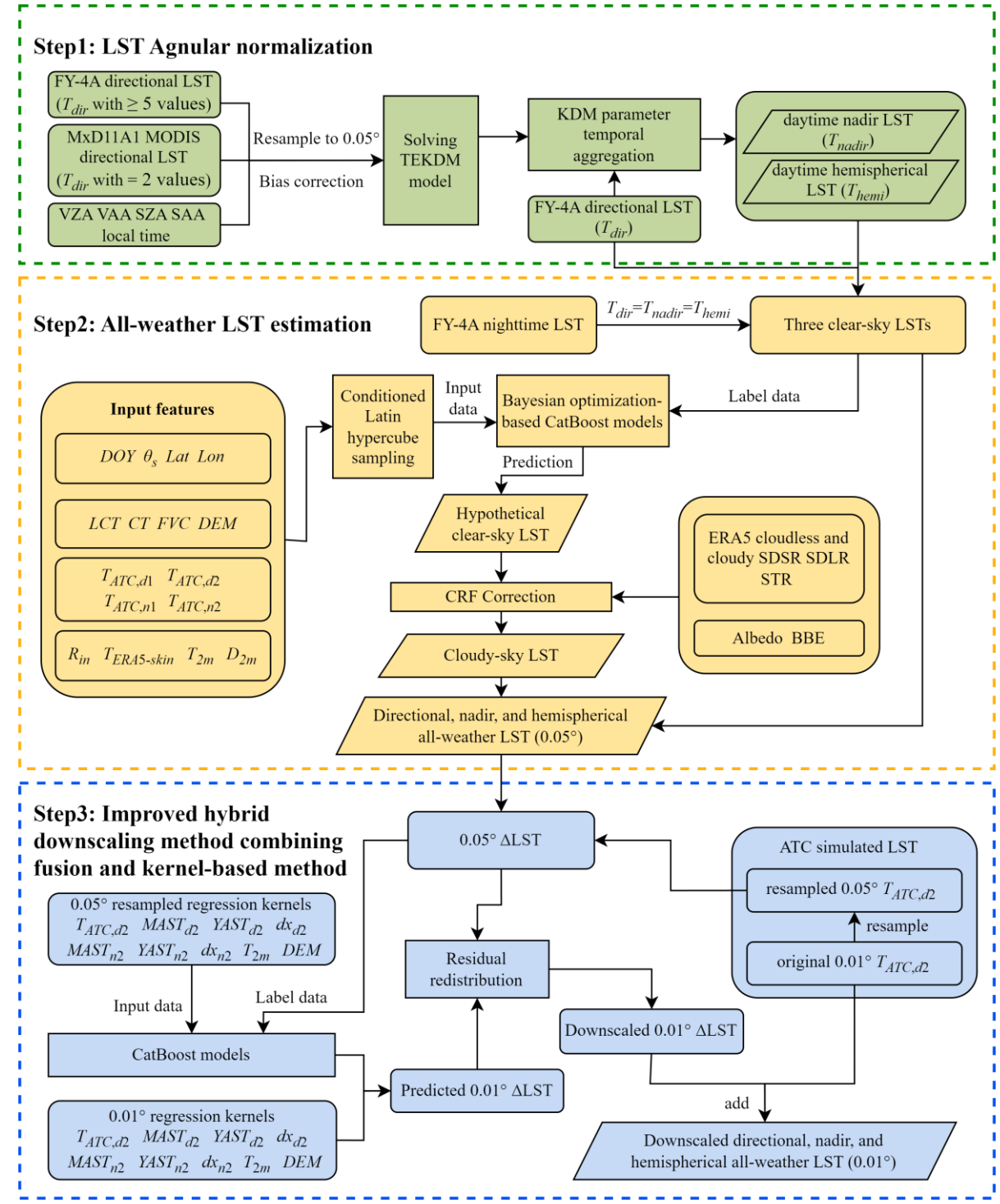


Figure 3: The overall workflow in this study.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

3.1 Angular normalization of clear-sky LST

The FY-4A LST disk has different viewing geometries for each pixel, and these values need to be normalized to a reference direction before subsequent applications. The KDM offers a solution by simulating the LST angular distribution through a linear combination of several kernel functions. However, traditional KDM typically requires three or four simultaneous clear-sky LST observations to calibrate the unknown kernel coefficients, which cannot be satisfied by FY-4A satellite. To enhance its applicability, Qin et al. (2023, 2025) proposed a time-evolving KDM (TEKDM) by coupling a diurnal temperature cycle (DTC) model (i.e., depicting LST temporal variation) and a KDM model (i.e., depicting LST angular variation) as shown in Eq. 5. Based on this TEKDM model, Na et al. (2024b) further estimated daytime nadir LST using GOES-16 and MODIS LST products, and obtained a significantly improved accuracy.

$$T_{dir}(t,\theta_s,\theta_v,\Delta\varphi) = f_{iso}(t)\times\left[1+\alpha\cdot K_{GapFraction}(\theta_v)+\beta\cdot\cos\theta_s\cdot K_{Hotspot}(\theta_s,\theta_v,\Delta\varphi,W)\right] \quad (5)$$

where $T_{dir}(t,\theta_s,\theta_v,\Delta\varphi)$ represents the directional LST observed at the local time of t, with the solar zenith angle (SZA) of θ_s, the VZA of θ_v, and the relative azimuth angle (RAA) of $\Delta\varphi$. The term $f_{iso}(t)$ is the isotropic kernel coefficient. The coefficients α and β are unknowns corresponding to the gap fraction kernel ($K_{GapFraction}$) and hotspot kernel ($K_{Hotspot}$), respectively. W is another unknown parameter related to the hotspot width. The temporal dynamics of isotropic kernel coefficient (i.e., $f_{iso}(t)$) during daytime can be effectively modeled using a DTC model as given in Eq. 6:

$$f_{iso}(t) = T_0 + T_a\times\left[\frac{\pi}{\omega}\cdot(t-t_m)\right] \quad (6)$$

where T_0 is the temperature at sunrise, T_a represents the amplitude of daily LST variation, ω is the length of the daytime period, t_m is the local time when the LST reaches its maximum. The nighttime LST was not processed here, as the angular effect is generally negligible during nighttime. In this study, the LSF kernel and the Chen kernel were selected to represent $K_{GapFraction}$ and $K_{Hotspot}$ respectively, as shown in Eq. 7-9.

$$K_{GapFraction}(\theta_v) = K_{LSF}(\theta_v) = \frac{1+2\cos\theta_v}{\sqrt{0.96}+1.92\cos\theta_v} - \frac{1}{4}\cdot\frac{\cos\theta_v}{1+2\cos\theta_v} + 0.15\cdot\left(1-e^{\frac{-0.75}{\cos\theta_v}}\right) - 1.0304 \quad (7)$$

$$K_{Hotspot}(\theta_s,\theta_v,\Delta\varphi,W) = K_{Chen}(\theta_s,\theta_v,\Delta\varphi,W) = e^{\frac{-\xi}{\pi\cdot W}} \quad (8)$$

$$\xi = \arccos\left(\cos\theta_v\cdot\cos\theta_s+\sin\theta_v\cdot\sin\theta_s\cdot\cos\Delta\varphi\right) \quad (9)$$

where ξ is the angular distance between the viewing direction and the solar direction. The TEKDM involves seven unknown parameters to be calibrated: T_0, T_a, ω, t_m, α, β, and W, therefore, at least seven clear-sky daytime observations are required as inputs. To abundant the angular information of the input data, two daytime clear-sky MODIS observations (i.e., TERRA and

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

AQUA) were utilized. Then, the TEKDM can be solved for the condition with ≥ 5 clear-sky FY-4A observations. A "trust-region-reflective" algorithm was employed for the non-linear optimization of Eq. 5, with initial values and parameter boundaries set according to Qin et al. (2025). Then, a 17-day moving average (i.e., covering 8 days before and after the target day) was applied to the parameters of α, β, and W for each pixel for two purposes: filling gaps in pixels with fewer than seven observations, and reducing the impact of outliers in the estimated TEKDM parameters. As shown in Eq. 10, the LST at any direction (i.e., the $T_{dir}(t,\theta_s,\theta_v,\Delta\varphi)$) can be simulated with the inputs of FY-4A direction LST (i.e., $T_{dir}(t,\theta_s,\theta_v^{FY},\Delta\varphi^{FY})$ and averaged α, β, and W parameters.

$$\frac{T_{dir}\left(t,\theta_s,\theta_v,\Delta\varphi\right)}{T_{dir}\left(t,\theta_s,\theta_v^{FY},\Delta\varphi^{FY}\right)}=\frac{1+\alpha\cdot K_{LSF}\left(\theta_v\right)+\beta\cdot\cos\theta_s\cdot K_{Chen}\left(\theta_s,\theta_v,\Delta\varphi,W\right)}{1+\alpha\cdot K_{LSF}\left(\theta_v^{FY}\right)+\beta\cdot\cos\theta_s^{FY}\cdot K_{Chen}\left(\theta_s,\theta_v^{FY},\Delta\varphi^{FY},W\right)} \tag{10}$$

Specifically, the nadir LST can be derived when setting $\theta_v = 0°$ as shown in Eq. 11:

$$T_{nadir}\left(t,\theta_s,0^{\circ},\Delta\varphi\right)=T_{dir}\left(t,\theta_s,\theta_v^{FY},\Delta\varphi_v^{FY}\right)\cdot\frac{1+\alpha\cdot K_{LSF}\left(0^{\circ}\right)+\beta\cdot\cos\theta_s\cdot K_{Chen}\left(\theta_s,0^{\circ},\Delta\varphi,W\right)}{1+\alpha\cdot K_{LSF}\left(\theta_v^{FY}\right)+\beta\cdot\cos\theta_s^{FY}\cdot K_{Chen}\left(\theta_s,\theta_v^{FY},\Delta\varphi_v^{FY},W\right)} \tag{11}$$

Correspondingly, the normalized hemispherical LST (T_{hemi}) could be calculated through the integration as shown in Eq. 12.

$$\begin{aligned}T_{hemi}\left(t,\theta_s\right)&=\left(\frac{1}{\pi}\cdot\int_0^{2\pi}\int_0^{\pi/2}T_{dir}^4\left(t,\theta_s,\theta_v,\Delta\varphi\right)\cdot\cos\theta_v\cdot\sin\theta_v d\theta_v d\varphi_v\right)^{\frac{1}{4}}\\&=\frac{T_{dir}\left(t,\theta_s,\theta_v^{FY},\Delta\varphi_v^{FY}\right)\cdot\left(\int_0^{2\pi}\int_0^{\pi/2}\left(1+\alpha\cdot K_{LSF}\left(\theta_v\right)+\beta\cdot\cos\theta_s\cdot K_{Chen}\left(\theta_s,\theta_v,\Delta\varphi,W\right)\right)^4\cdot\cos\theta_v\cdot\sin\theta_v d\theta_v d\varphi_v\right)^{\frac{1}{4}}}{\pi\cdot\left(1+\alpha\cdot K_{LSF}\left(\theta_v^{FY}\right)+\beta\cdot\cos\theta_s^{FY}\cdot K_{Chen}\left(\theta_s,\theta_v^{FY},\Delta\varphi_v^{FY},W\right)\right)}\end{aligned} \tag{12}$$

To accelerate the computation of T_{hemi}, the integral in the numerator of Eq. 12 was approximated using a third-order polynomial fit based on the input variables α, β, W, θ_s and K_{Chen} at nadir. This approach significantly reduces computational cost while introducing an uncertainty less than 0.05 K.

3.2 All-weather LST estimation method

The all-weather LST estimation method consists of two main sub-steps. First, a hypothetical clear-sky LST was predicted using a Bayesian optimization-based CatBoost model. Second, the CRF correction was applied to estimate the cloudy-sky LST, ultimately producing cloud-filled 0.05° all-weather directional, nadir, and hemispherical LST.

ML-based LST reconstructing have been widely used because it could accurately simulate the non-linear correlation between the hypothetical clear-sky LST and vegetative, meteorological, and topographical parameters (Li et al., 2024a; Ma et al., 2024a; Zhang et al., 2024). The CatBoost is a recently developed model based on the traditional gradient-boosting decision

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

tree (GBDT) framework with several enhancements: 1) Ordered target statistics, which encode categorical features without target leakage by using permutation-based strategies; 2) Ordered boosting, which minimizes gradient bias by employing permutation-driven training to reduce over fitting; 3) Oblivious trees, which apply the same splitting criterion across all nodes at each level, thereby improving speed and regularization (Prokhorenkova et al., 2018). These improvements enhance the model's robustness and efficiency, especially when handling large-scale data-sets containing categorical variables, and have proven effective in LST reconstruction studies (Dai et al., 2025). Here, the CatBoost model was used to capture the complex non-linear relationship (i.e., f_{clr}) between hypothetical clear-sky LST and auxiliary parameters, as represented in Eq. 13.

$$T_{clr} = f_{clr}\left(DOY, \theta_s, Lat, Lon, LCT, CT, FVC, DEM, T_{ATC,d1}, T_{ATC,d2}, T_{ATC,n1}, T_{ATC,n2}, R_{in}, T_{ERA5-skin}, T_{2m}, D_{2m}\right) \quad (13)$$

where the input features include the geolocation and temporal parameters, such as the *DOY*, θ_s, latitude (*Lat*), longitude (*Lon*). The other parameters in Eq. 13 have been introduced in Table 1.

During the model training, the target labels consisted of daytime T_{dir}, T_{nadir}, and T_{hemi}, along with nighttime FY-4A LST from 2018/1/1 to 2023/12/31. Considering the large volume of label data, a two-step approach was employed to determine the number and geolocation of sample points as introduced in Appendix A. As shown in Fig. A1, 37,000 locations distributed over the FY-4A disk were selected for each year. Then, the clear-sky LST and corresponding input features were extracted to train the CatBoost model in Eq. 13. To better align with the 0.25° spatial resolution of ERA5, a FY-4A pixel was strictly considered as clear-sky only when all surrounding pixels within a 5 × 5 neighbourhood were under clear-sky conditions. Six CatBoost models were trained for each year from 2018 to 2023. The average number of extracted samples per year was 138,817,085, which were randomly divided into training, testing, and Bayesian optimization sets in a ratio of 6:3:1. The Bayesian optimization method adjusted the CatBoost hyperparameters by reconstructing the posterior distribution of the cost function, defined as the average fitting RMSE from two-fold cross-validation. After the ML training, the CatBoost model can predict the hypothetical clear-sky T_{dir}, T_{nadir}, and T_{hemi} under cloudy-sky conditions, which requires CRF correction.

