Remote Sensing of Environment 303 (2024) 114001

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A mechanism-guided machine learning method for mapping gapless land surface temperature

Jun Ma[a], Huanfeng Shen[a,*], Menghui Jiang[a], Liupeng Lin[a], Chunlei Meng[b], Chao Zeng[a], Huifang Li[a], Penghai Wu[c]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[b] Institute of Urban Meteorology, China Meteorological Administration, Beijing 100089, China
[c] Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, Hefei 230601, China

ARTICLE INFO

Editor: Jing M. Chen

Keywords:
Land surface temperature
Gapless
Mechanistic guidance
Machine learning
Land surface modeling

ABSTRACT

More accurate, spatio-temporally, and physically consistent land surface temperature (LST) estimation has been a main interest in Earth system research. Developing physics-driven mechanism models and data-driven machine learning (ML) models are two major paradigms for gapless LST estimation, which have their respective advantages and disadvantages. In this paper, a mechanism-guided ML model, which combines the strengths in the mechanism model and ML model, is proposed to generate gapless LST with physical meanings and high accuracy. The hybrid model employs ML as the primary architecture, under which the input variable mechanistic guidance is incorporated to enhance the interpretability and extrapolation ability of the model. Specifically, the light gradient-boosting machine (LGBM) model, which only uses remote sensing data as input, serves as the pure ML model. Mechanistic guidance (MG) is coupled by further incorporating key Community Land Model (CLM) forcing data (cause) and CLM simulation data (effect) as inputs into the LGBM model. This integration forms the MG-LGBM model, which incorporates surface energy balance (SEB) guidance underlying the data in CLM-LST modeling within a biophysical framework. Results indicate that, MG-LGBM model shows a good accuracy for the sample-based validation, with a root-mean-square error of 1.23–2.03 K, and a Pearson correlation coefficient of 0.99. Validation with four independent ground measurements shows that MG-LGBM can generate clear-sky LST that is comparable to the original Moderate Resolution Imaging Spectroradiometer- (MODIS) LST under fully clear-sky conditions and can correct for the likely cloud-contaminated LST pixels. The generated LST also presents a high accuracy (RMSE $= 2.91$–3.66 K and $R = 0.97$–0.98) under cloudy-sky conditions. Compared with a pure mechanistic method and pure ML methods, the MG-LGBM model improves the prediction accuracy and mechanistic interpretability of LST. It also demonstrates a good extrapolation ability in the regions without valid samples, suggesting that the predictions of MG-LGBM model not only exhibit low errors on the training dataset but also align consistently with the known mechanistic laws in the unlabeled set. Compared with other popular ML methods and sophisticated gapless products, the MG-LGBM model delivers a superior validation accuracy and image quality. The proposed method represents an innovative way to map accurate and mechanistically interpretable gapless LST, and could provide insights to accelerate knowledge discovery in land surface processes and data mining in geographical parameter estimation.

1. Introduction

Land surface temperature (LST) is a crucial parameter in land-atmosphere interactions, reflecting the surface energy balance (SEB) and fluxes exchange and is widely used in multi-disciplinary research, such as climatology (Hansen et al., 2010), urbanology (Shen et al., 2016), agriculture (Karnieli et al., 2010), ecology (Connors et al., 2013) and hydrology (Anderson et al., 2012). Accordingly, LST has been identified as one of the most essential climate variables (ECVs) by the Global Climate Observing System (GCOS). For LST acquisition, thermal infrared (TIR)-based remote sensing has attracted increasing attention due to the relatively fine spatial resolution, high accuracy and data availability across the globe (Li et al., 2013b). Over the past decades, a wealth of LST products based on TIR algorithms have been developed

---

* Corresponding author.
E-mail address: shenhf@whu.edu.cn (H. Shen).

https://doi.org/10.1016/j.rse.2024.114001
Received 25 June 2023; Received in revised form 21 December 2023; Accepted 13 January 2024

and used, such as Moderate Resolution Imaging Spectroradiometer- (MODIS-), Spinning Enhanced Visible and Infrared Imager- (SEVIRI-), Advanced Very High-resolution Radiometer- (AVHRR-) and Visible Infrared Imaging Radiometer- (VIIRS-) LST. However, cloud contamination often leads to large gaps in TIR LST. For example, the MODIS LST covers <40% of the globe, which severely hinders the potential applications of the data (Duan et al., 2017). Therefore, filling LST gaps caused by cloud contamination is an urgent priority in relevant studies.

A series of approaches have been developed to fill the LST gaps and generate gapless LST products (Shen et al., 2015; Wu et al., 2021). The early mainstream methods typically utilized the spatio-temporal information of TIR LST itself, considering LST temporal variation laws, such as the diurnal temperature cycle (DTC) (Xu and Shen, 2013) and annual temperature cycle (ATC) (Xia et al., 2021), or spatial neighboring laws, such as kriging (Ke et al., 2013) and spline function (Neteler, 2010), or both of them (Kilibarda et al., 2014; Li et al., 2018; Weiss et al., 2014). However, the accuracy of these methods is significantly affected by the accessibility of neighboring clear pixels in space and time, with an unsatisfactory performance found in large-scale or long time-series data missing cases (Fu and Weng, 2016). Furthermore, this type of method cannot reflect the cloud effects, thus leading to a hypothetical clear-sky LST (Hong et al., 2021).

To obtain the real cloudy-sky LST, methods considering the difference between LST in clear and cloudy conditions were proposed. These methods normally take proxy data for clouds from other sources as the additional information, such as meteorological data, passive microwave (PMW)-based data, and land surface model (LSM)-simulated or assimilation data. By calculating and adding the cloud effect estimated from meteorological data, SEB-based methods have been proposed to correct the hypothetical clear-sky LST (Jia et al., 2021; Jin, 2000; Zeng et al., 2018). Nevertheless, some meteorological data, such as ground-based air temperature and radiation data, are difficult to obtain, thus impeding the applications of SEB-based methods in poorly gauged regions (Lu et al., 2011; Yu et al., 2014b). Due to the ability to penetrate clouds, PMW LST is another commonly used data source for mapping real and gapless LST (Duan et al., 2017; Wu et al., 2022; Xu and Cheng, 2021). However, the retrieval of high-quality PMW LST remains challenging due to the surface penetration, swath gaps and the relatively coarse spatial resolution, which can degrade the quality of the reconstructed LST (Zhang et al., 2020). LSMs can achieve the spatio-temporally continuous simulation of land surface parameters, such as LST, soil moisture (SM), and surface energy fluxes, and are regarded as a fundamental methodology in Earth system science (ESS) (Fisher and Koven, 2020). In addition, LSMs are rooted in scientific theory based on physical mechanisms and parameterization schemes, thereby providing more realistic LST estimates with physical meanings by learning explainable relationships between input and output variables. Some studies have incorporated LSM-simulated data as the complementary information to fill the gaps of TIR LST, which is an approach that have received much attention in recent years (Long et al., 2020; Ma et al., 2022; Zhang et al., 2021). However, LSMs mostly rely on the available mechanisms and lack capability to directly derive insights from data (Karpatne et al., 2017). For example, many LSMs use simplified mechanisms for representing the complex mechanistic processes that are not fully understood, thus resulting in a relatively low accuracy of LSM simulation data, which can introduce uncertainties in the reconstruction process.

Motivated by the remarkable data mining and non-linear representation capabilities of ML architectures, many achievements have been made in LST reconstruction studies. The reconstruction is usually implemented by establishing the non-linear function between TIR LST and gapless LST (e.g., PMW LST) or its spatio-temporal descriptors, such as the normalized difference vegetation index (NDVI), digital elevation model (DEM), day of year (DOY), and albedo in clear conditions, and applying the relationship to obtain the cloudy LST (Buo et al., 2021; Tan et al., 2021b; Wu et al., 2022; Wu et al., 2019; Xiao et al., 2023). Thanks to the emergence and introduction of proxy data for clouds, ML methods have advanced the ability to reconstruct the actual cloudy-sky LST (Cho et al., 2022; Fu et al., 2019; Shwetha and Kumar, 2016; Zhao and Duan, 2020). Nevertheless, the mainstream methods incorporate independent proxy data (e.g., solar radiation data or LSM simulated LST) which are only correlated with cloudy LST, but provide no prior assumptions or mechanistic understanding of the processes. As is well known, ML methods are empirically based and rely heavily on massive training data, making it difficult to transfer the model to other data-sparse regions (Lin et al., 2023). Furthermore, “correlation does not imply causation” (Aldrich, 1995; Altman and Krzywinski, 2015). The LST retrieval mechanisms and the process knowledge (such as conservation of energy) behind LST reasoning remain unclear, which may result in spurious predictions and extrapolation problems (Read et al., 2019). Therefore, it is necessary to incorporate causal data involved in the LST reasoning process as input, thereby coupling such mechanisms into the ML model to enhance its interpretability and transferability.

Coupling mechanistic knowledge into ML is one of the current research hotspots, and has been successfully applied in estimating various land surface parameters (De Bézenac et al., 2019; Karniadakis et al., 2021; Karpatne et al., 2017; Koppa et al., 2022; Shang et al., 2023; Shen and Zhang, 2023; Wang et al., 2021; Wang et al., 2023). Toward LST reconstruction, to the best of our knowledge, few studies have been conducted on coupling the mechanism and learning. Regarding the above issues, this study takes into consideration the rationalism and empiricism that inform a mechanism-guided ML method for mapping gapless LST. Specifically, an advanced LSM, i.e., the Community Land Model (CLM), which uses meteorological forcing data to calculate the SEB was first used to produce model-based estimates. The complete process of CLM-LST modeling was deduced to identify the optimal combination of mechanistic variables with causal relationships as inputs for the ML model. By incorporating key CLM forcing data (cause) and CLM simulation data (effect), the mechanistic guidance based on the SEB underlying the data were integrated into the ML model. The generated LST product under all-weather conditions tends to be mechanistically realistic while maintaining a high accuracy.