The core of the CRF correction lies in establishing the relationship between net radiation changes and CRF-induced LST changes, based on the surface radiation budget equation. The CRF-corrected LST is equal to the sum of hypothetical clear-sky LST with CRF-induced LST variation (i.e., ΔT_{CRF}). In this study, an analytical correction equation (Eq. 14) proposed by Zhang et al. (2024) was adopted for the CRF correction of T_{hemi}.

$$\gamma \cdot \Delta T_{CRF} = \Delta R_{in} - \varepsilon_{bb}\sigma_s\left[\left(T_{clr}^{hemi} + \Delta T_{CRF}\right)^4 - \left(T_{clr}^{hemi}\right)^4\right] \quad (14)$$

where ΔT_{CRF} is the correction value for hemispherical LST; γ is an energy transfer parameter, which could be estimated by Eq. 15. A temporal median filter with a 17-day window and a spatial median filter with a 3 × 3 window were applied to

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

replace outlier values in the calculation of γ. ΔR_{in} is the change in surface incoming radiation induced by the CRF effect, which could be estimated by Eq. 16; T_{clr}^{hemi} is predicted hypothetical clear-sky hemispherical LST using Eq. 13.

$$\gamma = \frac{\left[R_{in}(t_n) - R_{in}(t_{sr})\right] - \varepsilon_{bb}\sigma_s\left[T_{clr}^{hemi}(t_n)^4 - T_{clr}^{hemi}(t_{sr})^4\right]}{T_{clr}^{hemi}(t_n) - T_{clr}^{hemi}(t_{sr})} \tag{15}$$

$$\Delta R_{in} = (1 - Albedo)\cdot\left(SDSR_{clr} - SDSR_{cld}\right) + \varepsilon_{bb}\cdot\left(SDLR_{clr} - SDLR_{cld}\right) \tag{16}$$

where R_{in} is the clear-sky surface incoming radiation calculated by Eq. 2. t_n and t_{sr} denote the noon time and sunrise time. The subscripts "*clr*" or "*cld*" indicate that the parameter is under hypothetical clear-sky or actual cloudy-sky conditions. After the determination of γ and ΔR_{in}, the only unknown parameter ΔT_{CRF} in Eq. 14 can be estimated. Here, it was analytically solved through neglecting its quartic and cubic terms (see Eq. 17).

$$6\varepsilon_{bb}\sigma_s\left(T_{clr}^{hemi}\right)^2\Delta T_{CRF}^2 + \left(\gamma + 4\varepsilon_{bb}\sigma_s\left(T_{clr}^{hemi}\right)^3\right)\Delta T_{CRF} - \Delta R_{in} = 0 \tag{17}$$

After the pixel-by-pixel estimation of ΔT_{CRF}, the mean (μ_{CRF}) and standard deviation (σ_{CRF}) of ΔT_{CRF} could be calculated for each image. To reduce the impact of extreme values, the maximum and minimum limits of ΔT_{CRF} were set to $\mu_{CRF} \pm 3\sigma_{CRF}$. In the end, the cloudy-sky T_{dir}, T_{nadir}, and T_{hemi} are obtained by summing CatBoost-estimated hypothetical clear-sky LST (i.e., using Eq. 13) and ΔT_{CRF} (i.e., using Eq. 17).

3.3 IHDA combining fusion and kernel-based methods

To meet the 1-km spatial resolution and 1-hour temporal resolution requirement of GCOS (https://gcos.wmo.int/site/global-climate-observing-system-gcos/essential-climate-variables/land-surface-temperature), downscaling methods have garnered more and more attention for enhancing the spatial textures of low-resolution LST products (e.g., FY-4A LST). Fusion-based downscaling approaches typically require high-resolution LST data as input, whereas kernel-based methods require LST-related auxiliary data (i.e., regression kernels). Combining fusion-based and kernel-based downscaling methods could achieve higher accuracy than using either method alone (Li et al., 2023b; Xia et al., 2019). Dong et al. (2023b) recently proposed a simple and effective downscaling (SED) method that utilizes clear-sky, high-resolution Landsat 8 LST at an initial time to downscale low-resolution MODIS data at a target time. However, this requirement is difficult to be satisfied over the full-disk region due to the frequent presence of clouds. To address this limitation, an IHDA method is proposed leveraging the ATC-fitted spatiotemporal trend surface of LST (Liang et al., 2025). First, the ATC model is adopted to estimate gap-free LST at the initial time of 13:30 ($T_{ATC,d2}$) using Eq. 1 with the driven data of MYD11A1 LST products. Although ATC model may introduce additional uncertainty, recent studies have shown that incorporating meteorological data can further reduce the modeling residuals (Liu et al., 2019; Yang et al., 2024). Here, a CatBoost model is employed to establish the non-linear relationship (i.e., $f_{downscale}$) between the resampled 0.05° regression kernels and the 0.05° LST bias

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

(i.e., $\Delta LST_{0.05} = T_{0.05} - T_{ATC,d2}$, where $T_{0.05}$ is the 0.05° all-weather LST at the target time to be downscaled) as shown in Eq. 18.

$$\Delta LST_{0.05} = f_{downscale}\left(T_{ATC,d2}, MAST_{d2}, YAST_{d2}, dx_{d2}, MAST_{n2}, YAST_{n2}, dx_{n2}, T_{2m}, DEM\right) \quad (18)$$

The ATC parameters of daytime (13:30) and nighttime (01:30) MODIS LST were included as regression kernels to characterize surface thermal properties. Additionally, ERA5-Land air temperature (i.e., T_{2m}) was incorporated to capture cloud information, whereas the DEM was introduced to represent the topographic information. For model training, two typical days (i.e., the mid-month of January in winter and July in summer) were initially used to tune the hyperparameters of the CatBoost model using a Bayesian optimization approach. 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). For each image, 70% of the samples were used to train the CatBoost model, while the remaining 30% were reserved for testing. The 0.01° LST difference ($\Delta LST_{0.01}^{pred}$) was then predicted using the trained CatBoost model with 0.01° regression kernels as input:

$$\Delta LST_{0.01}^{pred} = f_{downscale}\left(T_{ATC,d2}^{0.01}, MAST_{d2}^{0.01}, YAST_{d2}^{0.01}, dx_{d2}^{0.01}, MAST_{n2}^{0.01}, YAST_{n2}^{0.01}, dx_{n2}^{0.01}, T_{2m}^{0.01}, DEM^{0.01}\right) \quad (19)$$

where the "0.01" means the input features are in 0.01° resolution. Finally, a residual redistribution process was applied to produce the final $\Delta LST_{0.01}$ as shown in Eq. 20.

$$\Delta LST_{0.01} = \Delta LST_{0.01}^{pred} + \left(\Delta LST_{0.05} - \Delta LST_{0.05}^{pred}\right)_{0.01} \quad (20)$$

where $\Delta LST_{0.05}$ represents the difference between the 0.05° all-weather LST and the ATC-simulated LST. $\Delta LST_{0.05}^{pred}$ means the 0.05° LST difference obtained by aggregating the predicted $\Delta LST_{0.01}^{pred}$. The second term in the right of Eq. 20 is the CatBoost modeling residual to be redistributed. A 3 × 3 median filter was applied to reduce the impact of extreme outliers in the residuals. The notation $(\cdot)_{0.01}$ indicates the process that the 0.05° residual (i.e., $\Delta LST_{0.05}$ - $\Delta LST_{0.05}^{pred}$) was resampled to 0.01° using the widely-used bilinear interpolation. Finally, the downscaled 0.01° all-weather directional, nadir, and hemispherical LST ($T_{0.01}$) could be calculated by summing the 0.01° ATC simulated LST at initial time ($T_{ATC,d2}(0.01°)$) and the Eq. 20 estimated $\Delta LST_{0.01}$ as below.

$$T_{0.01} = T_{ATC,d2}(0.01^{\circ}) + \Delta LST_{0.01} \quad (21)$$

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

4 Results

4.1 Results of normalized nadir LST

The spatial distributions of FY-4A directional LST T_{dir}, normalized nadir LST T_{nadir}, the LST difference between T_{nadir} and T_{dir}, the FY-4A VZA, and the histogram of these two LST products on 2020/6/24 at the UTC time of 3:00 (i.e., the Beijing time of 11:00) are shown in Fig. 4. As shown in Fig. 4c-d, the T_{nadir} closely resembles T_{dir} (with a small value of T_{nadir}-T_{dir}) when the VZA is within 40°. T_{nadir} becomes significantly higher than T_{dir} as VZA increases, with a correction value of T_{nadir}-T_{dir} exceeding 5 K near the edge of the FY-4A disk. When the VZA is small, the FY-4A observation angle is close to the nadir direction, resulting in a small correction value from T_{dir} to T_{nadir}. However, more cool vegetated elements were viewed as the VZA increases (i.e., the well-known gap fraction effect), which resulted in a much lower T_{dir} compared to T_{nadir} and further led to a larger angular correction value. The MBD between nadir and directional LST of the whole image is 1.51 K as shown in Fig. 4e. The fraction of lower LST values between 290 K and 300 K was significantly reduced after the angular correction of TRD effect.

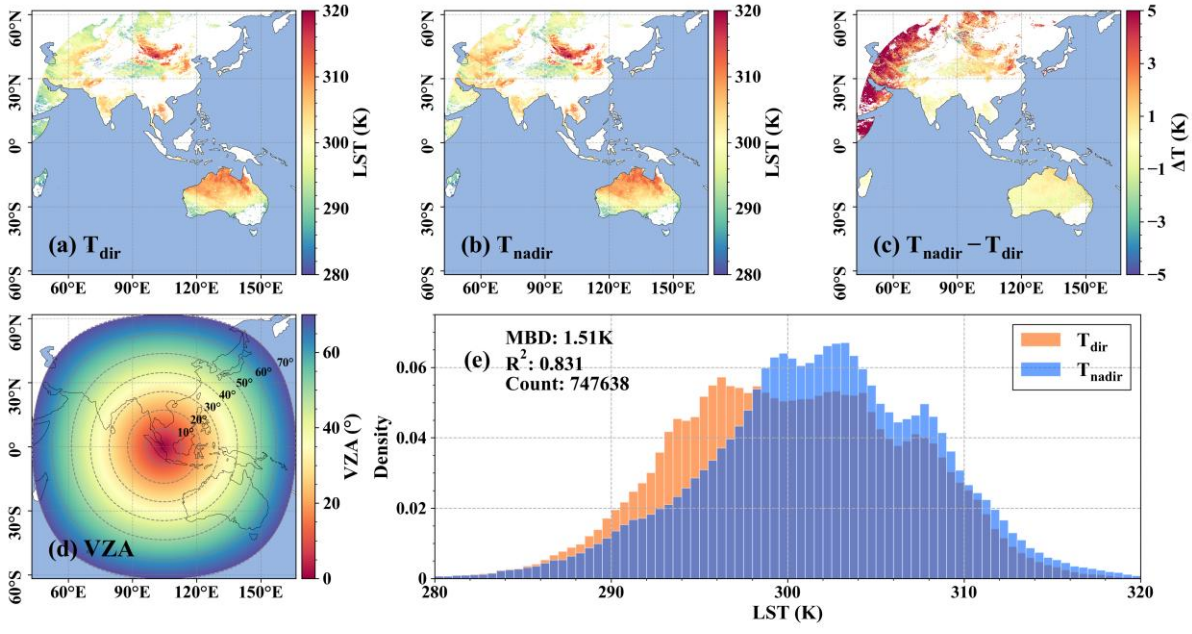


Figure 4: The spatial distribution of (a) directional LST, (b) nadir LST, (c) LST difference between nadir and directional LST, and (d) VZA on 2020/6/24 at the UTC time of 3:00. (e) histogram comparison of directional LST and nadir LST.

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. 5. The root mean squared difference (RMSD) and mean bias difference (MBD) are 6.21 K and -4.04 K for T_{dir} as shown in Fig. 5a. However, the RMSD and MBD for T_{nadir} is 3.48 K and -2.13 K as shown in Fig. 5b, with a 2.73 K (i.e., 44%) reduction in RMSD and a 1.91 K (i.e., 47%) reduction in the MBD. Fig. 5c 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

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

less than 40°. However, when VZA exceeds 40°, the RMSD for T_{dir} increases substantially (from 4.8 K to 8.6 K), while the RMSD for normalized T_{nadir} remains lower than 4.0 K. The large RMSD of T_{dir} at large VZA can be partially attributed to its significant underestimation as shown in Fig. 5d. The MBD of T_{dir} exceeds -6 K when the VZA is between 60° and 70°, whereas it is only -1.1 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.

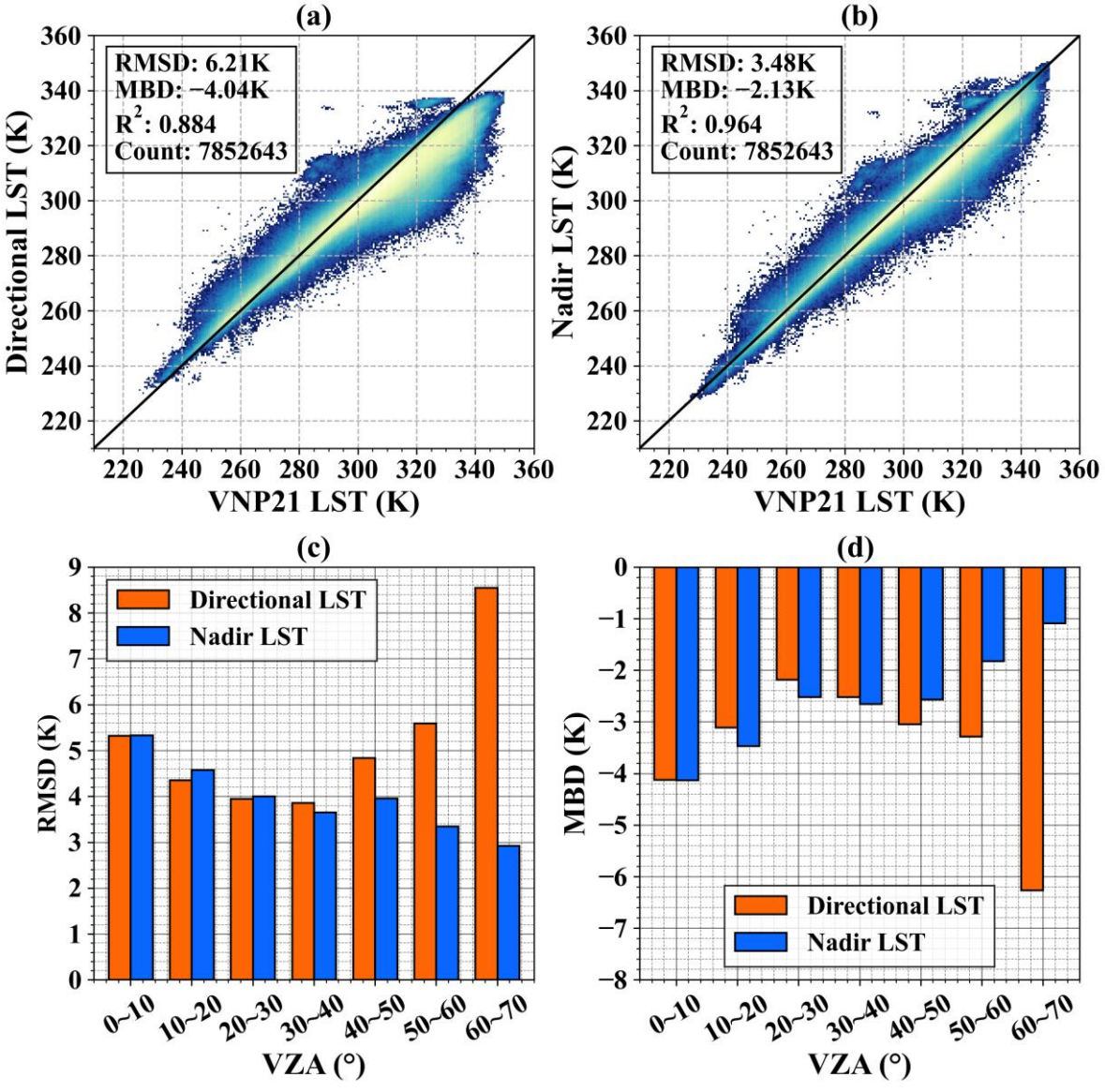


Figure 5: The cross-validation of (a) directional and (b) nadir LST against VNP21 near-nadir LST at 0.05° resolution. The angular dependence of (c) RMSD and (d) MBE for directional and nadir LST at 0.05° resolution.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

The accuracy of f_{clr} in reconstructing hypothetical clear-sky T_{dir} and T_{nadir} over the test set from 2018 to 2023 is summarized in Table 3. The average annual sample size is 41,312,444. The RMSE ranges from 2.22 K to 2.46 K for T_{dir}, and from 2.28 K to 2.52 K for T_{nadir}. RMSE values remain relatively stable from 2018 to 2021, followed by an increase of approximately 0.2 K in 2022 – 2023. On average, the RMSE for T_{nadir} is 0.06 K higher than that for T_{dir}, indicating comparable fitting performance across different CatBoost models. The mean absolute error (MAE) ranges from 1.54 K to 1.69 K for T_{dir} and from 1.60 K to 1.75 K for T_{nadir}, with an average difference of 0.06 K too. The MBE is zero for all models, indicating that no significant systematic bias was introduced in the predicted hypothetical clear-sky LST.

Table 3. The prediction accuracy for reconstructing hypothetical clear-sky nadir and directional LST

Year	RMSE (K)		MAE (K)		MBE (K)		Counts
	T_{dir}	T_{nadir}	T_{dir}	T_{nadir}	T_{dir}	T_{nadir}	
2018	2.22	2.28	1.54	1.60	0.00	0.00	40,534,869
2019	2.25	2.30	1.56	1.61	0.00	0.00	43,285,823
2020	2.27	2.34	1.57	1.63	0.00	0.00	41,883,982
2021	2.26	2.32	1.57	1.63	0.00	0.00	39,331,441
2022	2.46	2.52	1.69	1.75	0.00	0.00	41,833,769
2023	2.44	2.50	1.68	1.74	0.00	0.00	41,004,779
mean	2.32	2.38	1.60	1.66	0.00	0.00	41,312,444

As shown in Fig. 6, two typical locations (i.e., corresponding to the DM and SDQ sites) were selected to compare the time series of all-weather T_{dir} and T_{nadir}. The VNP21A1 near-nadir LST is also exhibited as a reference in Fig. 6. The hourly T_{dir} and T_{nadir} values were interpolated to match the VIIRS overpass times using linear interpolation. Overall, T_{nadir} closely aligns with T_{dir} as shown in Fig. 6a-b. However, T_{nadir} tends to be higher than T_{dir} in summer (see Fig. 6c-d) and lower than T_{dir} in winter (see Fig. 6e-f). This seasonal variation may be attributed to the varying temperature contrasts between soil and vegetation (Liu et al., 2020, 2025). In summer, more hot soil components are viewed in nadir direction, resulting in a higher T_{nadir} compared to T_{dir}. Conversely, in winter, soil temperatures become cooler than vegetation temperatures, leading to a lower T_{nadir} compared to T_{dir}. The VNP21A1 near-nadir LST confirmed this explanation of the seasonal variation of T_{nadir}.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

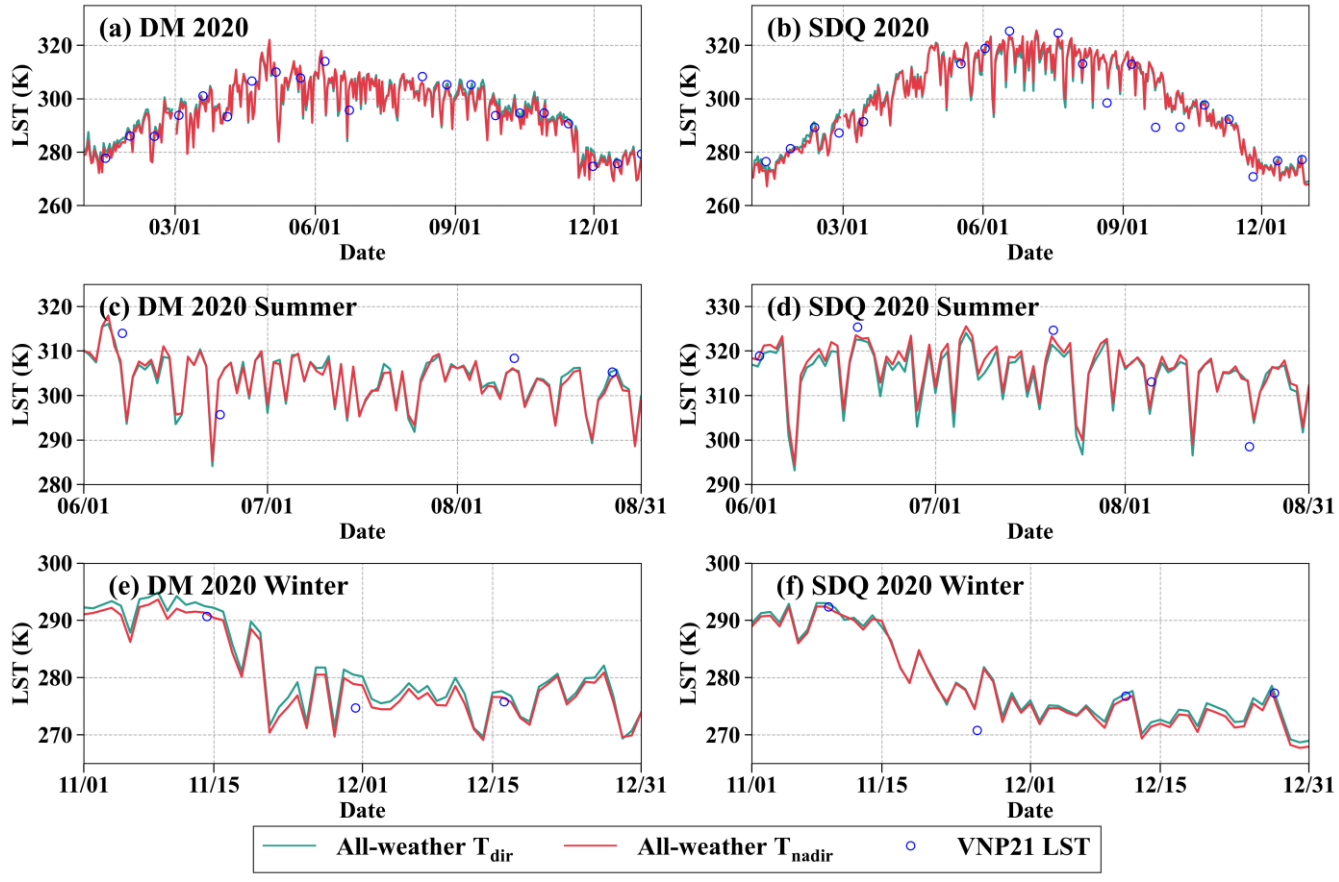


Figure 6: The time series of T_{dir}, T_{nadir}, and VNP21A1 near-nadir LST in 2020 in (a) DM site, (b) SDQ site, (c) the summer of DM site, (d) the summer of SDQ site, (e) the winter of DM site, (f) the winter of SDQ site.

In the spatial resolution downscaling process, the CatBoost model was used to establish the relationship between regression kernels and the LST difference (e.g., the nadir LST minus ATC-simulated LST) at 0.05° resolution. The prediction accuracy of each month over the test set from 2018/1/1 to 2023/12/31 for daytime and nighttime conditions is shown in Fig. 7. The nighttime RMSE ranges from 1.97 K to 2.43 K, which is consistently lower than the daytime RMSE (2.62 K – 3.04 K). This discrepancy can be partially attributed to the fact that daytime LST tends to be more heterogeneous than nighttime LST, leading to greater modeling errors during the day. The daytime RMSE increases from January to May and then decreases, with a total variation of 0.42 K. In contrast, the nighttime RMSE decreases from January to July and then rises, with a variation of 0.46 K. Overall, the results demonstrate the reliable capability of the CatBoost model in the spatial resolution downscaling.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

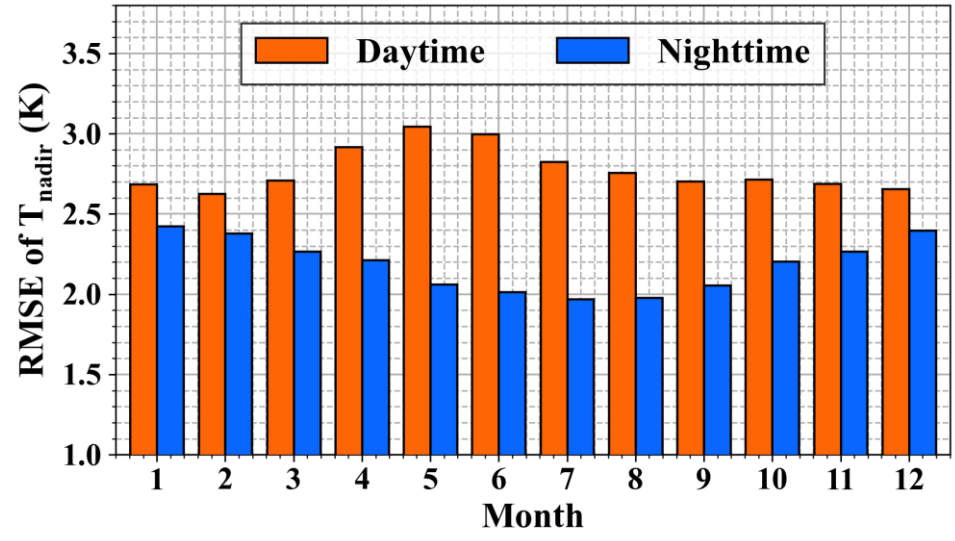


Figure 7: The prediction RMSE over test set for nadir LST during the day and night for each month.

The spatial distributions of all-weather 0.05° daytime T_{nadir} (Fig. 8a), nighttime T_{nadir} (Fig. 8c), and downscaled 0.01° daytime T_{nadir} (Fig. 8b), nighttime T_{nadir} (Fig. 8d) on 2020/6/24 are shown in Fig. 8. Generally, the spatial distribution is almost the same between 0.05° and 0.01° T_{nadir}. It demonstrates pronounced patterns of latitudinal and elevational variation (i.e., the T_{nadir} is lower in high-latitude and high-altitude regions both during the day and night). The downscaled T_{nadir} (Fig. 8b and 8d) reveals rich spatial texture, particularly in areas with rugged terrain such as in southeastern Tibetan Plateau. The temperature contrast between mountain tops and bases is more pronounced than in the 0.05° LST, which verifies the effectiveness of the IHDA downscaling method.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

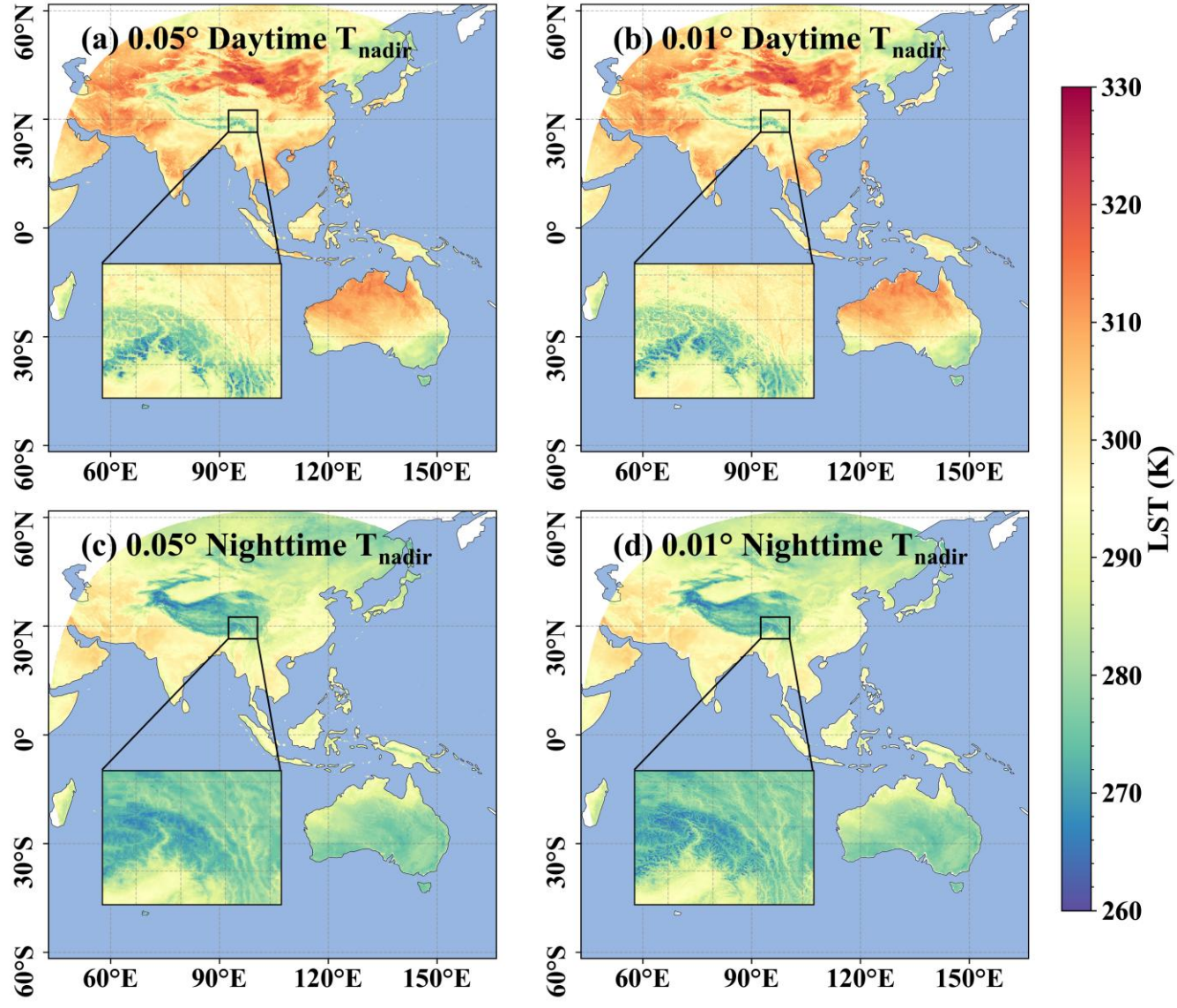


Figure 8: The spatial distribution of (a) 0.05° nadir LST, (b) 0.01° nadir LST at the UTC time of 4:00 on 2020/6/24, and (c) 0.05° nadir LST, (d) 0.01° nadir LST at the UTC time of 16:00 on 2020/6/24.