2. Study area and data

2.1. Study area

Fig. 1 shows the study area, which covers most of the middle and upper reaches of the Heihe River Basin (HRB-MU), and is located within 37.5°N–39.5°N and 99°E–101°E in arid Northwest China. Featuring glaciers, alpine meadows, grassland and forest in the upstream and dominated by irrigated crops and desert in the middle part of the HRB, the terrain of the HRB-MU is complex, with the elevation ranging from 1325 to 5040 m. The HRB-MU has a continental climate with mean annual precipitation of ~400 mm and mean annual temperature of ~275 K (Tan et al., 2021a). The in-situ measurements used in this study were derived from four weather stations in a well-known watershed observatory network, namely the Watershed Allied Telemetry Experimental Research (WATER) network (Li et al., 2009). The observatory network in the HRB-MU provides an ideal testbed for the validation of the proposed method in estimating gapless LSTs under various topographic conditions (Liu et al., 2018).

2.2. Data

The data used in the study consisted of: 1) satellite data for model training and validation; 2) meteorological forcing data for model training, validation and CLM driving; 3) CLM simulations for model training and validation; and 4) in-situ LST observations for model validation. Before the model implementation, the satellite data, meteorological forcing data and CLM simulation data from 2008 to 2011 were all resampled to a $0.01^\circ \times 0.01^\circ$ spatial resolution using a nearest neighbor

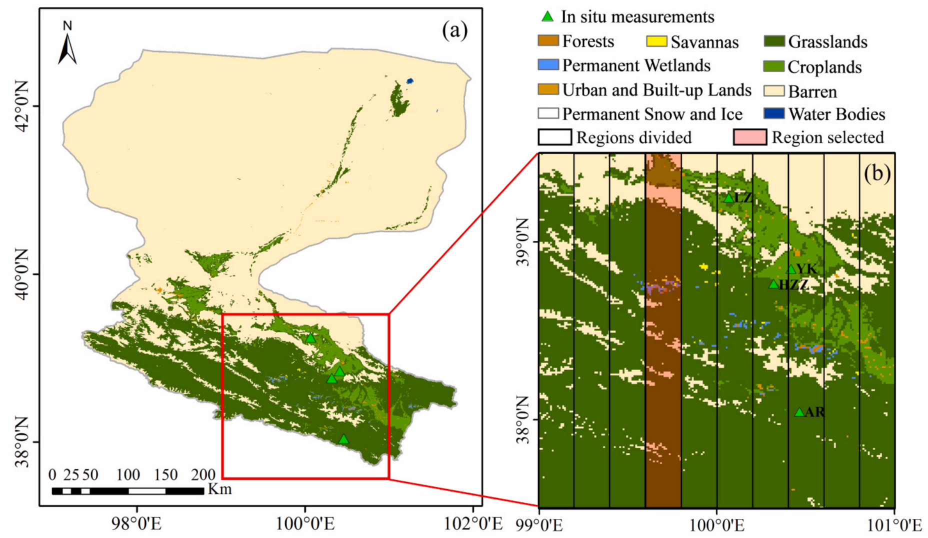


Fig. 1. Location of (a) the Heihe River Basin (HRB) and (b) the study area (HRB-MU). Land-cover types for 2010 are from the MODIS land-cover type product (MCD12Q1) with a 500 m spatial resolution. The rectangles filled with null values represent the divided regions. The rectangle filled with light red represent the selected region for testing, while the other rectangles represent the regions for training in the space-based validation (Section 3.3). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interpolation and re-projected to the WGS84 coordinate system. Table 1 summarizes the basic information about the multi-source data used in this study.

2.2.1. *Satellite data*

The satellite data used in this study from 2008 to 2011 are listed in Table 1. The 1-km MODIS daily LST/emissivity product (MYD11A1) in Collection 6 was selected as the model label data. The MYD11A1 v6 product was retrieved from Aqua MODIS using the generalized split-window algorithm (Wan and Dozier, 1996), with an approximate overpass time of 1:30 p.m. (local solar time) in ascending orbit and 1:30 a.m. in descending orbit. Some studies have indicated that the MODIS LST product has an accuracy of within 1 K in homogeneous areas (Wan, 2014). The quality control (QC) flags with “LST produced, good quality (QC = 0)” and “average LST error <= 1 K" were used to identify the clear-sky MODIS LST data.

Four surface variables that are highly correlated with surface thermal properties were chosen as the LST predictors. The 16-day 1-km NDVI data were retrieved from the MODIS vegetation index product (MOD13A2). The Global LAnd Surface Satellite (GLASS) (Liang et al., 2021) black-sky surface albedo product with visible (B_VIS) and near-infrared (B_NIR) spectral ranges was obtained from GLASS02A06. A normalized difference snow index (NDSI) product produced by a Spatio-Temporal Adaptive fusion method with error correction (STAR-NDSI) (Jing et al., 2022) was adopted in this study. The DEM and latitude (LAT) data were used to account for the LST terrain effect and temperature gradients (Minder et al., 2010; Zhao et al., 2019). We acquired the Shuttle Radar Topography Mission DEM (SRTM DEM) data with a 90-m spatial resolution from the United States Geological Survey (USGS). For the temporal feature, the DOY was selected.

Table 1
Summary of the satellite data, meteorological forcing data, model simulation data and in-situ measurements used in this study.[a]

Short name	Variables	Spatial resolution	Temporal resolution	Usage
MYD11A1	LST	1 km	Daily	Model label
MOD13A2	NDVI	1 km	16-day	Model input
GLASS02A06	Albedo	1 km	8-day	Model input
STAR NDSI	NDSI	500 m	Daily	Model input
SRTM DEM	DEM, latitude	90 m	–	Model input
CLDAS forcing	TMP, RHU, PRS, SRA, WIN, PRE	0.0625°	Hourly	CLM driving Model input,
CLM simulated	CLM-LST, CLM-SM	0.05°	Hourly	Model input
WATER	ULW, DLW	Point	30-min	Model validation

[a] The definitions of the variables are in the text and the download links are attached in the Acknowledgements.

The temporal resolutions of the GLASS product and MODIS vegetation index product vary from 8 days to 16 days. To obtain these products at a daily time scale, different interpolation methods were employed. The nearest neighbor interpolation approach was used for albedo due to its relative invariance within 8 days. As for the NDVI, the linear interpolation approach was applied, as it can represent the true trend of this variable. After the preprocessing, all the auxiliary variables were gapless with a 1-km and daily spatio-temporal resolution.

2.2.2. *Meteorological forcing data and surface data*

The China Meteorological Administration Land Data Assimilation System version 2.0 (CLDAS v2.0) (Shi et al., 2011) was used as the forcing data for driving the Community Land Model Version 5.0 (CLM5.0) and the input data for the ML model. The CLDAS v2.0 product was developed and maintained by the China Meteorological Administration (CMA), covering East Asia (0–65°N, 60–160°E) and spanning from 2008 to 2018. It was produced by merging reanalysis data, satellite-based products, and in-situ measurements from >2400 national automatic stations of the CMA, with a spatial resolution of 0.0625° × 0.0625° and a temporal resolution of 1 h (Liu et al., 2019). The dataset consists of six meteorological forcing variables, i.e., 2-m air temperature (TMP), 2-m specific humidity (RHU), 10-m wind speed (WIN), precipitation rate (PRE), air pressure (PRS), and downward shortwave radiation (SRA). Compared with the Global Land Data Assimilation System (GLDAS) forcing data, the CLDAS product has been proven to be more accurate in China due to the integration of more site observations (Yang et al., 2017). The soil properties data, i.e., percent sand, percent clay and organic matter density with a 30 arc-second resolution, were derived from the China Dataset of Soil Properties for Land Surface Modeling (Shangguan et al., 2013). The static land-cover data, including percent crop/lake/wetland/glacier/urban/natural vegetation and different percent plant functional types (PFTs) were collected from the MODIS land-cover type product (MCD12Q1) for 2010. The other surface data, such as elevation, slope, and the monthly leaf and stem area index, were obtained from the CLM surface data pool (Oleson et al., 2010).

The above-mentioned high-quality forcing and surface data were introduced to replace the default datasets in the CLM5.0 data pool, to improve the model performance. After the CLM5.0 build and spin-up, a four-year numerical simulation was conducted to produce model-based LST and SM estimates with a 0.05° and 1-h spatio-temporal resolution. To match the satellite data, the CLDAS forcing data and CLM simulation data for ML model training were temporally interpolated to the value at the MYD11A1 view time by a cubic spline interpolation method.

2.2.3. In-situ data

To access the accuracy of the reconstructed LST, four in-situ longwave radiation measurements with different land-cover types were collected from WATER, who provided continuous and high-quality in-situ measurements for LST assessment studies (Duan et al., 2017; Wu et al., 2022; Yu et al., 2014a). The A'rou freeze station (AR), Huazhaizi desert station (HZZ), Yingke oasis station (YK), and Linze grassland station (LZ) are equipped with Kipp & Zonen (CNR1/CNR4) or CAMPBELL (CG3) net radiometers for measuring upwelling (ULW) and downwelling longwave radiation (DLW) every 30 min. Fig. 1 shows the locations of the stations, and Table 2 lists the basic information about the four stations.

The in-situ LSTs were retrieved using Stefan-Boltzmann law as follows (Wang and Liang, 2009):

$$T_s = \left[\frac{F^{\uparrow} - (1-\varepsilon_b)F_{\downarrow}}{\varepsilon_b \bullet \sigma}\right]^{1/4} \tag{1}$$

where T_s is the surface skin temperature or LST; $F^{\uparrow}$ and $F_{\downarrow}$ denote the surface upwelling and atmospheric downwelling longwave radiation, respectively; σ is the Stefan-Boltzmann constant (5.67×10^{-8} W m^{-2} K^{-4}); and ε_b is broadband emissivity (BBE), which was acquired from a GLASS BBE product representing the emissivity value at 8–13.5 μm (Cheng et al., 2015). Finally, the "3σ (standard deviations) edit rule" method (Pearson, 2002) was utilized to remove outliers due to cloud contamination.