465

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

4.2 Results of normalized hemispherical LST

The spatial distribution of FY-4A directional LST T_{dir}, normalized hemispherical LST T_{hemi}, the LST difference between T_{hemi} and T_{dir}, the FY-4A VZA, and the histogram of these two LST products on 2020/6/24 at the UTC time of 3:00 are shown in Fig. 9. In Australia, the T_{hemi} (Fig. 9b) is approximately 2 K lower than T_{dir} (Fig. 9a). In the northern China, the T_{hemi} is very close to T_{dir}. However, the T_{hemi} is around 3 K higher than T_{dir} at the edge of the FY-4A disk (where the VZA exceeds 60°, see Fig. 9c-d). This spatial pattern can be partially explained by the value of the hemispherical equivalent angle (i.e., the VZA at which $T_{dir} = T_{hemi}$), which has been reported to range from 44° to 53° in relevant studies (Hu et al., 2023; Zhang et al., 2025). When the VZA is smaller than 44°, T_{dir} tends to be higher than T_{hemi} due to more hotter soil component was viewed. Conversely, for the large VZAs over 53° (e.g., at the edge of FY-4A disk), T_{dir} becomes lower than T_{hemi} as more cooler vegetation component was viewed. It is interesting that the distribution of T_{hemi} is more concentrated that that of T_{dir} since the relative lower (i.e., in the range of 290 K – 295 K) and higher (i.e., in the range of 307 K – 312 K) values of T_{dir} were normalized to be close to 300 K (Fig. 9e). Thus, the MBD between the unnormalized T_{dir} and the normalized T_{hemi} is as small as 0.2 K.

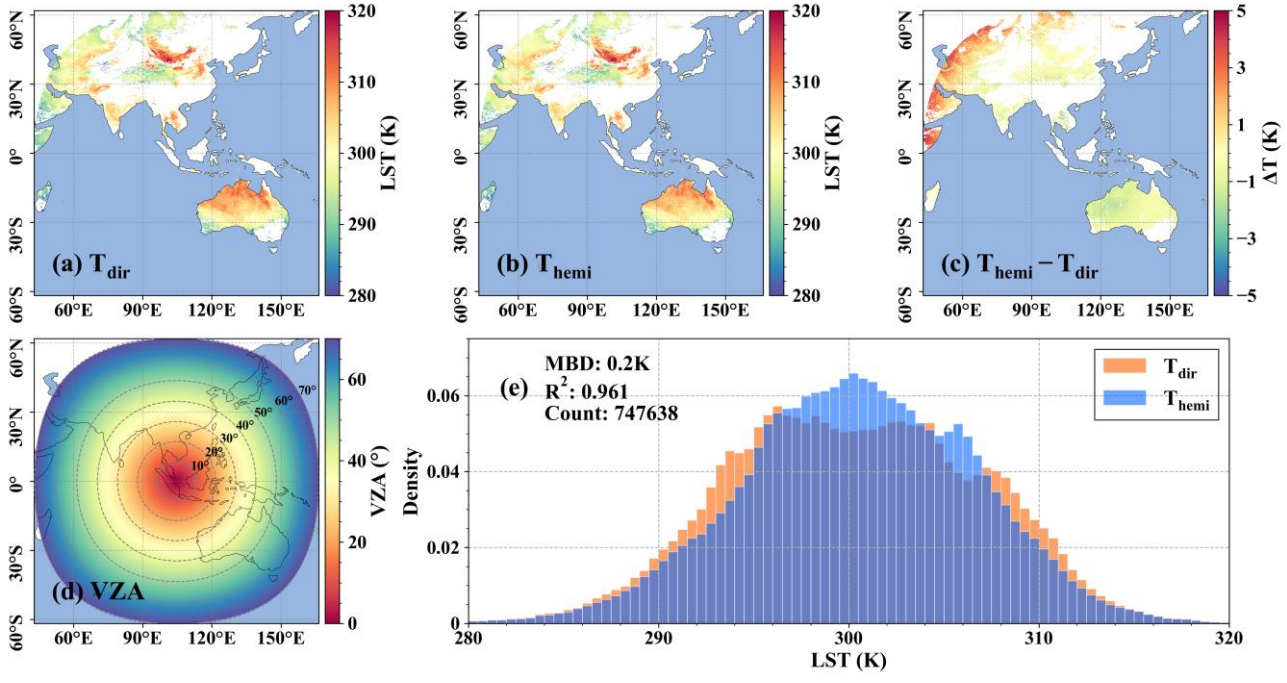
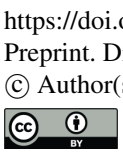


Figure 9: The spatial distribution of (a) 0.05° directional LST, (b) 0.05° hemispherical LST, (c) LST difference between hemispherical and directional LST, (d) VZA on 2020/6/24 at the UTC time of 3:00. (e) histogram comparison of directional LST and hemispherical LST.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026



Open Access Earth System Science Data Discussions

To quantitatively evaluate the 0.05° FY-4A T_{dir} and T_{hemi}, they were compared with daytime in-situ LST measurements over 15 stations (Fig. 10a-b). The results show that the RMSE decreased from 3.45 K to 3.19 K, with an improvement of 0.26 K. The MBE was reduced from 0.43 K to -0.36 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. 10c-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 a RMSE reduced by approximately 0.4-0.8 K (Fig. 10c). As shown in Fig. 10d, the MBEs of T_{hemi} is much smaller than those of T_{dir} in the morning, with a maximum reduction of 2.0 K at 11:00.

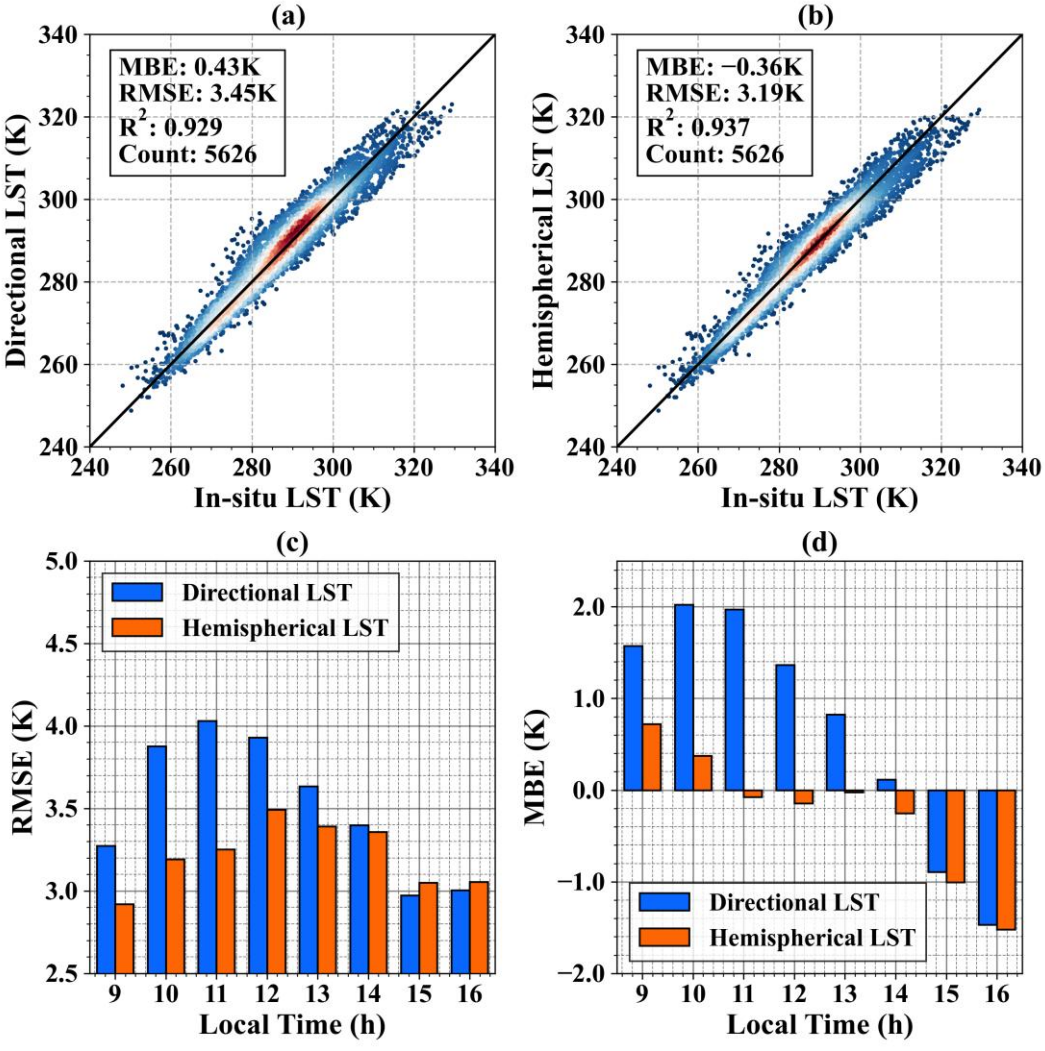


Figure 10: The accuracy of (a) FY-4A directional LST, and (b) FY-4A hemispherical LST. The temporal variation of (c) RMSE and (d) MBE for directional and hemispherical LST.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

The accuracy of reconstructing hypothetical clear-sky 0.05° T_{dir} and 0.05° T_{hemi} over the test set from 2018 to 2023 is summarized in Table 4. The RMSE ranges from 2.22 K to 2.46 K for T_{dir} and from 2.23 K to 2.46 K for T_{hemi}. The average RMSE for T_{hemi} is equal to that of T_{dir} (i.e., 2.32 K), indicating comparable performance across the different CatBoost models. The MAE ranges from 1.54 K to 1.69 K for T_{dir} (with a mean value of 1.60 K), and ranges from 1.55 K to 1.70 K for T_{hemi} (with a mean value of 1.61 K). The MBE is zero for all models, which indicates that the reconstructed hypothetical clear-sky T_{dir} and T_{hemi} are unbiased.

Table 4. The prediction accuracy for reconstructing hypothetical clear-sky hemispherical and directional LST

Year	RMSE (K)		MAE (K)		MBE (K)		Counts
	T_{dir}	T_{hemi}	T_{dir}	T_{hemi}	T_{dir}	T_{hemi}	
2018	2.22	2.23	1.54	1.55	0.00	0.00	40,534,869
2019	2.25	2.25	1.56	1.56	0.00	0.00	43,285,823
2020	2.27	2.28	1.57	1.58	0.00	0.00	41,883,982
2021	2.26	2.26	1.57	1.58	0.00	0.00	39,331,441
2022	2.46	2.46	1.69	1.70	0.00	0.00	41,833,769
2023	2.44	2.44	1.68	1.68	0.00	0.00	41,004,779
mean	2.32	2.32	1.60	1.61	0.00	0.00	41,312,444

The estimated 0.05° all-weather T_{hemi} results were further evaluated by the in-situ hemispherical LSTs from 15 in-situ stations (Fig. 11). The samples are divided into clear-sky (Fig. 11a) and cloudy-sky (Fig. 11b) conditions. The "3σ-Hampel identifier" method applied to minimize the influence of outliers at each in-situ site. Both daytime and nighttime T_{hemi} values were considered here. Under clear-sky conditions, 0.05° T_{hemi} shows an RMSE of 2.99 K and an MBE of -0.77 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 4.56 K and a more significant negative MBE of -1.56 K. It is challenging to maintain the same accuracy under cloudy-sky condition as under clear-sky condition since the uncertainty under cloudy skies is composed of the estimation of CatBoost-predicted hypothetical LST and the estimation of ΔT_{CRF}.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

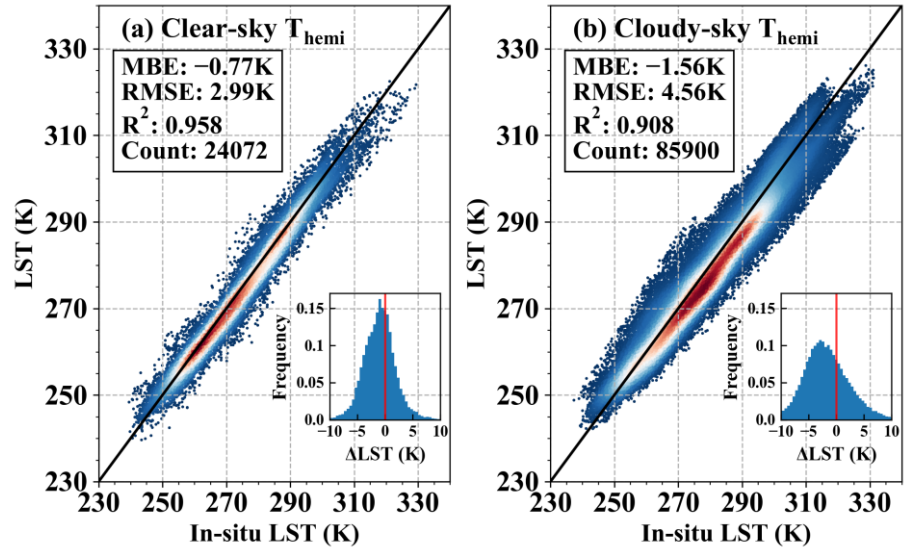


Figure 11: The accuracy of generated all-weather hemispherical LST under (a) clear-sky condition and (b) cloudy-sky condition.

Fig. 12a-b show the temporal consistency of FY-4A official LST, T_{hemi}, and in-situ LST, at the DM and SDQ sites throughout 2020 at 04:00 and 16:00 UTC (i.e., 12:00 and 00:00 Beijing time). Overall, the all-weather T_{hemi} aligns well with the in-situ LST. At the DM site during daytime (Fig. 12a), the clear-sky RMSE (MBE) of FY-4A official LST is 3.20 K (2.43 K), whereas the corresponding values for T_{hemi} are 2.15 K (0.46 K), indicating that the systematic overestimation is significantly mitigated after angular normalization. Similarly, at the SDQ site during daytime (Fig. 12b), the clear-sky RMSE (MBE) decreases from 4.77 K (4.02 K) for the FY-4A official LST to 3.60 K (2.55 K) for T_{hemi}. During nighttime, the clear-sky T_{hemi} is the same as FY-4A official LST, showing lower RMSE than in daytime. The RMSE (MBE) values are 2.38 K (-0.94 K) at the DM site and 2.2 K (-0.17 K) at the SDQ site. Under cloud-sky conditions, the RMSE (MBE) values are 3.54 K (-0.53 K) at the DM site and 4.09 K (1.97 K) at the SDQ site during daytime, and 4.58 K (-3.77 K) at the DM site and 3.46 K (-1.36 K) at the SDQ site during nighttime. In summary, the all-weather LST could effectively capture the annual temperature cycle characteristics.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

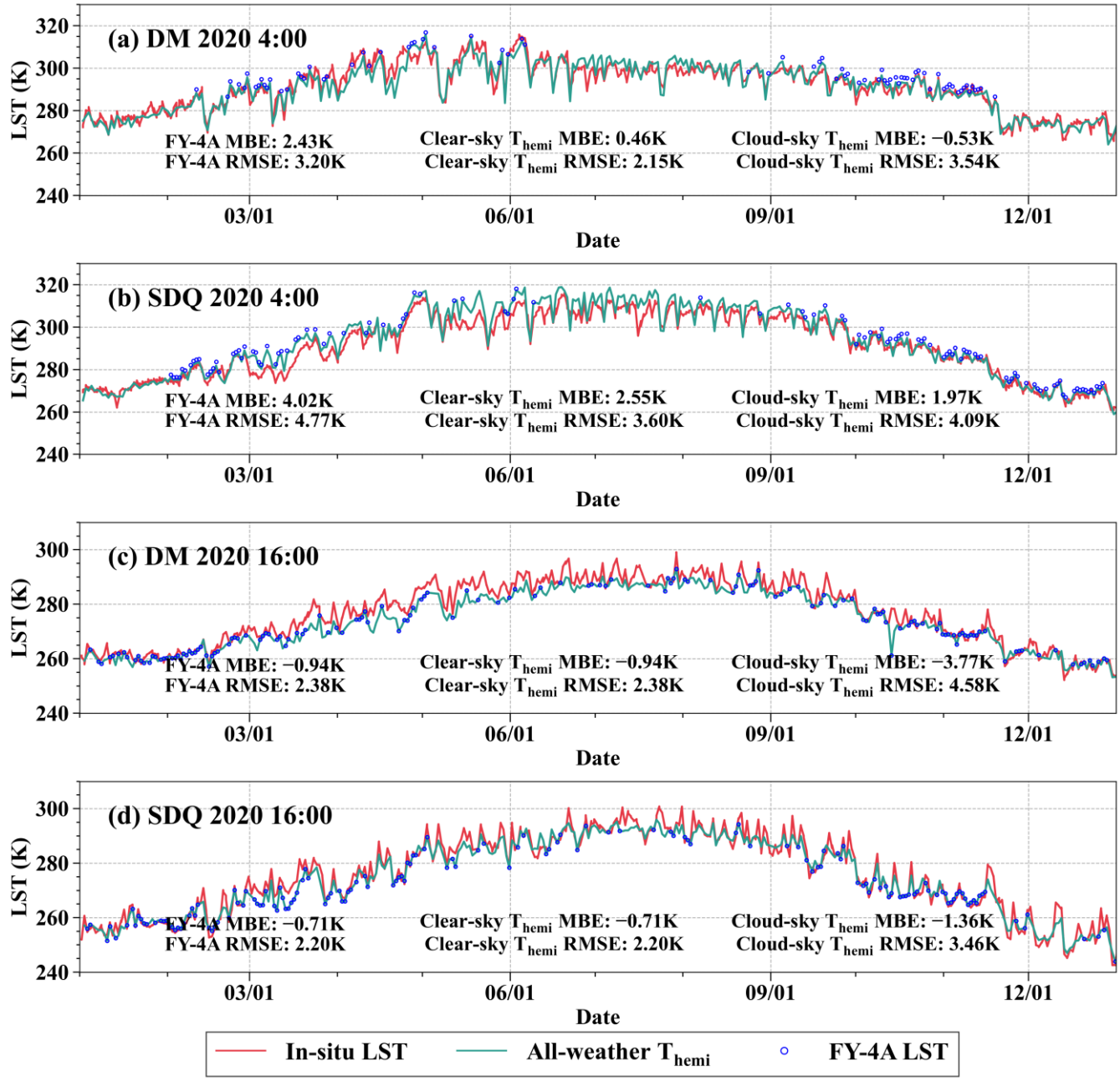


Figure 12: The temporal variation trend of in-situ LST, T_{hemi}, and FY-4A official LST in 2020 for (a) DM site at the UTC time of 4:00, (b) SDQ site at the UTC time of 4:00, (c) DM site at the UTC time of 16:00, (d) SDQ site at the UTC time of 16:00. Note that the clear-sky T_{hemi} is the same as FY-4A LST during nighttime because the TRD effect is ignored at night.

The 0.05° LST difference (i.e., the hemispherical LST minus ATC-simulated LST) fitting accuracy of the CatBoost model in the downscaling process over the test set is shown in Fig. 13a. The nighttime RMSE (1.87 K – 2.40 K) is lower than the

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

daytime RMSE (2.59 K – 2.99 K), with a difference ranging from 0.26 K to 1.00 K. The maximum daytime RMSE occurs in May, while the minimum RMSE is observed in February, with a variation of 0.4 K. The maximum nighttime RMSE is in January, and the minimum RMSE occurs in July, with a variation of 0.53 K. As shown in Fig. 13b-c, the accuracy of the 0.01° all-weather T_{hemi} is comparable to that of the all-weather 0.05° T_{hemi}, indicating that the downscaling method did not introduce additional uncertainty. The RMSE slightly decreased from 4.23 K at the 0.05° resolution to 3.99 K at the 0.01° resolution, and the absolute value of the MBE reduced by 0.08 K. Since a residual redistribution (Zheng et al., 2024) was performed during the downscaling process, the accuracy of the 0.01° T_{hemi} remains highly dependent on the original 0.05° LST product.

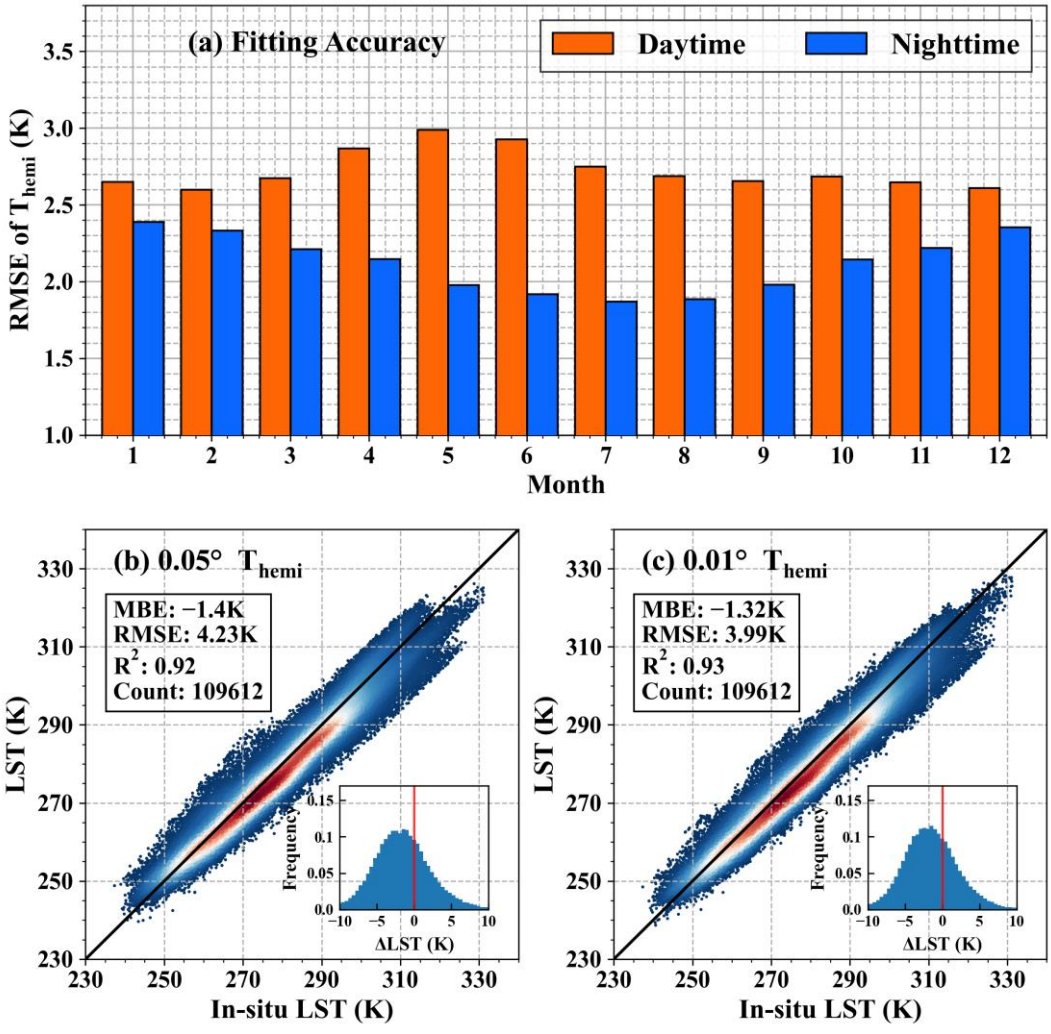


Figure 13: The fitting RMSE over test set for hemispherical LST during the day and night for each month (a). The accuracy of all-weather hemispherical LST products before (b) and after (c) downscaling.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

The spatial distributions of all-weather 0.05° daytime T_{hemi} (Fig. 14a), nighttime T_{hemi} (Fig. 14c), and downscaled 0.01° daytime T_{hemi} (Fig. 14b), nighttime T_{hemi} (Fig. 14d) on 2020/6/24 are shown in Fig. 14. The T_{hemi} is much lower in high-latitude and high-altitude regions for both daytime and nighttime, which demonstrates pronounced patterns of latitudinal and elevational variation of LST. Generally, the spatial distribution of 0.01° T_{hemi} is consistent with that of 0.05° T_{hemi}. The downscaled 0.01° T_{hemi} (Fig. 14b and 14e) reveals richer spatial texture (e.g., in southeastern Tibetan Plateau).

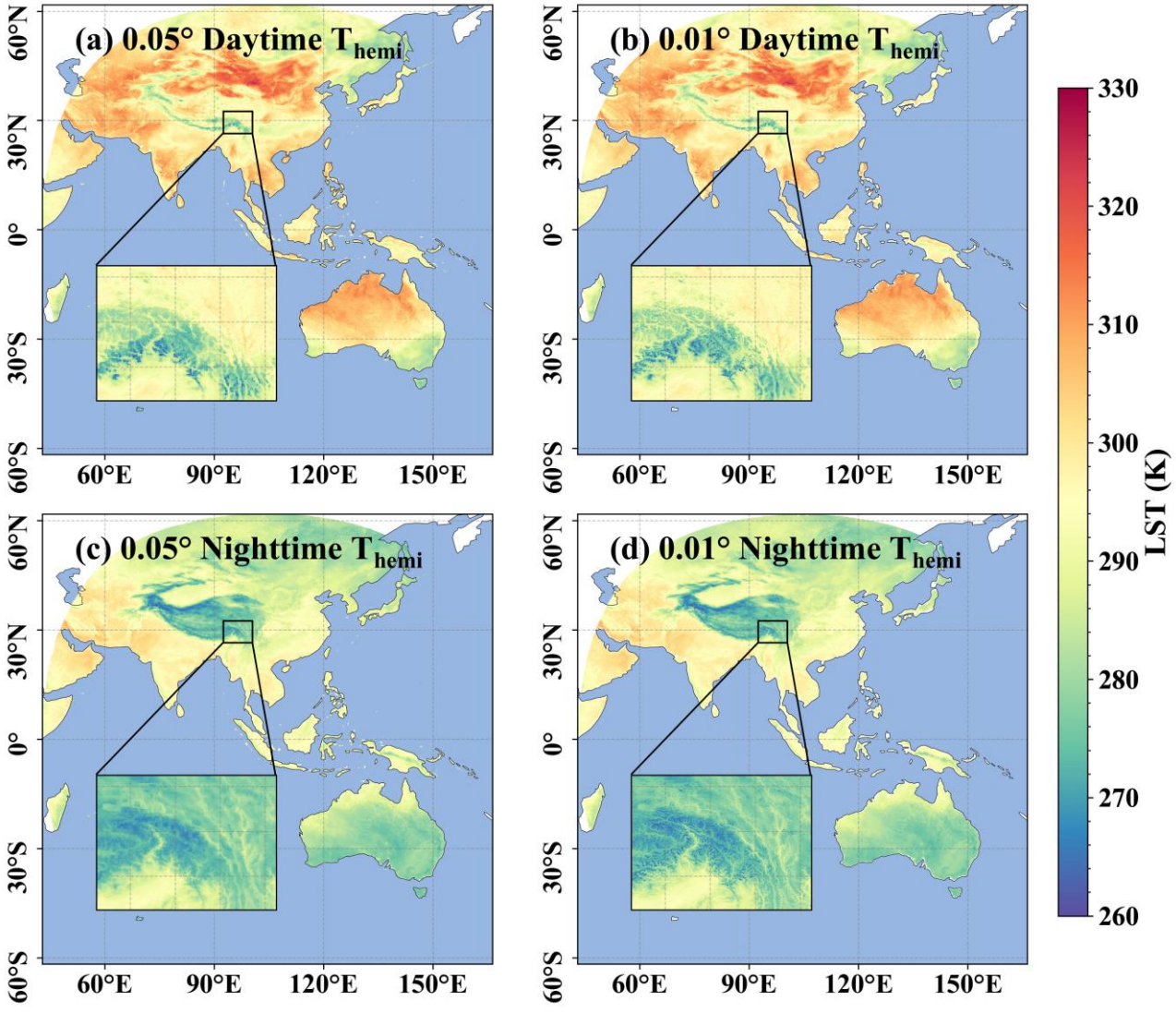


Figure 14: The spatial distribution of (a) 0.05° T_{hemi}, (b) 0.01° T_{hemi} at the UTC time of 4:00 on 2020/6/24, and (c) 0.05° T_{hemi}, (d) 0.01° T_{hemi} at the UTC time of 16:00 on 2020/6/24.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

5 Discussion

5.1 The temporal aggregation method for KDM parameters

In this study, a temporal aggregation method spanning 8 days before and after the target day (a total of 17 days) was employed to aggregate the KDM parameters (i.e., α, β, and W). It not only filtered out extreme values in the TEKDM results but also filled the gaps for pixels where TEKDM could not be solved. The 17-day compositing method was comprehensively compared with three other aggregation approaches: 1) 9-day composition: averaging values from four days before and after the target day (a total of 9 days). 2) Monthly composition (i.e., ~30 days in total): calculating the monthly average of KDM parameters, assuming they remain constant throughout each month. 3) Yearly composition (i.e., ~365 days in total): calculating the yearly average of KDM parameters, assuming they remain constant throughout the whole year. After calibrating the KDM parameters using each approach, FY-4A directional LST could be normalized to T_{nadir}. Then, the T_{nadir} result of 2020 were cross-validated using VNP21A1 near-nadir LST (i.e., VZA < 5°) as did in Fig 5.

Fig. 15a-e shows the cross-validation result between the T_{dir}/T_{nadir} and VNP21A1 near-nadir LST, whereas Fig. 15f shows the fraction of valid pixels (i.e., the ratio of pixels with valid KDM parameters to the total number of land surface pixels) of above temporal aggregation methods. The RMSD (MBD) of the FY-4A official LST is 6.21 K (-4.04 K), while the RMSD (MBD) for the four compositing methods ranges from 3.45 K to 3.79 K (-2.31 K to -2.12 K). The RMSD in ascending order is monthly method< 17-day method< 9-day method< yearly method. In terms of valid pixel fractions, the order is the 9-day method (89.3%)< 17-day method (93.6%)< monthly method (95.9%)< yearly method (99.5%). The significant lower fractions for the 9-day and 17-day methods in 2020/8 (Fig. 15f) are due to the absence of MYD11A1 LST products between 2020/8/16 and 2020/8/31 (induced by a Formatter-Multiplexer Unit/Solid State Recorder (FMU/SSR) error, more details can be found at https://ladsweb.modaps.eosdis.nasa.gov/alerts-and-issues/168889). The yearly method was not adopted since it has the largest RMSD, whereas the 9-day method was not used since it has the lowest valid pixel fraction. The 17-day and monthly methods could achieve comparable and acceptable accuracy. The 17-day method was ultimately selected referenced to the parameter aggregation method in optical domain (Liu et al., 2013).