3. Methodology

We propose a mechanism-guided ML model for mapping gapless LST. The hybrid model takes ML as the main architecture, under which the underlying mechanistic laws (e.g., the SEB) between the key CLM forcing data (SRA, TMP, RHU, PRS) and CLM simulation data (CLM-LST, CLM-SM) in the CLM-LST modeling are incorporated and guide the ML model. The LST estimations of the hybrid model were compared with those of a pure mechanistic method, pure ML methods, and other advanced ML methods. The well-trained model was applied for the HRB-MU from 2008 to 2011. The generated LST data were validated against in-situ LST and compared with other gapless LST data. The details of the proposed method are provided below.

Table 2
The basic information about the four weather stations used in this study.

Site	Location	Elevation (m)	Land-cover	Period
AR	100.4647°E, 38.0444°N	3033	Alpine meadow	2008/01–2011/12
HZZ	100.3201°E, 38.7659°N	1731	Desert steppe	2008/01–2011/12
YK	100.4167°E, 38.85°N	1519	Cropland	2008/01–2011/12
LZ	100.0667°E, 39.25°N	1394	Grassland	2008/01–2008/10

3.1. Pure ML model

ML has gained increasing popularity in gapless LST estimation (Li et al., 2021; Zhang et al., 2020; Zhao and Duan, 2020). To date, the ML has usually been implemented by establishing the relationship between the clear-sky LST and auxiliary variables. The established relationship was then subsequently applied to the gapless auxiliary variables for mapping all-weather LST. The early studies normally used gapless data (e.g., remote sensing data) as the explanatory variables. The empirical model can be constructed as follows:

$$LST = \mathbf{ML}(\mathrm{DEM}, \mathrm{LAT}, \mathrm{NDVI}, \mathrm{NDSI}, \mathrm{B\_VIS}, \mathrm{B\_NIR}, \mathrm{DOY}) \tag{2}$$

where ***ML*** is the pure ML model without mechanistic constraints; *LST* is the MODIS LST with the highest quality; *DEM*, *LAT*, *NDVI*, *NDSI*, *B_VIS*, and *B_NIR* are the remote sensing data, which represent the surface thermal properties and terrain effect of LST; and the day of year (*DOY*) represents the temporal characterization.

In this study, the light gradient-boosting machine (LGBM) model was used as the learner. Fig. 2 shows a schematic diagram of the LGBM model. As the successor of the extreme gradient boosting (XGBoost) model, the LGBM model is a newly developed ML algorithm based on the gradient-boosting decision tree (GBDT) (Ke et al., 2017). Compared with conventional deep learning methods, such as a deep belief network (DBN) and generalized regression neural network (GRNN), the LGBM model is advantageous in dealing with massive data and countering overfitting problems. A second-order approximation is used to minimize the objective function in the LGBM model, which is formulated as follows:

$$L^{(t)} = \sum_{i=1}^{n}\left[l\left(y_i, \widehat{y}^{(t-1)}\right) + \partial_{\hat{y}(t-1)} l\left(y_i, \widehat{y}^{(t-1)}\right) f_t(x_i) + \frac{1}{2}\partial_{\hat{y}}^2 (t-1) l \left(y_i, \widehat{y}^{(t-1)}\right) f_t^2(x_i)\right] + \Omega(f_t) \tag{3}$$

where $L^{(t)}$ is the objective function of the $t-th$ iteration solution; x_i is the sample; n denotes the number of samples; l denotes the loss function used to measure the difference between the actual value y_i and the predicted value $\widehat{y}$, which is the root-mean-square error (RMSE) in this case; ∂ and ∂^2 denote the first- and second-order gradients of l, respectively; f_t is an independent regression tree and $f_t(x)$ is the corresponding increment; and $\Omega(f)$ denotes the regularization term, which is defined as follows:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{4}$$

where T denotes the number of leaf nodes; w denotes the leaf weights; γ and λ are hyper-parameters, and larger γ and λ indicate that the model prefers simpler tree structures, which are adopted from a leaf-wise tree growth strategy(Chen and Guestrin, 2016).

The effectiveness of the LGBM model mainly stems from Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) (Fig. 2). GOSS is a novel sampling technique that keeps a good balance between reducing the number of data instances and maintaining a high accuracy. Firstly, it sorts the data instances based on their gradient absolute values and selects the top m % instances. Secondly, it randomly samples n % instances from the remaining data. Finally, the sample instances with small gradients are amplified by a $\frac{1-\mathrm{m}\%}{\mathrm{n}\%}$ weight to ensure that more emphasis is placed on the under-trained instances without changing the original data distribution. EFB was designed to safely filter the features (such as NDVI, NDSI, B_VIS, B_NIR, etc.) based on the sparsity of high-dimensional data, which is founded on a histogram-based algorithm. It first clusters the data to find those sparse but frequently co-occurring features, which are usually mutually exclusive. It then bundles these exclusive features into a single feature to avoid unnecessary computation for zero feature values. This bundling

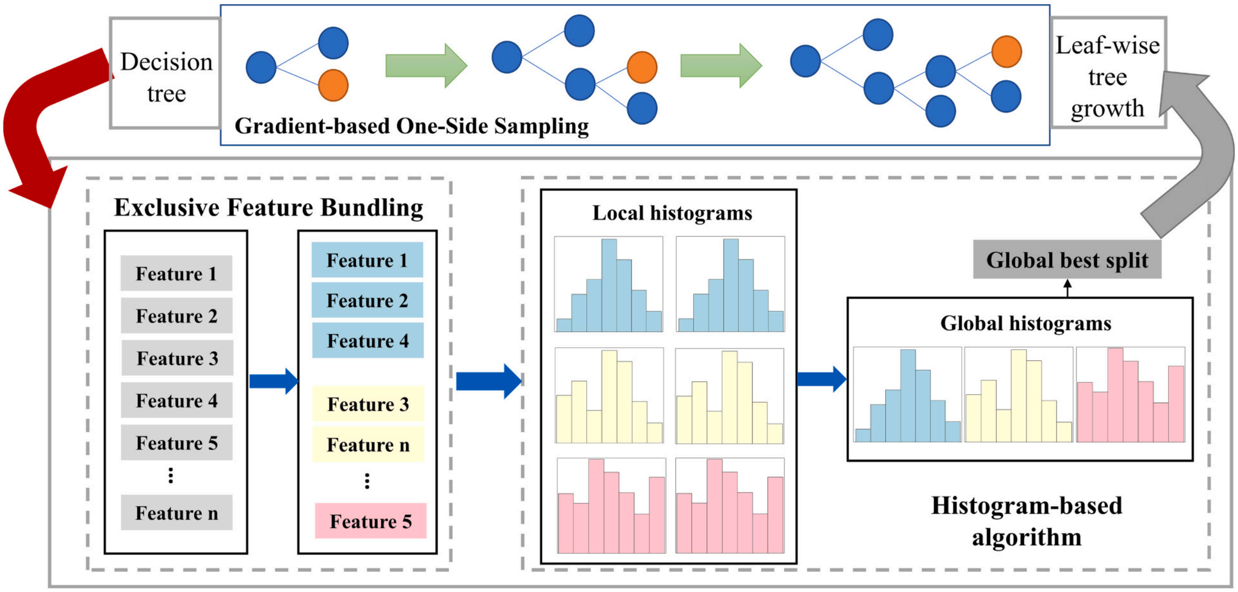


Fig. 2. Schematic of the light gradient-boosting machine model.

process can reduce the dimension of the samples, thereby reducing memory usage, preventing overfitting and retaining accuracy.

3.2. *Mechanism-guided ML model*

Despite filling the gaps of the MODIS LST, the obtained cloudy LST based on pure ML (Eq. (2)) is far away from realistic due to the unavailability of data that can distinguish between clear-sky and cloudy-sky conditions. Furthermore, pure ML methods can only learn the relationship between discrete LST samples and the explanatory data. The causal relationship and the underling mechanistic process between LST retrieval input and output remain unclear. As a result, the obtained LST lacks mechanistic interpretability, which is usually ignored in LST reconstruction studies using ML methods. In our previous work, three guidance methods were proposed to integrate the mechanistic knowledge into the ML model, i.e., input variable guidance, objective function guidance, and model structure guidance (Shen and Zhang, 2023). In this case, the input variable mechanistic guidance is employed. In this section, we first describe the overall framework of input variable mechanism-guided ML model, termed mechanism-guided LGBM (MG-LGBM). Then, the LST retrieval mechanisms and the involved process knowledge in CLM5.0 are deduced in mechanism-based equations to help understand the process of coupling.

3.2.1. *Overall framework*

Consider a predictive learning problem in which we have a set of input forcing data, CF, that are mechanistically related to a target variable of interest, LST. A common approach is to train a data driven model, such as an LGBM model, $f_{LG}: CF \rightarrow LST$, which can be used to generate estimates of the target variable, $\widehat{LST}$. Alternatively, we can also use a mechanistic model, e.g., $f_{CLM}: CF \rightarrow LST$, to simulate the value of the target variable, LST_{CLM}, given its mechanistic relationships with F. However, LST_{CLM} may provide an incomplete representation of LST due to the approximations of parameterized forms in f_{CLM} , thus resulting in discrepancies between model simulations and observations (Karpatne et al., 2017). Therefore, the primary objective of MG-LGBM is to combine f_{CLM} and f_{LG} so as to overcome their respective limitations and leverage information in both mechanism and data. One simple yet effective way for combining f_{CLM} and f_{LG} is to use the simulated outputs of the mechanistic model, LST_{CLM}, along with the forcing data, F, as joint inputs for the ML model. The relationship is improved as:

$$LST = \mathbf{MG}-\mathbf{LGBM}(\mathbf{CF},\mathbf{CS},\mathrm{DEM},\mathrm{LAT},\mathrm{NDVI},\mathrm{NDSI},\mathrm{B\_VIS},\mathrm{B\_NIR},\mathrm{DOY}) \tag{5}$$

where $\boldsymbol{CF} = \{SRA, TMP, RHU, PRS\}$ denotes the four sets of CLM forcing data that are mechanistically dominant in CLM-LST simulation; and $\boldsymbol{CS} = \{CLM-LST, CLM-SM\}$ represents the two sets of CLM simulation data that correspond to the CLM forcing data. Both ***CF*** and ***CS*** can characterize the real thermal state under clouds. The downward shortwave radiation SRA is only used in the daytime model. ***MG-LGBM*** is the mechanism-guided ML model obtained by imposing SEB guidance in mechanism-based equations to the LGBM model (Fig. 3). Mentioning that if the mechanistic model and its output is accurate enough, ***MG-LGBM*** can learn to predict $\widehat{LST}$ with the same level of accuracy as LST_{CLM}. However, if there are systematic bias in LST_{CLM}, then ***MG-LGBM*** can learn to mitigate these differences by extracting latent features from the input forcing data space, thereby narrowing the knowledge gap.

3.2.2. *The CLM5.0 model and its configuration*

In this study, the Community Land Model Version 5.0 (CLM5.0) was utilized to produce the model-based LST and SM estimates. CLM5.0 is the land component coupling in the Community Earth System Model 2 (CESM2, https://www.cesm.ucar.edu/models/cesm2), which is the latest version released by the National Center for Atmospheric Research (NCAR) in 2019 (Lawrence et al., 2019). Compared with CLM4.5, CLM5.0 is updated with several parameterizations in soil and plant hydrology, snow density, river modeling, carbon and nitrogen cycling and crop modeling (Lawrence et al., 2019). Previous studies have confirmed that CLM5.0 performs better than CLM 4.5 in soil temperature simulation (Deng et al., 2020; Luo et al., 2020). A 7-year (2001–2007) spin-up simulation at the regional scale over the HRB-MU was implemented to reach an initialization equilibrium of the thermal regime for CLM5.0. Subsequently, offline numerical simulations with a 0.05° × 0.05° grid resolution and a 1-h time interval during 2008–2011 were conducted in the prescribed satellite phenology mode.

3.2.3. *Reasoning of the mechanism guidance coupled in the ML model*

As mentioned, the input variable guidance is applied to incorporate mechanistic knowledge into the LGBM model. Thus, it is important to explore which forcing data are used to carry out the mechanism simulation of the dynamics and what the mechanistic guidance coupled into the ML model actually represents. Here, we provide a detailed reasoning

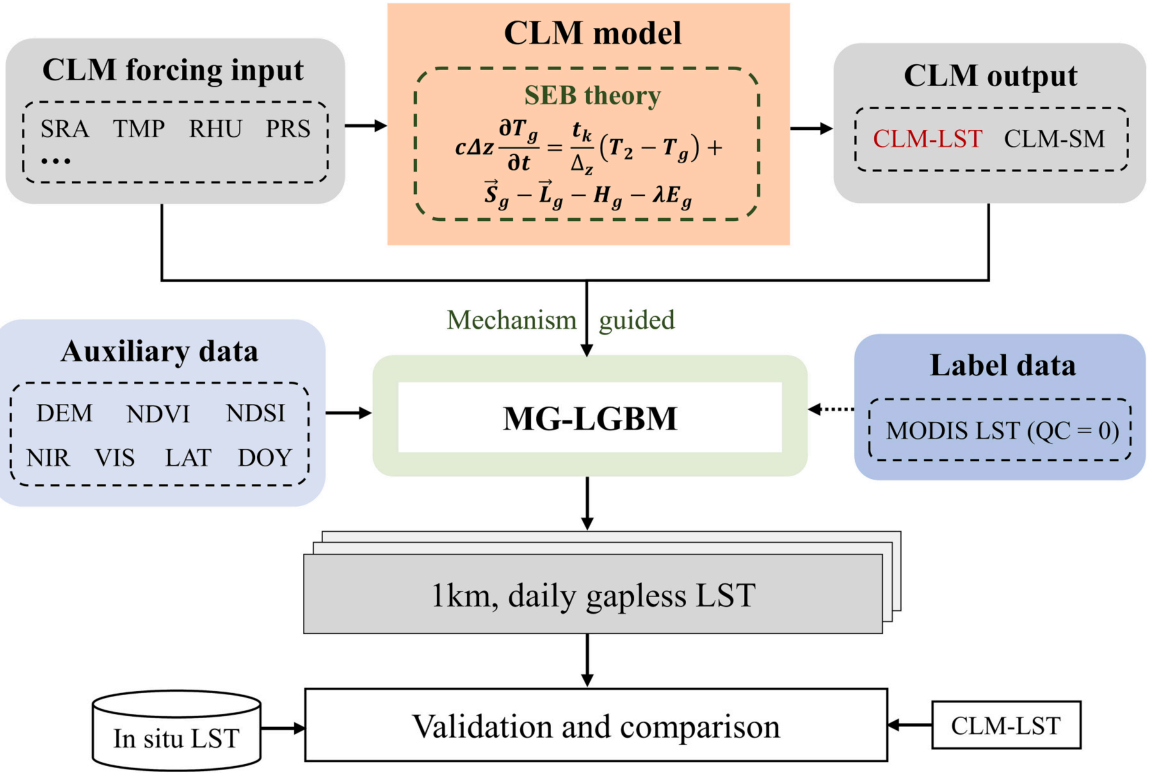


Fig. 3. Framework of the proposed mechanism-guided ML method for mapping gapless LST. Δz is the layer thickness between two layers; T_g and T_2 are the soil temperatures at the first and second layer, respectively; $\vec{S}_g$ and $\vec{L}_g$ are the net solar radiation and the net longwave radiation absorbed by the ground, respectively; and H_g and λE_g are the sensible and latent heat fluxes, respectively.

process for the retrieval of LST in the CLM, to identify the meteorological forcing data that are mechanistically related to CLM-LST and the mechanistic laws containing in the mechanism-based equations. The mathematical derivation is given below.

In CLM 5.0, LST is computed as follows:

$$T_s = \left(\frac{L\uparrow}{\sigma}\right)^{1/4} \tag{6}$$

where T_s is the skin temperature (LST); and $L\uparrow$ is the surface upward longwave radiation, which can be estimated as:

$$L\uparrow = \begin{cases} (1-\varepsilon_g)L_{atm}\downarrow + \varepsilon_g\sigma(T_g)^4, \mathrm{L+S} < 0.05 \\ (1-\varepsilon_g)(1-\varepsilon_v)^2 L_{atm}\downarrow + \varepsilon_v\left[1+(1-\varepsilon_g)(1-\varepsilon_v)\right]\sigma(T_v)^4 + \varepsilon_g(1-\varepsilon_v)\sigma(T_g)^4, \text{otherwise} \end{cases} \tag{7}$$

where L and S denote the exposed leaf area index and stem area index, respectively; $L+S<0.05$ represents the non-vegetated surfaces, and otherwise vegetated surfaces; T_g and T_v are the ground (soil, snow and surface water) and vegetation temperatures, respectively; σ is the Stefan-Boltzmann constant; ε_g and ε_v are the ground and vegetation emissivity, respectively; and $L_{atm}\downarrow$ is the downward atmospheric longwave radiation, which is calculated based on Idso (1981):

$$L_{atm}\downarrow = \left[0.7 + 5.95\times10^{-5}\times0.01\times e_{atm}\times exp\left(\frac{1500}{T_{atm}}\right)\right]\sigma\, T_{atm}{}^4 \tag{8}$$

where $e_{atm}=\frac{P_{atm}q_{atm}}{0.622+0.378q_{atm}}$ is the atmospheric vapor pressure; and P_{atm}, q_{atm}, and T_{atm} are the air pressure (PRS), the air specific humidity (RHU), and the air temperature (TMP), which were obtained from the CLDAS meteorological forcing dataset.

The numerical solutions for T_g and T_v are based on the SEB equation. Taking T_g as an example, the SEB is defined as follows:

$$G = \vec{S}_g - \vec{L}_g - H_g - \lambda E_g \tag{9}$$

where G is the ground heat flux, which is used for the soil temperature calculation; $\vec{S}_g$ and $\vec{L}_g$ are the net solar radiation and the net longwave radiation absorbed by the ground, respectively; H_g and λE_g are the sensible and latent heat fluxes, respectively, where E_g is the water vapor flux, and λ is a multiplier for converting the water vapor flux to an energy flux. The surface net radiation ($\vec{S}_g - \vec{L}_g$) is given as:

$$R_{n,g} = (1-\alpha)S\downarrow + \varepsilon_g L_{atm}\downarrow - \varepsilon_g\sigma(T_g)^4 \tag{10}$$

where $R_{n,g}$ is the surface net radiation; $S\downarrow$ is the incident solar flux (W m^{-2}), which is obtained from the solar shortwave radiation (SRA); and α is the ground albedo, which is associated with the land-cover type and soil color.