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

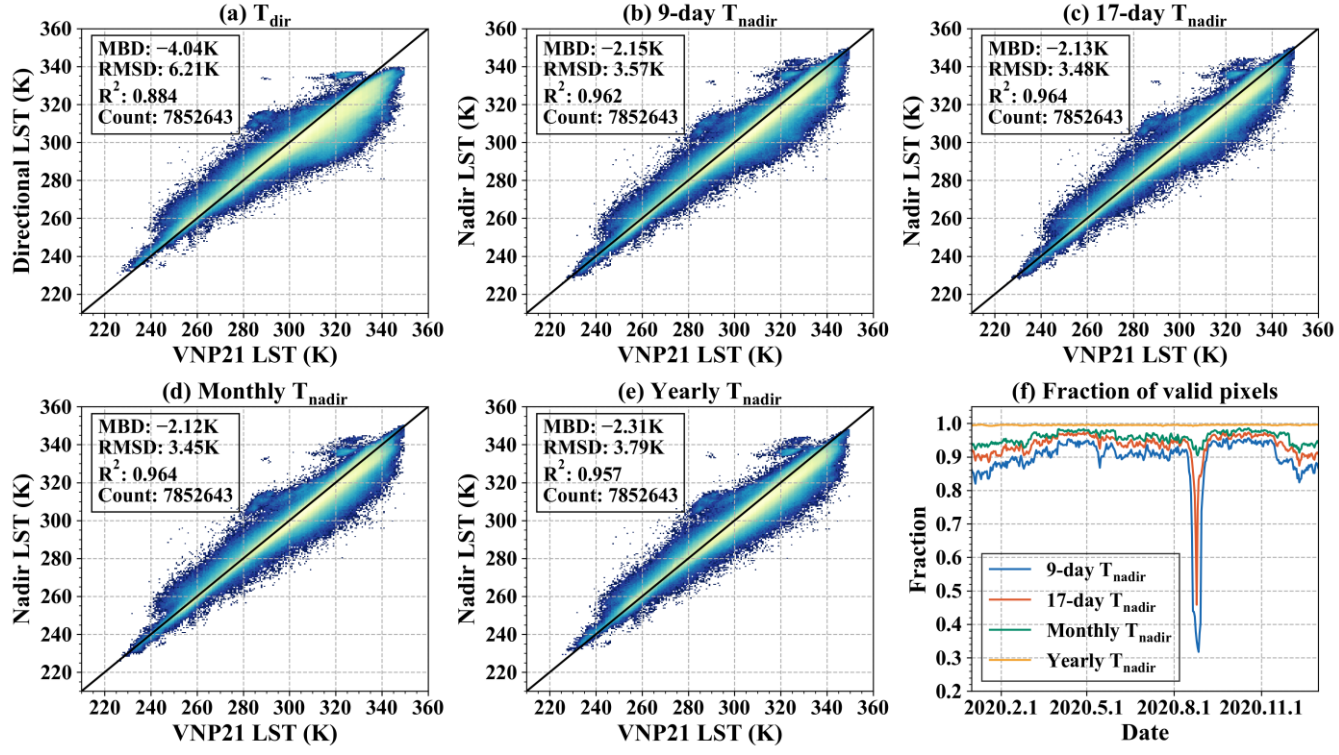


Figure 15: The cross-validation results of (a) FY-4A directional LST, (b) T_{nadir} generated using 9-day composition method, (c) using 17-day composition method, (d) using monthly composition method, (e) using yearly composition method. (f) The fraction of valid pixels for different composition methods.

5.2 Feature Importance for CatBoost models in hypothetical clear-sky LST estimation and LST downscaling

The relative feature importance of CatBoost model was determined based on each feature's contribution to the reduction of the loss function. We employed the logarithmic value of $\log(1 + \text{score}_{\text{norm}})$ to quantify the importance of each model (where $\text{score}_{\text{norm}}$ is the max-min normalized feature scores). The average relative feature importance across all CatBoost models is shown in Fig. 16. In the estimation of hypothetical clear-sky LST (Fig. 16a), $T_{ERA5\text{-}skin}$ exhibited the highest importance. The T_{2m} and R_{in} were also identified as influential predictors, aligning with the findings of Zhang et al. (2024). The θ_s accounted for a significant 9.9% of the importance, reflecting the strong influence of solar radiation on LST. Most other features had relatively low importance (below 5%), and ATC_{d1}, CT, FVC, and LCT contribute less than 1%. In the LST downscaling task from 0.05° to 0.01° using CatBoost model (Fig. 16b), the ATC-simulated LST at the initial time had the highest importance of 30.8%. The air temperature (T_{2m}) and $MAST_{d2}$ also had high importance exceeding 10%. The $YAST_{d2}$, DEM, and dx_{n2} showed moderate importance between 5% and 10%, while the dx_{d2}, $MAST_{n2}$, and $YAST_{n2}$ contributed less than 5%. Eliminating those features with low importance has the potential to enhance computational efficiency and reduce dependency on related auxiliary datasets, which should be carefully assessed in future studies.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

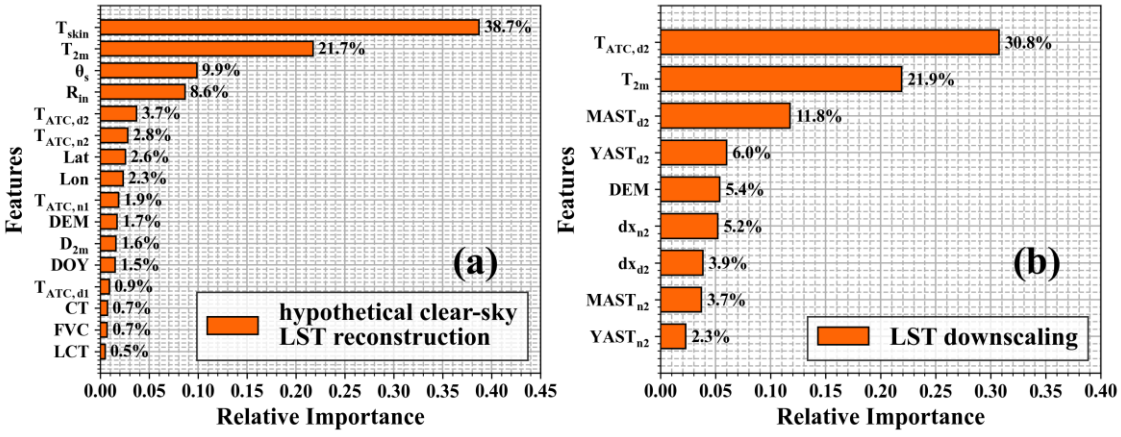


Figure 16: The relative importance for the CatBoost model when (a) estimating the hypothetical clear-sky LST using Eq. 13, and (b) downscaling the 0.05° LST difference using Eq. 18.

5.3 The spatial representativeness of in-situ sites

The T_{nadir} was evaluated by the VNP21A1 near-nadir LST, whereas the 0.05° and 0.01° T_{hemi} were validated using the in-situ measurements. The T-based validation of T_{hemi} requires that in-situ sites exhibit high spatial representativeness. To assess the spatial heterogeneity of 15 sites, the standard deviation (STD) of Landsat 8 LST values in 2020 was calculated for each site at both 0.01° and 0.05° resolution (see Fig. 17a). At the 0.05° scale, STD values range from 1.0 K to 6.4 K, which are significantly higher than those at 0.01° resolution (i.e., 0.5 K – 2.7 K). Almost half of the sites at 0.05° scale have STD values greater than 3 K, indicating a potential influence on validation accuracy due to the limited representativeness. As spatial resolution improved, landscapes within a single pixel tend to become more homogeneous, leading to an decrease in STD. At 0.01° resolution, all site have a STD below 3 K (11 sites among them have STD values below 2 K). Therefore, downscaling LST products could mitigate the impact of site representativeness error on T-based validation result. Fig. 17b shows the relationship between the STD of each site and corresponding RMSE of T_{hemi} at 0.01° resolution. A significant linear relationship is observed, with a Pearson's R of 0.54 and a p-value< 0.05. The slope of the linear regression is 0.73, indicating that a decrease of 1 K in STD of in-situ site (i.e., better spatial representativeness) corresponds to an RMSE decrease of 0.73 K for T_{hemi}.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

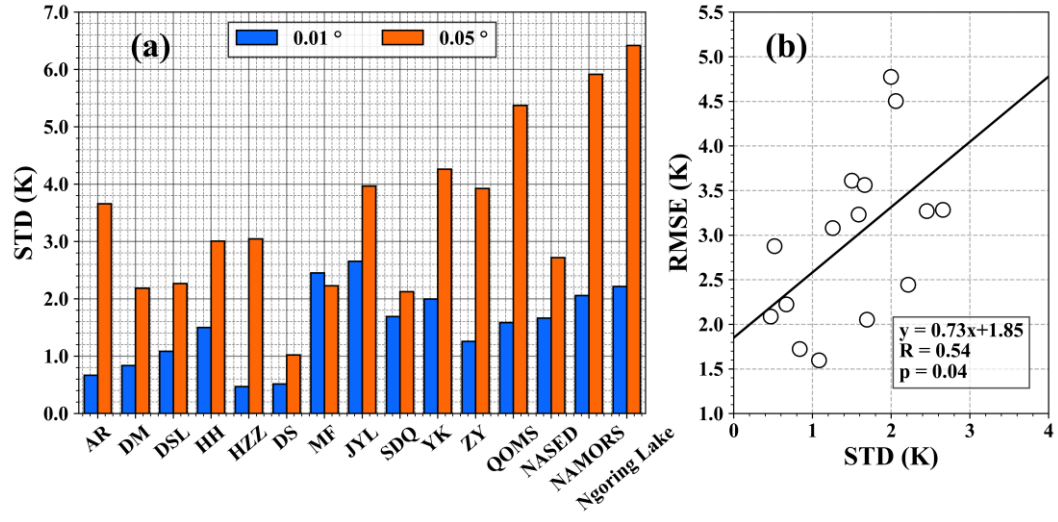


Figure 17: (a) The STD of Landsat 8 LST at 0.01° and 0.05° resolution for all 15 sites in 2020. (b) The relationship between the STD of each site and corresponding RMSE of hemispherical LST at 0.01° resolution.

5.4 The inter-comparison with other all-weather LST products

The inter-comparison between the generated all-weather T_{hemi} product and four existing all-weather LST products against the in-situ LST measurements in the year 2020 is shown in Fig. 18. Although the absolute accuracy is influenced by the spatial representativeness of 15 sites as analyzed in section 5.3, the relative accuracy of different LST products is comparable. Two of the employed datasets are reanalysis products: ERA5-Land LST at 0.1° resolution and China Land Surface Data Assimilation System (CLDAS) LST at 0.0625° resolution. Additionally, two remote sensing products were included: a global all-weather 0.05° LST dataset generated by Jia et al. (2023), and a regional 0.02° LST product for East Asia region developed by Dong et al. (2022). Jia's method first estimates hypothetical clear-sky LST using a Kalman filter with ERA5, Himawari-8 and MODIS LSTs as inputs. An iterative CRF correction is then applied to produce the cloudy-sky LST. Dong's approach first retrieves clear-sky LST from Himawari-8 using a TES algorithm, and then estimates cloudy-sky LST using a multiresolution Kalman filter based on CLDAS LST. For consistency, all four products were resampled to 0.05° resolution to match the 0.05° T_{hemi} for inter-comparison. Under clear-sky conditions (Fig. 18a), ERA5 and CLDAS show similar accuracy with RMSE values of 4.8 K – 5.0 K during the day and 4.6 K – 4.8 K at night. Our T_{hemi} product has the best accuracy with RMSE values of 3.3 K during the day and 2.8 K at night. Dong's LST and Jia's LST have RMSE values between the reanalysis products and the T_{hemi}. Under cloudy-sky conditions (Fig. 18b), ERA5 LST has the lowest accuracy with an RMSE of 6.7 K during the day and an RMSE of 5.0 K at night. CLDAS LST has RMSEs of 5.2 K during daytime and 4.2 K at night, which performs slightly better than Jia's LST. The RMSE of T_{hemi} is similar to Dong's LST during the day (5.1 K) and 0.2 K higher than Dong's LST at night. The superior performance of our T_{hemi} product during both clear-sky and cloudy-sky conditions highlights its good potential in the subsequent applications.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

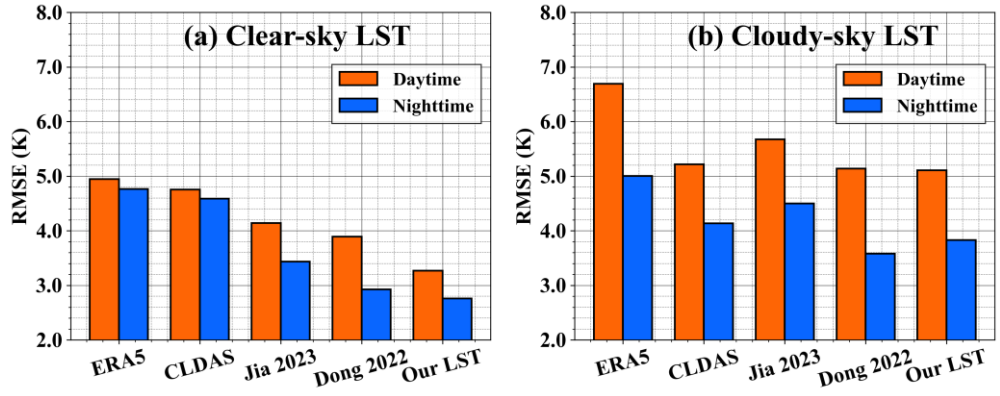


Figure 18: The RMSE comparison of five products during the day and night under (a) clear-sky condition and (b) cloudy-sky condition.

5.5 Limitations and future work

The evaluation results of the generated ANCFDS-LST demonstrate its high accuracy and detailed texture. Several limitations remain to be addressed in future work:

1) The TEKDM used in this study is developed for vegetated canopies, which is dominated by the gap fraction and hot spot effects. However, the applicability of this model over other land covers, such as bare soil, snow, and urban areas requires further exploration and validation (Du et al., 2023, 2024; Wang et al., 2024a; Xiong et al., 2024). Moreover, the reliability of TEKDM in complex terrain also demands in-depth studies (Zhan et al., 2025).
2) This study extensively employs machine learning methods, making it challenging to analyse error propagation processes. 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).
3) The T-based validation in this study is based on 15 in-situ sites located in HRB and Tibetan Plateau, which inevitably introduces uncertainty due to limited spatial representativeness. Future studies could incorporate radiance-based (R-based) validation to mitigate this issue, offering a valuable complement to T-based validation (Li et al., 2021).