The calculations of H_g and E_g are based on Monin-Obukhov similarity theory (Monin and Obukhov, 1954) using the ground temperature from the previous time step, in conjunction with the atmospheric potential temperature, specific humidity, and thermodynamic and aerodynamic resistances. The formulas for H_g and E_g can be found in CLM5.0 technical note (https://escomp.github.io/ctsm-docs/versions/release-clm5.0/html/tech_note/). H_g and E_g are then used as the surface forcing for the solution of the ground temperature at the current time step. The law of heat conduction in one-dimensional form is:

$$c\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left[t_k \frac{\partial T}{\partial z}\right] \tag{11}$$

where c denotes the volumetric heat capacity of snow/soil (J m^{-3} K^{-1}); t denotes the time, which was 3600 s in this study; t_k denotes the thermal conductivity (W m^{-1} K^{-1}), depending on the soil property; and z denotes the vertical direction depth (m). Combining Eqs. (9) and (11) yields the energy balance equation for LST calculation:

$$c\Delta z\frac{\partial T_g}{\partial t} = \frac{t_k}{\Delta_z}\left(T_2 - T_g\right) + \overrightarrow{S}_g - \overrightarrow{L}_g - H_g - \lambda E_g \tag{12}$$

where Δz denotes the layer thickness (m) between the two layers; T_g and T_2 are the soil temperature (K) at the first and second layer, respectively. In CLM5.0, there are 25 layers in total, which are thinker from the top layer to the bottom layer. This equation is solved numerically using the Crank-Nicolson method to calculate the ground temperature, with the boundary conditions of G as the ground heat flux into the top ground surface from the overlying atmosphere.

According to Eq. (12), LST is determined by subtracting the outgoing energy from the incident energy, which represents the amount of energy that is absorbed by the surface (Jin, 2000). To represent the incident energy, the SRA is used (Eq. (10)). In CLM5.0, the downward atmospheric longwave radiation $L_{atm}\downarrow$ also exerts a crucial role on LST (Eqs. (7) and (10)). The $L_{atm}\downarrow$ is parameterized using TMP, PRS, and RHU (Eq. (8)). Furthermore, TMP and RHU are also required in the sensible and latent heat flux parameterizations. In this respect, the CLM forcing data (specifically, SRA, TMP, PRS and RHU) with a strong mechanistic relationship on CLM-LST are selected as the input variables for the MG-LGBM model, making it feasible in the identification of causation from the correlation between the model forcing data and simulation data. Based on Eq. (12), we can also observe that interpretable process knowledge (i.e., the SEB and the law of heat conduction) is concealed within the simulation process of CLM-LST.

3.3. Evaluation strategies

LST varies significantly on diurnal and intra-annual scales (Göttsche and Olesen, 2001; Zhan et al., 2014). In this respect, we divided all the samples from 2008 to 2011 into eight data subsets based on spring (March–May), summer (June–August), autumn (September–November), and winter (December–February), as well as daytime and nighttime. Therefore, eight models in total were trained and evaluated.

Sample-based validation, space-based validation and independent site validation were used for the model validation. Fig. 4. shows the schematic diagram of sample-based validation and space-based validation. For the sample-based validation, 10-fold cross-validation (CV) was utilized, i.e., all the matched (spatially and temporally) samples were randomly and evenly divided into 10 folds, with nine folds used for model training and one for model validation. This process was repeated 10 times so that each fold can be validated, and the average accuracy was calculated. The sample-based CV shows superiority in evaluating the overall model performance and overfitting problems (Rodriguez et al., 2009). However, it is less effective in tackling spatial extrapolation problems in particular locations (Shen et al., 2022). Accordingly, space-based validation was applied to further test the model's spatial prediction accuracy in regions without valid observations. The procedure of space-based validation is similar to that of sample-based validation. The only difference is that the spatial grids within the study area were evenly divided into 10 folds (Fig. 1), instead of all the matched samples (Li et al., 2020c). The generated LST product was then further compared with independent in-situ LST and the CLM-LST. The Pearson correlation coefficient (R), the RMSE, the mean absolute error (MAE) and the overall bias (BIAS) were selected as the validation indicators.

4. Results

4.1. Spatio-temporal patterns of the MG-LGBM estimated gapless LST

One representative year (2010) was selected to enable the investigation into the spatio-temporal patterns of the generated gapless LST. Fig. 5 shows the spatial distribution of the CLM-simulated LST, MODIS LST, and estimated gapless LST in the different seasons (DOYs 16, 105, 196, and 287) in the daytime and nighttime of 2010. Due to the interference of clouds, missing data are prevalent in the MODIS LST. Despite spatial completeness, the CLM-LST has relatively few spatial details. Comparatively, the MG-LGBM LST is spatially complete and retains a high consistency with the MODIS LST in spatial details as well as LST magnitude. In addition, there are no block effects or artifacts in the estimated LST image, indicating that the MG-LGBM model can successfully address the scale inconsistencies between the input data and MODIS data. As shown in Fig. 5, the LST is highly variable in space and time, especially during the summer daytime, whereas relatively gentle variation is found during the nighttime. Relatively high LST values are observed in the low-altitude northeast region of the HRB-MU, where bare land is dominant. Likewise, relatively low LST values are observed in the high-altitude southern HRB-MU, where grassland and forest are widespread. Overall, it is viable to use the MG-LGBM model to generate gapless LST with full spatial continuity and to accurately depict the spatial thermal dynamics of the land surface.

Fig. 6 displays the temporal variation of the MG-LGBM estimated all-weather LST during the daytime and nighttime in 2010 over the three sites. The in-situ LST, pure ML LST (e.g., ML predicted LST which only uses remote sensing data) and the corresponding MODIS LST are also provided for a comparison. The black, blue, and red lines represent the in-situ LST, pure ML LST, and estimated all-weather LST, respectively. The clear-sky MODIS LST is represented as a black hollow circle. The results show that the estimated all-weather LST can basically capture the seasonal and daily variations during both the daytime and nighttime, with an accuracy comparable to that of clear-sky MODIS LST, but with stronger time continuity. It is shown that the estimated daytime LST varies more sharply than the nighttime LST, with an RMSE of 2.92–3.79

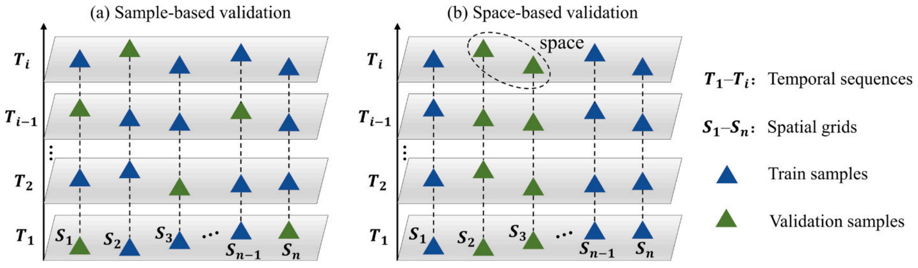
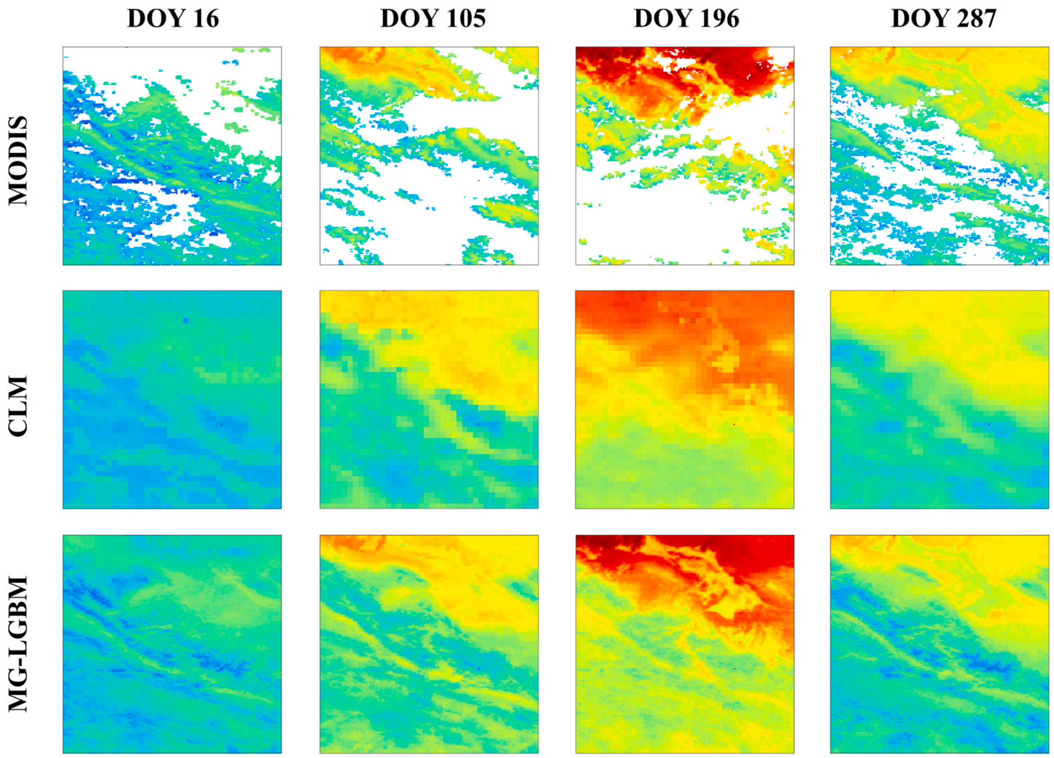
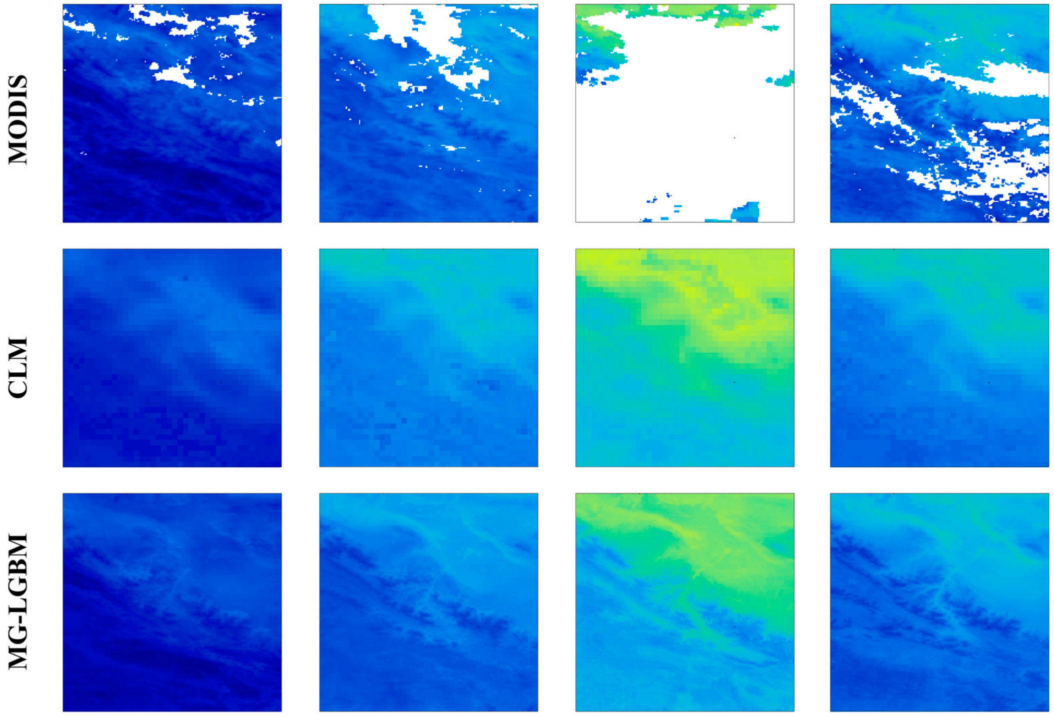