6 Data availability

The hourly, 0.01°, all-weather T_{dir}, T_{nadir}, and T_{hemi} product (ANCFDS-LST) from 2018 to 2023 is freely available at https://doi.org/10.11888/RemoteSen.tpdc.303249 (last access: 30 January 2026; Na et al., 2026). Data are stored in GeoTIFF format with three sequential bands (T_{dir}, T_{nadir}, and T_{hemi} respectively), and the units of each file is kelvin. A scale factor of 0.1 and offset of 0 have been applied for value encoding.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

7 Conclusion

In this study, the angular-normalized, cloud-filled, and 0.01°-downscaled LST was generated from 2018 to 2023 on the basis of the official FY-4A LST dataset. Firstly, a TEKDM-based angular normalization method was used to generate nadir and hemispherical LST products. Secondly, an all-weather LST estimation approach was adopted to produce 0.05° cloud-filled directional, nadir, and hemispherical LST. Finally, an IHDA method was employed to enhance the spatial resolution of the LST products from 0.05° to 0.01°. The main conclusions are summarized as follows:

1) The TEKDM model significantly normalized the angular dependence of daytime clear-sky T_{dir}. Taking the near-nadir VNP21A1 LST as reference, the RMSD (MBD) decreased from 6.21 K (-4.04 K) of the T_{dir} to 3.48 K (-2.13 K) of normalized 0.05° T_{nadir}. Taking the in-situ hemispherical LSTs in the Heihe River Basin and the Tibetan Plateau as reference, the RMSE (MBE) decreased from 3.45 K (0.43 K) of the T_{dir} to 3.19 K (-0.36 K) of the normalized 0.05° T_{hemi}.
2) For the all-weather 0.05° T_{hemi}, the T-based validation shows an RMSE (MBE) of 2.99 K (-0.77 K) under clear-sky conditions and 4.56 K (-1.56 K) under cloudy-sky conditions. The generated all-weather T_{hemi} outperformed the existing four all-weather LST products (i.e., ERA5, CLDAS, Jia2023 and Dong2022).
3) The downscaled all-weather LST successfully recovered fine spatial details with a comparable accuracy. The T-based validation shows that the RMSE (MBE) for T_{hemi} slightly decreased from 4.23 K (-1.4 K) at 0.05° resolution to 3.99 K (-1.32 K) at 0.01° resolution.

Despite the generated directional, nadir, and hemispherical LST products have high accuracy and rich spatial texture, their spatial coverage is currently limited to the FY-4A disk. Further work is needed to expand this methodology for producing global-scale LST products, e.g., gathering Himawari, GOES and MSG satellite datasets. In addition, the T-based validation was performed over limited in-situ sites in this study, and a more comprehensive approach such as R-based validation is still lacking. Future research should focus on performing robust validation over more land covers and also over complex terrains.

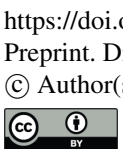
Appendix A

The sampling of label data is essential in the estimation of hypothetical clear-sky LST since the large volume of clear-sky LST observations. In this study, a representative sub-dataset was determined through a two-step sampling approach:

Step 1) Determining the number of samples by evaluating the CatBoost model's performance in fitting the annual maximum LST (i.e., $T_{max} = MAST + YAST$). Due to its highest level of heterogeneity, T_{max} is the most challenging variable for CatBoost to fit. Here, a CatBoost model was trained with the inputs of following features: *Lat*, *Lon*, *LCT*, *CT*, *FVC*, *DEM*, and the annual mean values of R_{in}, T_{2m}, and D_{2m}, as shown in Eq. A1.

$$T_{max} = f_{CatBoost}\left(Lat, Lon, LCT, CT, FVC, DEM, R_{in}, T_{2m}, D_{2m}\right) \quad \text{(A1)}$$

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026



Firstly, 10% of the samples (974,403 samples) were used to tune the hyperparameters of the CatBoost model using a Bayesian optimization method. Another 30% of the samples (2,923,400 samples) were allocated as the test set. The training set includes a total of 5,847,441 samples (i.e., 60% of the samples). Aimed to determine the optimal sample size, directly transverse all the possible sample sizes over training set is impractical. In this study, the number of training samples was gradually increased from 2,000 to 100,000 with an increment of 2,000. As shown in Fig. A1(a), the optimal sample size (i.e., 37,000) was determined based on the point at which the RMSE variation (i.e., |ΔRMSE|) remained less than 0.001 K.

Step 2) determining the exact geolocations of 37,000 points for each year (i.e., 6 years from 2018 to 2023). Here, a conditioned Latin hypercube sampling (cLHS) method was employed, which first divides the features into equal partitions according to their values and then randomly selects sample points from each partition. It has been demonstrated to be both efficient and reliable for large datasets (Minasny and McBratney, 2006; Yang et al., 2020). The input features for this step are identical to those in Eq. A1. The final spatial distribution of 37,000 selected points per year is shown in Fig. A1b-g, indicating a uniform distribution across the FY-4A disk without significant clustering.

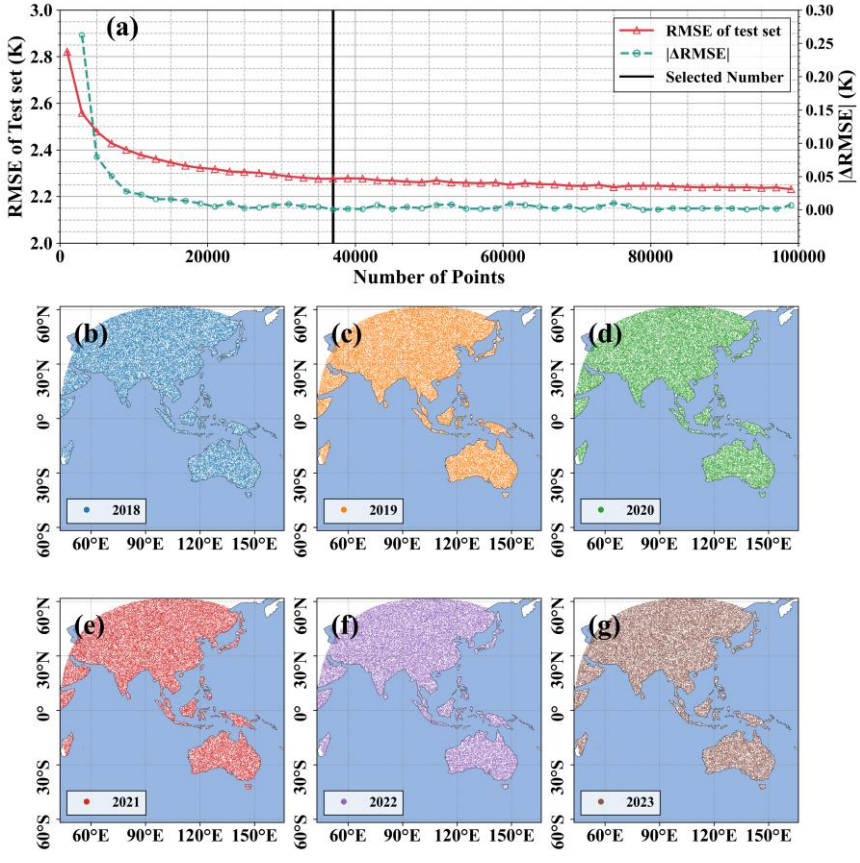


Figure A1: The selection of optimal sample number (a) and the geolocation of 37000 sample points in (b) 2018, (c) 2019, (d) 2020, (e) 2021, (f) 2022, (g) 2023.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Open Access Earth System Science Data Discussions

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 42422107.

References

Agam, N., Kustas, W. P., Anderson, M. C., Li, F., and Neale, C. M. U.: A vegetation index based technique for spatial sharpening of thermal imagery, Remote Sensing of Environment, 107, 545–558, https://doi.org/10.1016/j.rse.2006.10.006, 2007.

Cao, B., Liu, Q., Du, Y., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Trigo, I. F., Zhan, W., Yu, Y., Cheng, J., Jacob, F., Lagouarde, J.-P., Bian, Z., Li, H., Hu, T., and Xiao, Q.: A review of earth surface thermal radiation directionality observing and modeling: Historical development, current status and perspectives, Remote Sensing of Environment, 232, 111304, https://doi.org/10.1016/j.rse.2019.111304, 2019a.

Cao, B., Gastellu-Etchegorry, J.-P., Du, Y., Li, H., Bian, Z., Hu, T., Fan, W., Xiao, Q., and Liu, Q.: Evaluation of Four Kernel-Driven Models in the Thermal Infrared Band, IEEE Trans. Geosci. Remote Sensing, 57, 5456–5475, https://doi.org/10.1109/TGRS.2019.2899600, 2019b.

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., and Xiao, Q.: A general framework of kernel-driven modeling in the thermal infrared domain, Remote Sensing of Environment, 252, 112157, https://doi.org/10.1016/j.rse.2020.112157, 2021.

Chang, Y., Weng, Q., Voogt, J. A., and Xiao, J.: Urban thermal anisotropies by local climate zones: An assessment using multi-angle land surface temperatures from ECOSTRESS, Remote Sensing of Environment, 322, 114705, https://doi.org/10.1016/j.rse.2025.114705, 2025.

Che, T., Li, X., Liu, S., Li, H., Xu, Z., Tan, J., Zhang, Y., Ren, Z., Xiao, L., Deng, J., Jin, R., Ma, M., Wang, J., and Yang, X.: Integrated hydrometeorological, snow and frozen-ground observations in the alpine region of the Heihe River Basin, China, Earth Syst. Sci. Data, 11, 1483–1499, https://doi.org/10.5194/essd-11-1483-2019, 2019.

Coll, C., Galve, J. M., Niclòs, R., Valor, E., and Barberà, M. J.: Angular variations of brightness surface temperatures derived from dual-view measurements of the Advanced Along-Track Scanning Radiometer using a new single band atmospheric correction method, Remote Sensing of Environment, 223, 274–290, https://doi.org/10.1016/j.rse.2019.01.021, 2019.

Dai, Y., Yang, Y., Pan, X., Hu, P., Meng, X., Li, F., and Wang, Z.: Retrieval of Land Surface Temperature From Passive Microwave Observations Using CatBoost-Based Adaptive Feature Selection, IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 18, 4949–4963, https://doi.org/10.1109/JSTARS.2025.3532605, 2025.

Davies, L. and Gather, U.: The Identification of Multiple Outliers, Journal of the American Statistical Association, 88, 782–792, https://doi.org/10.1080/01621459.1993.10476339, 1993.

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

Dong, L., Hu, J., Tang, S., and Min, M.: Field validation of the GLASS land surface broadband emissivity database using pseudo-invariant sand dune sites in northern China, International Journal of Digital Earth, 6, 96–112, https://doi.org/10.1080/17538947.2013.822573, 2013.

Dong, L., Tang, S., Wang, F., Cosh, M., Li, X., and Min, M.: Inversion and Validation of FY-4A Official Land Surface Temperature Product, Remote Sensing, 15, 2437, https://doi.org/10.3390/rs15092437, 2023a.

Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S. K., and Zhao, L.: 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, 2020.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Dong, P., Zhan, W., Wang, C., Jiang, S., Du, H., Liu, Z., Chen, Y., Li, L., Wang, S., and Ji, Y.: 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, 2023b.

Dong, S., Cheng, J., Shi, J., Shi, C., Sun, S., and Liu, W.: A Data Fusion Method for Generating Hourly Seamless Land Surface Temperature from Himawari-8 AHI Data, Remote Sensing, 14, 5170, https://doi.org/10.3390/rs14205170, 2022.

Du, H., Zhan, W., Liu, Z., Scott Krayenhoff, E., Chakraborty, T., Zhao, L., Jiang, L., Dong, P., Li, L., Huang, F., Wang, S., and Xu, Y.: Global mapping of urban thermal anisotropy reveals substantial potential biases for remotely sensed urban climates, Science Bulletin, S2095927323004152, https://doi.org/10.1016/j.scib.2023.06.032, 2023.

Du, H., Zhan, W., Liu, Z., Wang, C., and Huang, F.: 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 Sensing of Environment, 328, 114840, https://doi.org/10.1016/j.rse.2025.114840, 2025.

Du, W., Liu, X., Li, Z.-L., Qin, Z., and Fan, J.: An Improved Integrated Model for Temporal Normalization of Satellite-derived Land Surface Temperature, IEEE Trans. Geosci. Remote Sensing, 1–1, https://doi.org/10.1109/TGRS.2024.3372071, 2024.

Ebrahimy, H. and Azadbakht, M.: Downscaling MODIS land surface temperature over a heterogeneous area: An investigation of machine learning techniques, feature selection, and impacts of mixed pixels, Computers & Geosciences, 124, 93–102, https://doi.org/10.1016/j.cageo.2019.01.004, 2019.

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

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

Ermida, S. L., Trigo, I. F., DaCamara, C. C., and Roujean, J.-L.: Assessing the potential of parametric models to correct directional effects on local to global remotely sensed LST, Remote Sensing of Environment, 209, 410–422, https://doi.org/10.1016/j.rse.2018.02.066, 2018b.

Freitas, S. C., Trigo, I. F., Macedo, J., Barroso, C., Silva, R., and Perdigão, R.: Land surface temperature from multiple geostationary satellites, International Journal of Remote Sensing, 34, 3051–3068, https://doi.org/10.1080/01431161.2012.716925, 2013.

Guillevic, P., Göttsche, F., Hulley, J., Ghent, G., Román, J., Camacho, M., Guillevic, P., Göttsche, F.-M., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J., Remedios, J., Román, M., and Camacho de Coca, F.: Land surface temperature product validation best practice protocol. Version 1.1, 2018.

Hong, F., Zhan, W., Göttsche, F.-M., Lai, J., Liu, Z., Hu, L., Fu, P., Huang, F., Li, J., Li, H., and Wu, H.: A simple yet robust framework to estimate accurate daily mean land surface temperature from thermal observations of tandem polar orbiters, Remote Sensing of Environment, 264, 112612, https://doi.org/10.1016/j.rse.2021.112612, 2021.

Hong, F., Zhan, W., Göttsche, F.-M., Liu, Z., Dong, P., Fu, H., Huang, F., and Zhang, X.: A global dataset of spatiotemporally seamless daily mean land surface temperatures: generation, validation, and analysis, Earth System Science Data, 14, 3091–3113, https://doi.org/10.5194/essd-14-3091-2022, 2022.

Hu, T., Roujean, J.-L., Cao, B., Mallick, K., Boulet, G., Li, H., Xu, Z., Du, Y., and Liu, Q.: Correction for LST directionality impact on the estimation of surface upwelling longwave radiation over vegetated surfaces at the satellite scale, Remote Sensing of Environment, 295, 113649, https://doi.org/10.1016/j.rse.2023.113649, 2023.

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

Jia, A., Liang, S., Wang, D., Ma, L., Wang, Z., and Xu, S.: 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, 2023.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Jia, A., Liang, S., Wang, D., Mallick, K., Zhou, S., Hu, T., and Xu, S.: 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, 2024.