Fig. 4. Schematic diagram of sample-based validation and space-based validation in our study.



(a) Daytime



(b) Nighttime

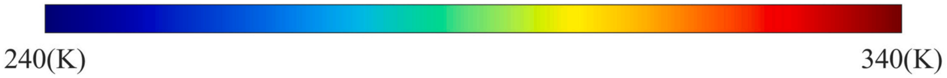
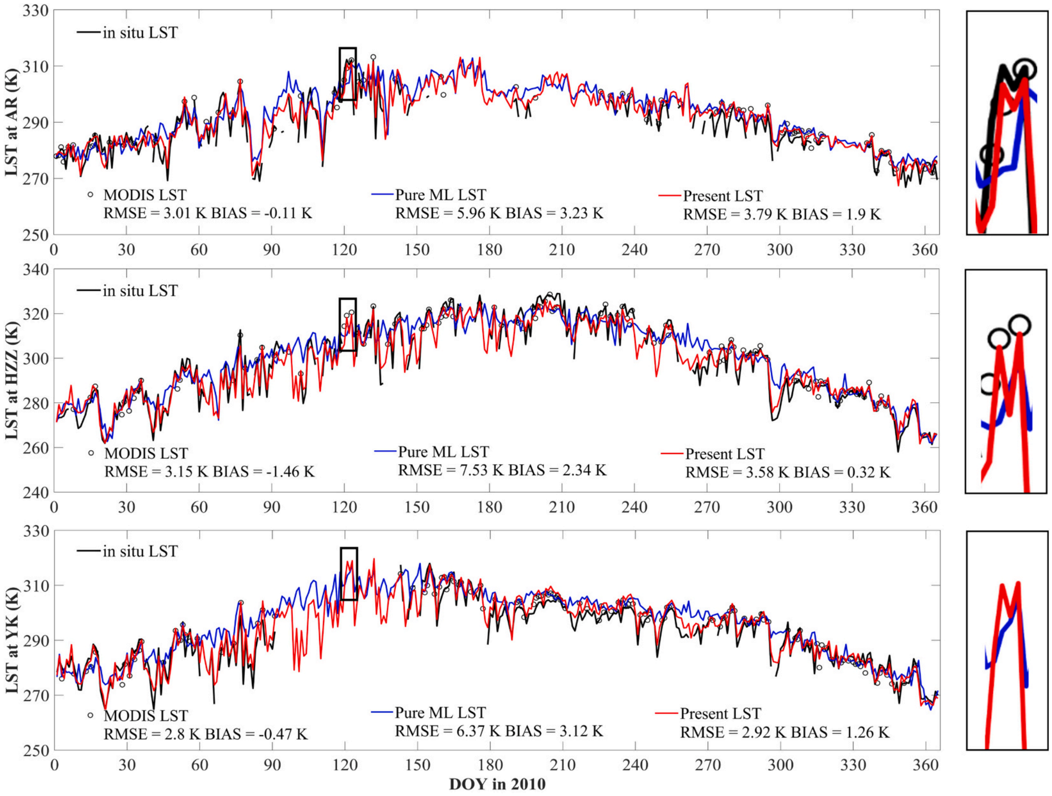
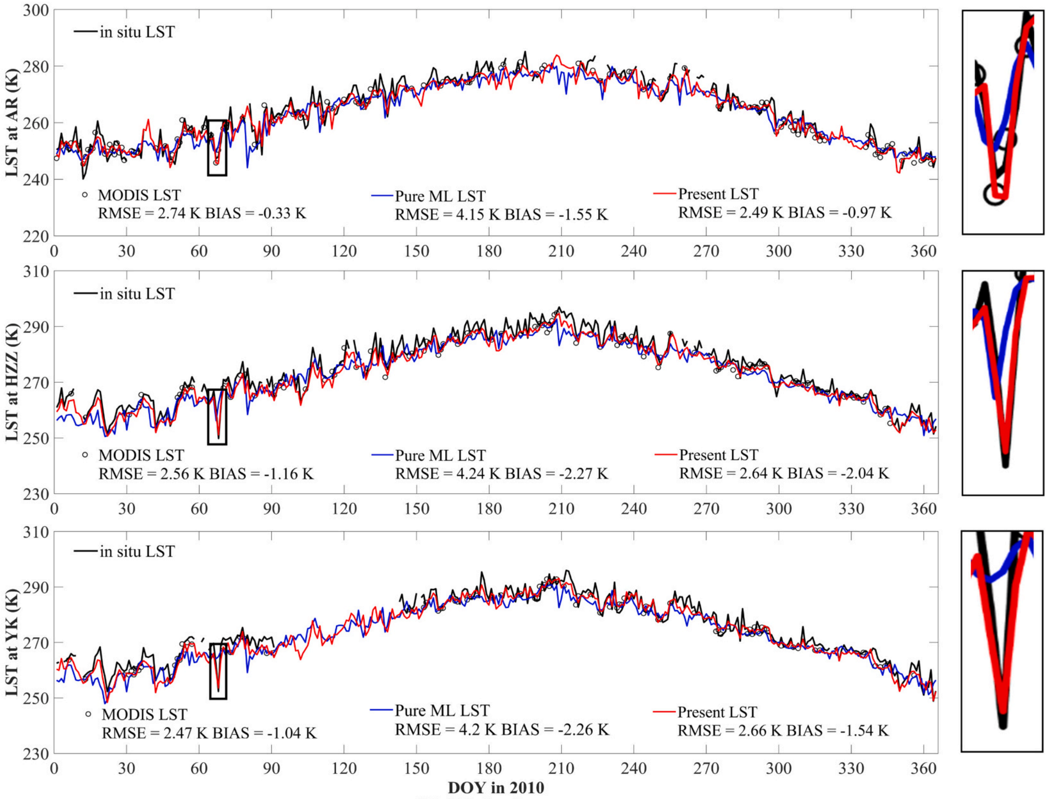


Fig. 5. Spatial patterns of the original MODIS LST, CLM-simulated LST, and estimated gapless LST for DOYs 16, 105, 196, and 287 in 2010 during (a) the daytime (around 13:30 local time) and (b) nighttime (around 01:30 local time). The land cover map can be referenced in Fig. 1.



(a) Daytime



(b) Nighttime

Fig. 6. Time series of the in-situ LST, clear-sky MODIS LST, pure ML LST, and the MG-LGBM estimated all-weather LST at the three sites in 2010 during (a) the daytime and (b) the nighttime. The first, second and third rows in each subgraph represent the AR, HZZ and YK sites, respectively. DOY denotes the day of year. The discontinuity of the black line is mainly due to the missing in-situ LST data.

K and 2.49–2.66 K during the daytime and nighttime, respectively. Benefiting from the incorporation of the mechanistic laws, the proposed method can capture some of the sudden rises (e.g., DOY 121 during the daytime) or drops (e.g., DOY 67 during the nighttime) in LST. Compared with the MG-LGBM estimated LST, the pure ML LST exhibits an inferior performance, in terms of an RMSE of 4.15–7.53 K and a BIAS of −2.27–3.23 K. By extracting complex features from CLM forcing data and CLM simulated data, the MG-LGBM estimated LST is to largely compensate for these errors.

4.2. Quantitative evaluation results

The performance of the MG-LGBM model was validated by simulation and real-data experiments using MODIS LST and in-situ LST measurements, respectively. Fig. 7 shows the sample-based and space-based validation results for the daytime and nighttime. As witnessed in Fig. 7, the sample size of the validation is adequate (>10 million), which demonstrates the reliability of the validation results. For the sample-based CV results, the averages of the R, RMSE, MAE, and BIAS are 0.99, 2.03 K, 1.49 K, and 0 K, respectively, for the daytime. Meanwhile, these values for the nighttime are 0.99, 1.23 K, 0.92 K, and 0 K, respectively. The results indicate that the MG-LGBM model shows a favorable performance in LST prediction. Moreover, the MG-LGBM model shows good results under the space-based validation, with an R of 0.98 (0.99), RMSE of 2.62 K (1.56 K), MAE of 1.94 K (1.16 K), and BIAS of −0.05 K (0.02 K) for the daytime (nighttime), which verifies the predictive ability of the model in regions without training samples. Compared with the model performance in the daytime, a higher accuracy can be observed in the nighttime, which could be associated with the high thermal heterogeneity and TIR directional anisotropy during daytime (Cao et al., 2019; Zhang et al., 2021).

In-situ LST measurements from four sites with different land-cover types were further used to assess the estimated gapless LST. Fig. 8 shows the scatter plots of the estimated LST against in-situ LST measurements under all-weather conditions. The CLM-simulated LST under cloudy-sky conditions is also provided for a comparison. As can be seen, the MG-LGBM estimated LST is in good agreement with the in-situ LST under clear-sky conditions, with an RMSE of 2.45–2.79 K, MAE of 2.05–2.21 K and R of 0.99. Through the powerful ability of the ML model in data mining, MG-LGBM largely corrected the systematic discrepancies

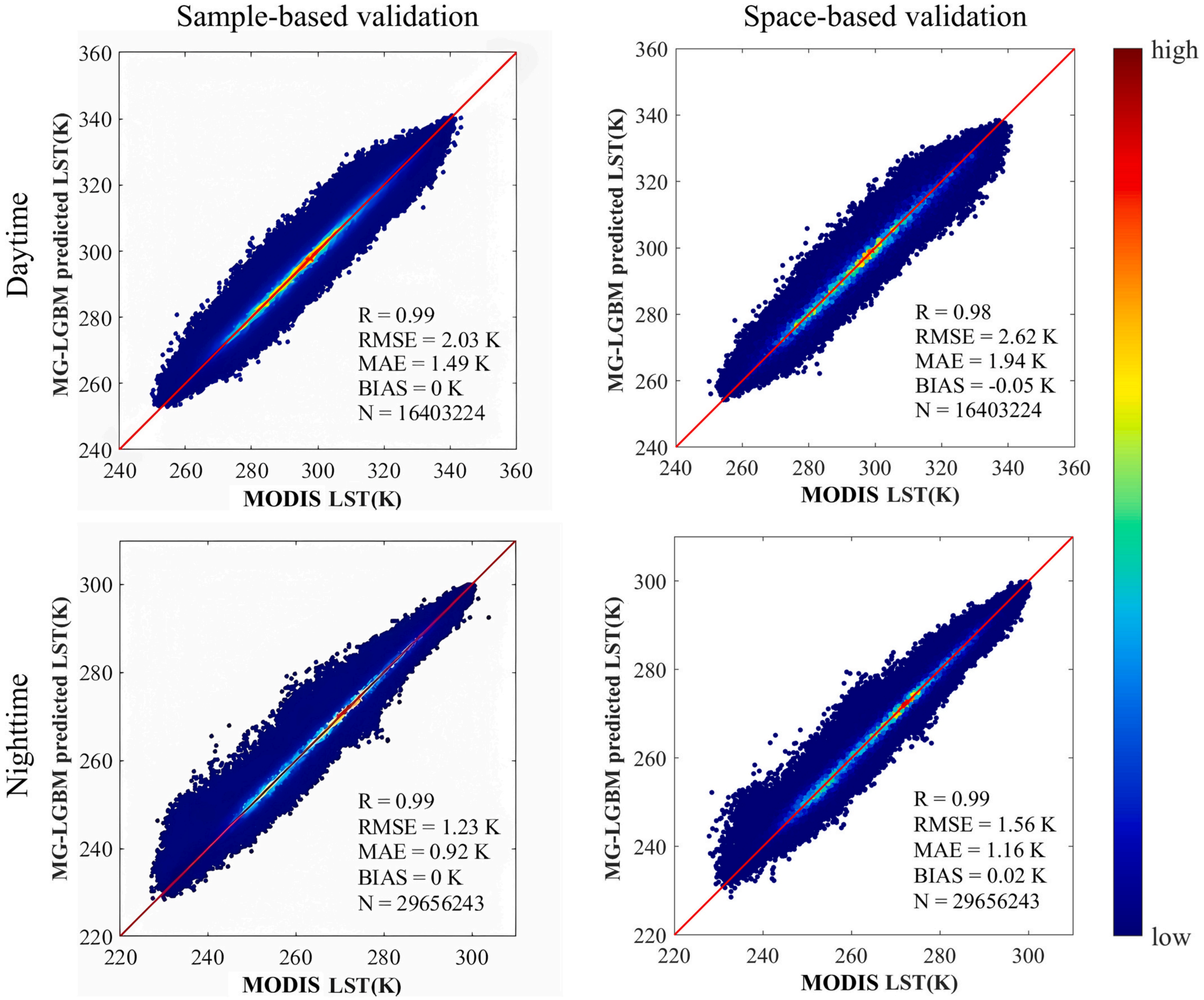


Fig. 7. Scatter plots of the sample-based CV results and space-based CV results for the daytime and nighttime. The red solid line denotes the 1:1 line. The color bar is the density of the samples. N represents the sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

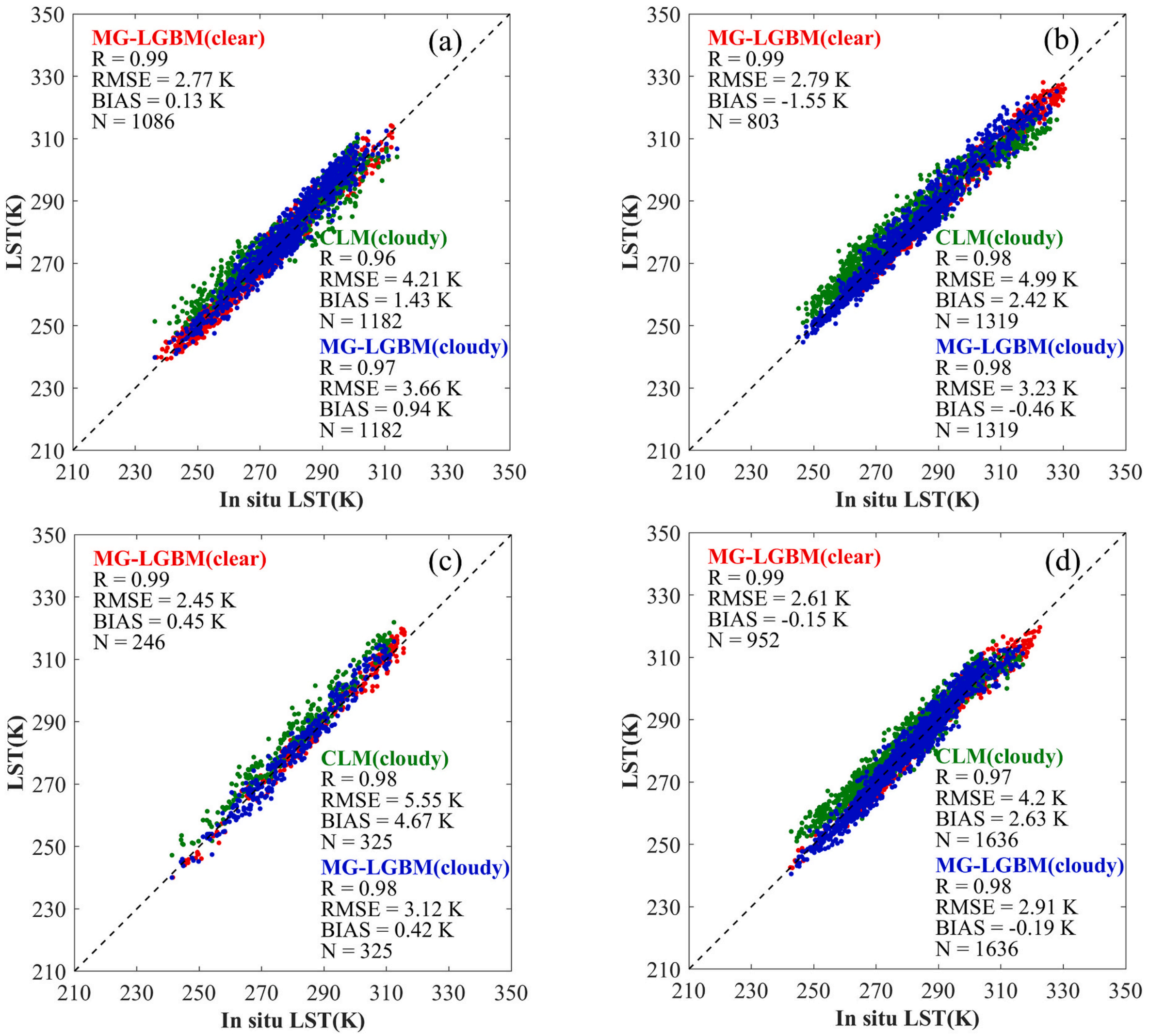


Fig. 8. Scatter plots of the estimated LST against in-situ LST measurements under all-weather conditions and CLM-LST against in-situ LST measurements under cloudy-sky conditions during 2008 to 2011 at the four sites: (a) AR (alpine meadow), (b) HZZ (desert steppe), (c) LZ (cropland), and (d) YK (grassland).

in CLM-LST. As a result, the accuracy of the cloudy-sky LST is generally comparable to that of the clear-sky LST, in terms of RMSE differences of 0.3–0.89 K, suggesting that the proposed approach can successfully recover cloudy LST from the established clear-sky model. In more detail, a better performance is observed at the YK and LZ sites, with an RMSE of 2.91 K and 3.12 K, respectively, whereas a relatively poor performance is observed at the HZZ and AR sites, with an RMSE of 3.23 K and 3.66 K, respectively. The inferior accuracy can be attributed to the errors of the MODIS LST retrievals arising from the uncertainty in emissivity estimation in specific land-covers (e.g., the HZZ site in a desert region, and the AR site in an alpine snowy region) (Duan et al., 2019; Li et al., 2019). In general, the estimated gapless LST resembles the in-situ LST reasonably well under both clear-sky and cloudy-sky conditions, demonstrating the reliability of the proposed method for estimating all-weather LST.