Li, C., Wu, P., Duan, S.-B., Jia, Y., Sun, S., Shi, C., Yin, Z., Li, H., and Shen, H.: 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, 2025.

Li, H., Sun, D., Yu, Y., Wang, H., Liu, Y., Liu, Q., Du, Y., Wang, H., and Cao, B.: Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China, Remote Sensing of Environment, 142, 111–121, https://doi.org/10.1016/j.rse.2013.11.014, 2014.

Li, H., Liu, Q., Yang, Y., Li, R., Wang, H., Cao, B., Bian, Z., Hu, T., Du, Y., and Sun, L.: Comparison of the MuSyQ and MODIS Collection 6 Land Surface Temperature Products Over Barren Surfaces in the Heihe River Basin, China, IEEE Trans. Geosci. Remote Sensing, 57, 8081–8094, https://doi.org/10.1109/TGRS.2019.2918259, 2019.

Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., and Liu, Q.: Temperature-Based and Radiance-Based Validation of the Collection 6 MYD11 and MYD21 Land Surface Temperature Products Over Barren Surfaces in Northwestern China, IEEE Trans. Geosci. Remote Sensing, 59, 1794–1807, https://doi.org/10.1109/TGRS.2020.2998945, 2021.

Li, M., Zhao, W., Yang, Y., Wu, T., and Luo, J.: A solar radiation-based method for generating spatially seamless and temporally consistent land surface temperature, IEEE Trans. Geosci. Remote Sensing, 1–1, https://doi.org/10.1109/TGRS.2024.3392845, 2024a.

Li, R., Li, H., Hu, T., Bian, Z., Liu, F., Cao, B., Du, Y., Sun, L., and Liu, Q.: Land Surface Temperature Retrieval From Sentinel-3A SLSTR Data: Comparison Among Split-Window, Dual-Window, Three-Channel, and Dual-Angle Algorithms, IEEE Trans. Geosci. Remote Sensing, 61, 1–14, https://doi.org/10.1109/TGRS.2023.3288584, 2023a.

Li, X., Liu, F., Ma, C., Hou, J., Zheng, D., Ma, H., Bai, Y., Han, X., Vereecken, H., Yang, K., Duan, Q., and Huang, C.: Land Data Assimilation: Harmonizing Theory and Data in Land Surface Process Studies, Reviews of Geophysics, 62, e2022RG000801, https://doi.org/10.1029/2022RG000801, 2024b.

Li, Y., Wu, H., Chen, H., and Zhu, X.: A Robust Framework for Resolution Enhancement of Land Surface Temperature by Combining Spatial Downscaling and Spatiotemporal Fusion Methods, IEEE Trans. Geosci. Remote Sensing, 61, 1–14, https://doi.org/10.1109/TGRS.2023.3283614, 2023b.

Li, Z., Wu, H., Duan, S., 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., Yan, G., and Zhou, C.: Satellite Remote Sensing of Global Land Surface Temperature: Definition, Methods, Products, and Applications, Reviews of Geophysics, 61, e2022RG000777, https://doi.org/10.1029/2022RG000777, 2023c.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sensing of Environment, 131, 14–37, https://doi.org/10.1016/j.rse.2012.12.008, 2013.

Liang, D., Cao, B., Wang, Q., Jia, K., Qi, J., Zhao, W., and Yan, K.: A New Detection Method for Land Surface Anomalies From the Perspective of Thermal Infrared Remote Sensing, IEEE Trans. Geosci. Remote Sensing, 63, 1–0, https://doi.org/10.1109/TGRS.2025.3605164, 2025.

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, Hydrol. Earth Syst. Sci., 17, 2121–2129, https://doi.org/10.5194/hess-17-2121-2013, 2013.

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

Liu, X., Tang, B.-H., Li, Z.-L., Zhou, C., Wu, W., and Rasmussen, M. O.: An improved method for separating soil and vegetation component temperatures based on diurnal temperature cycle model and spatial correlation, Remote Sensing of Environment, 248, 111979, https://doi.org/10.1016/j.rse.2020.111979, 2020.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Liu, X., Li, Z.-L., Duan, S.-B., Leng, P., and Si, M.: Retrieval of global surface soil and vegetation temperatures based on multisource data fusion, Remote Sensing of Environment, 318, 114564, https://doi.org/10.1016/j.rse.2024.114564, 2025.

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

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

Ma, Y., Xie, Z., Chen, Y., Liu, S., Che, T., Xu, Z., Shang, L., He, X., Meng, X., Ma, W., Xu, B., Zhao, H., Wang, J., Wu, G., and Li, X.: Dataset of spatially extensive long-term quality-assured land–atmosphere interactions over the Tibetan Plateau, Earth Syst. Sci. Data, 16, 3017–3043, https://doi.org/10.5194/essd-16-3017-2024, 2024b.

Michel, J., Hagolle, O., Hook, S. J., Roujean, J.-L., and Gamet, P.: Quantifying Thermal Infra-Red directional anisotropy using Master and Landsat-8 simultaneous acquisitions, Remote Sensing of Environment, 297, 113765, https://doi.org/10.1016/j.rse.2023.113765, 2023.

Minasny, B. and McBratney, A. B.: A conditioned Latin hypercube method for sampling in the presence of ancillary information, Computers & Geosciences, 32, 1378–1388, https://doi.org/10.1016/j.cageo.2005.12.009, 2006.

Minder, J. R., Mote, P. W., and Lundquist, J. D.: Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains, J. Geophys. Res., 115, 2009JD013493, https://doi.org/10.1029/2009JD013493, 2010.

Na, Q., Li, H., Cao, B., Qin, B., Zheng, L., Bian, Z., Du, Y., Xiao, Q., and Liu, Q.: Comprehensive Analysis of Current 1-km Land Surface Temperature Products in Sparsely Vegetated Area: T-Based Evaluation, Thermal Anisotropy, and Joint Application, IEEE Trans. Geosci. Remote Sensing, 62, 1–14, https://doi.org/10.1109/TGRS.2024.3418964, 2024a.

Na, Q., Cao, B., Qin, B., Mo, F., Zheng, L., Du, Y., Li, H., Bian, Z., Xiao, Q., and Liu, Q.: 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, 2024b.

Na, Q., Cao, B., Qin, B., Hu, T., Li, H., Dong, L., Zhang, H., Zhan, W., and Liu, Q.: Angular-normalized, cloud-filled, 0.01°-downscaled land surface temperature from 2018 to 2023 based on official FY-4A dataset, https://doi.org/10.11888/RemoteSen.tpdc.303249, 2026.

Pearson, R. K.: Outliers in process modeling and identification, IEEE Trans. Contr. Syst. Technol., 10, 55–63, https://doi.org/10.1109/87.974338, 2002.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A.: CatBoost: unbiased boosting with categorical features, Advances in neural information processing systems, 31, 2018.

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., and Liu, Q.: A thermal radiation directionality correction method for the surface upward longwave radiation of geostationary satellite based on a time-evolving kernel-driven model, Remote Sensing of Environment, 294, 113599, https://doi.org/10.1016/j.rse.2023.113599, 2023.

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., and Liu, Q.: Angular normalization of GOES-16 and GOES-17 land surface temperature over overlapping region using an extended time-evolving kernel-driven model, Remote Sensing of Environment, 318, 114532, https://doi.org/10.1016/j.rse.2024.114532, 2025.

Quan, J., Zhan, W., Ma, T., Du, Y., Guo, Z., and Qin, B.: An integrated model for generating hourly Landsat-like land surface temperatures over heterogeneous landscapes, Remote Sensing of Environment, 206, 403–423, https://doi.org/10.1016/j.rse.2017.12.003, 2018.

Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, Bulletin of the American Meteorological Society, 94, 1031–1049, https://doi.org/10.1175/BAMS-D-12-00117.1, 2013.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Sun, Y., Deng, K., Ren, K., Liu, J., Deng, C., and Jin, Y.: Deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data: A systematic review, ISPRS Journal of Photogrammetry and Remote Sensing, 208, 14–38, https://doi.org/10.1016/j.isprsjprs.2023.12.011, 2024.

Tang, Y., Wang, Q., Tong, X., and Atkinson, P. M.: Generation of 100-m, Hourly Land Surface Temperature Based on Spatio-Temporal Fusion, IEEE Trans. Geosci. Remote Sensing, 62, 1–16, https://doi.org/10.1109/TGRS.2024.3357735, 2024.

Teng, Y., Ren, H., Zhu, J., Jiang, C., Ye, X., and Zeng, H.: A practical method for angular normalization on land surface temperature using space between thermal radiance and fraction of vegetation cover, Remote Sensing of Environment, 291, 113558, https://doi.org/10.1016/j.rse.2023.113558, 2023.

Trigo, I. F., Dacamara, C. C., Viterbo, P., Roujean, J.-L., Olesen, F., Barroso, C., Camacho-de-Coca, F., Carrer, D., Freitas, S. C., García-Haro, J., Geiger, B., Gellens-Meulenberghs, F., Ghilain, N., Meliá, J., Pessanha, L., Siljamo, N., and Arboleda, A.: The Satellite Application Facility for Land Surface Analysis, International Journal of Remote Sensing, 32, 2725–2744, https://doi.org/10.1080/01431161003743199, 2011.

Vinnikov, K. Y., Yu, Y., Goldberg, M. D., Tarpley, D., Romanov, P., Laszlo, I., and Chen, M.: Angular anisotropy of satellite observations of land surface temperature, Geophys. Res. Lett., 39, L23802, https://doi.org/10.1029/2012GL054059, 2012.

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, IEEE Trans. Geosci. Remote Sensing, 34, 892–905, https://doi.org/10.1109/36.508406, 1996.

Wang, D., Hu, L., Voogt, J. A., Chen, Y., Zhou, J., Chang, G., Quan, J., Zhan, W., and Kang, Z.: Simulation of urban thermal anisotropy at remote sensing pixel scales: Evaluating three schemes using GUTA-T over Toulouse city, Remote Sensing of Environment, 300, 113893, https://doi.org/10.1016/j.rse.2023.113893, 2024a.

Wang, Q., Tang, Y., Tong, X., and Atkinson, P. M.: Filling gaps in cloudy Landsat LST product by spatial-temporal fusion of multi-scale data, Remote Sensing of Environment, 306, 114142, https://doi.org/10.1016/j.rse.2024.114142, 2024b.

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., and Cribb, M.: Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach, Remote Sensing of Environment, 231, 111221, https://doi.org/10.1016/j.rse.2019.111221, 2019.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., and Song, Y.: Improved 1 km resolution $PM_{2.5}$ estimates across China using enhanced space–time extremely randomized trees, Atmos. Chem. Phys., 20, 3273–3289, https://doi.org/10.5194/acp-20-3273-2020, 2020.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sensing of Environment, 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021.

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

Wu, P., Shen, H., Zhang, L., and Göttsche, F.-M.: Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature, Remote Sensing of Environment, 156, 169–181, https://doi.org/10.1016/j.rse.2014.09.013, 2015.

Wu, P., Yin, Z., Zeng, C., Duan, S.-B., Gottsche, F.-M., Ma, X., Li, X., Yang, H., and Shen, H.: 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, 2021.

Xia, H., Chen, Y., Li, Y., and Quan, J.: Combining kernel-driven and fusion-based methods to generate daily high-spatial-resolution land surface temperatures, Remote Sensing of Environment, 224, 259–274, https://doi.org/10.1016/j.rse.2019.02.006, 2019.

Xiong, C., Yuan, L., Wang, Z., and Shi, J.: Modeling the Thermal Infrared Emissivity of Snow and Ice Using Photon Tracking, IEEE Trans. Geosci. Remote Sensing, 1–1, https://doi.org/10.1109/TGRS.2024.3454791, 2024.

https://doi.org/10.5194/essd-2026-74
Preprint. Discussion started: 16 February 2026

Xu, F., Zhu, X., Chen, J., and Zhan, W.: A stepwise unmixing model to address the scale gap issue present in downscaling of geostationary meteorological satellite surface temperature images, Remote Sensing of Environment, 306, 114141, https://doi.org/10.1016/j.rse.2024.114141, 2024.

Yang, L., Li, X., Shi, J., Shen, F., Qi, F., Gao, B., Chen, Z., Zhu, A.-X., and Zhou, C.: Evaluation of conditioned Latin hypercube sampling for soil mapping based on a machine learning method, Geoderma, 369, 114337, https://doi.org/10.1016/j.geoderma.2020.114337, 2020.

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

Yu, Y., Tarpley, D., Privette, J. L., Goldberg, M. D., Rama Varma Raja, M. K., Vinnikov, K. Y., and Xu, H.: Developing Algorithm for Operational GOES-R Land Surface Temperature Product, IEEE Trans. Geosci. Remote Sensing, 47, 936–951, https://doi.org/10.1109/TGRS.2008.2006180, 2009.

Zhan, W., Chen, Y., Zhou, J., Wang, J., Liu, W., Voogt, J., Zhu, X., Quan, J., and Li, J.: Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats, Remote Sensing of Environment, 131, 119–139, https://doi.org/10.1016/j.rse.2012.12.014, 2013.

Zhan, W., Huang, F., Quan, J., Zhu, X., Gao, L., Zhou, J., and Ju, W.: Disaggregation of remotely sensed land surface temperature: A new dynamic methodology, JGR Atmospheres, 121, https://doi.org/10.1002/2016JD024891, 2016.

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

Zhang, H., Tang, B.-H., and Li, Z.-L.: 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, 2024.

Zhang, Q., Wang, N., Cheng, J., and Xu, S.: A Stepwise Downscaling Method for Generating High-Resolution Land Surface Temperature from AMSR-E Data, IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 1–1, https://doi.org/10.1109/JSTARS.2020.3022997, 2020.

Zhang, X., Cao, B., Na, Q., Zheng, L., Yang, Z., Qin, B., Bian, Z., Du, Y., Li, H., Xiao, Q., and Liu, Q.: Determination of the Hemispherical Equivalent Angle for Surface Upward Longwave Radiation, IEEE Geosci. Remote Sensing Lett., 1–1, https://doi.org/10.1109/LGRS.2025.3558980, 2025.

Zheng, L., Cao, B., Na, Q., Qin, B., Bai, J., Du, Y., Li, H., Bian, Z., Xiao, Q., and Liu, Q.: Estimation and Evaluation of 15 Minute, 40 Meter Surface Upward Longwave Radiation Downscaled from the Geostationary FY-4B AGRI, Remote Sensing, 16, 1158, https://doi.org/10.3390/rs16071158, 2024.

Zhou, J., Menenti, M., Jia, L., Gao, B., Zhao, F., Cui, Y., Xiong, X., Liu, X., and Li, D.: A scalable software package for time series reconstruction of remote sensing datasets on the Google Earth Engine platform, International Journal of Digital Earth, 16, 988–1007, https://doi.org/10.1080/17538947.2023.2192004, 2023.

Zhou, Q., Zhu, Z., Xian, G., and Li, C.: A novel regression method for harmonic analysis of time series, ISPRS Journal of Photogrammetry and Remote Sensing, 185, 48–61, https://doi.org/10.1016/j.isprsjprs.2022.01.006, 2022.