4.3. Comparison with other methods and datasets

4.3.1. Comparison with pure ML methods

In this study, the key CLM forcing data and CLM simulation data were jointly compiled for training an ML model, according to their strong causal associations. In order to better understand the contributions of mechanistic guidance made to the LGBM model, the performance of ML models with different combinations of data was evaluated.

Table 3
Space-based model accuracy for different combinations of data, based on the LGBM model. The bold metrics denote the best performances among all the data combinations.

Time	Input combinations	RMSE (K)	MAE (K)	R
Daytime	RS	3.88	2.94	0.97
	RS + CF	2.7	2	0.98
	RS + CS	2.95	2.19	0.98
	RS + CF + CS (MG-LGBM)	**2.62**	**1.94**	**0.98**
Nighttime	RS	2.62	1.96	0.98
	RS + CF	1.65	1.22	0.99
	RS + CS	1.85	1.36	0.99
	RS + CF + CS (MG-LGBM)	**1.56**	**1.16**	**0.99**

RS: remote sensing data; CF: key CLM forcing data, i.e., SRA, TMP, PRS and RHU; CS: CLM simulation data, i.e., CLM-LST and CLM-SM. The numbers of samples are 16,403,324 (29656243) for the daytime (nighttime), respectively.

Table 3 summarizes the model accuracy metrics for different combinations of data under the space-based validation. It should be clarified that RS, CF, and CS represent remote sensing data, key CLM forcing data (SRA, TMP, PRS and RHU), and CLM simulation data (CLM-LST and CLM-SM), respectively. Due to the introduction of input variable mechanistic guidance (i.e., the underlying process knowledge between CF and CS in CLM modeling) in the ML model, RS + CF + CS belongs to a mechanism-guided ML method (see Section 3.2), while RS, RS + CF, and RS + CS belong to pure ML methods. The results show that all the methods achieve reasonable accuracies, with the RMSE (MAE) lower than 3.88 K (2.94 K) and *R* >0.97. In more detail, the worst accuracy is found when using only remote sensing data as input, with an RMSE of 3.88 K and 2.62 K for the daytime and nighttime, respectively. It is demonstrated that the daily remote sensing data interpolated from an 8-day or 16-day temporal resolution cannot sufficiently depict the high daily fluctuations in LST. When instantaneous key CLM forcing data or CLM simulation data are incorporated into the ML model, obvious improvements can be observed, compared with the basic result, followed by an RMSE decrease of 0.77–1.18 K, a MAE decrease of 0.6–0.94 K, and an R increase of 0.01–0.02. Compared with the pure ML methods, it is evident that the proposed MG-LGBM model (i.e., RS + CF + CS) performs the best among all the data combinations, which demonstrates the superiority of coupling mechanistic knowledge with the ML model in LST estimation.

To further evaluate the influence of the mechanistic guidance on the estimated LST under all-weather conditions, scatter plots of the all-weather LST with different data inputs against in-situ LST are presented in Fig. 9. The CLM modeling is also added as a pure mechanistic method for comparison. The accuracy of CLM-LST in cloudy-sky conditions (RMSE = 4.56 K, *R* = 0.97) is comparable with that under clear-sky conditions (RMSE = 4.63 K, *R* = 0.98), indicating that the model

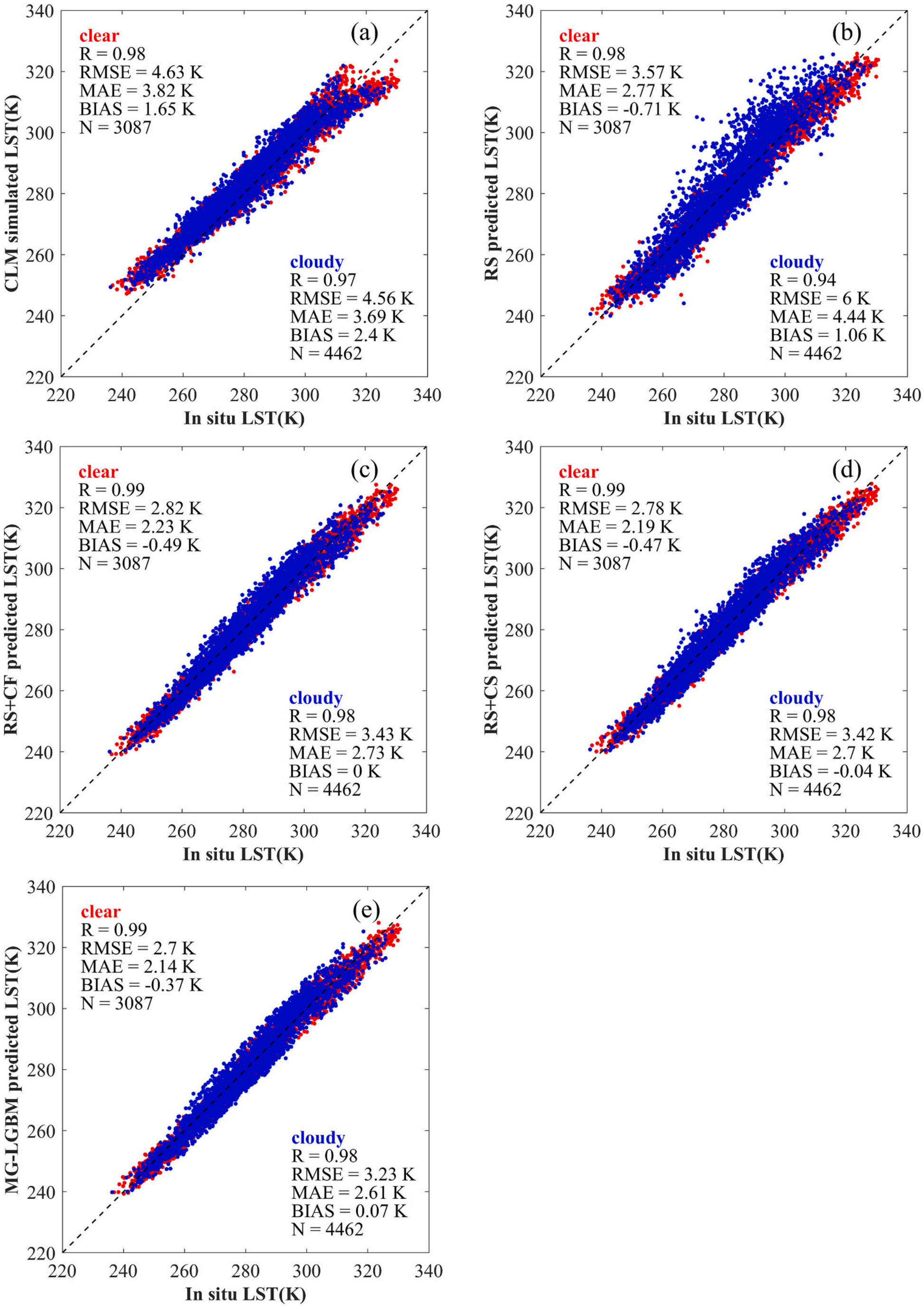


Fig. 9. Scatter plots of the all-weather LST against in-situ LST (four sites) during 2008 to 2011 with different data inputs: (a) CLM simulated; (b) RS; (c) RS + CF; (d) RS + CS; and (e) RS + CF + CS.

Table 4
The bias between predicted LSTs and in-situ measurements under cloudy-sky conditions during the daytime and nighttime, respectively.

Time	Input combinations	BIAS (K)
Daytime	RS	4.32
	RS + CF	1.65
	RS + CS	1.67
	RS + CF + CS (MG-LGBM)	**1.57**
Nighttime	RS	−1.95
	RS + CF	−1.53
	RS + CS	−1.53
	RS + CF + CS (MG-LGBM)	**−1.43**

The bold metrics denote the best performances among all the data combinations.

forcing and simulation data can effectively characterize the actual thermal information under clouds (Fig. 9a). This phenomenon also provides evidence in support of the reliability of applying the clear-sky ML model to recover LST in cloudy-sky conditions. Among all the estimated LSTs, the RS LST performs worst, with an RMSE of 3.57 K (6 K) and R of 0.94 (0.98) in clear-sky (cloudy-sky) conditions. Table 4 further lists the bias of predicted LSTs under cloudy-sky conditions during the daytime and nighttime, respectively. The RS LST demonstrates poor performances in reflecting the cooling and warming effects of clouds during the daytime and nighttime, respectively. This results in a significant overestimation (BIAS = 4.32 K) during the daytime and an underestimation (BIAS = −1.95 K) during the nighttime in LST in cloudy conditions. Once the model forcing data or simulation data are incorporated into the ML model, significant improvements in LST accuracy can be found, with an RMSE, MAE, and R of 2.78–3.43 K, 2.19–2.73 K, and 0.98–0.99, respectively (Fig. 9c−d). Additionally, this incorporation substantially reduces the bias of RS LST (Table 4). The best accuracy is found for RS + CF + CS LST, for which the RMSE, MAE, BIAS, and R are 2.7–3.23 K, 2.14–2.6 K, −0.37 − 0.07 K, and 0.98–0.99, respectively, which ranks first under both clear-sky and cloudy-sky conditions in all the scenarios (Fig. 9e). This again confirms that integrating key LSM forcing data and LSM simulation data simultaneously into the ML model actually makes sense in all-weather LST estimation. A reasonable illustration is that the MG-LGBM model is capable of implicitly learning the explicit process knowledge (e.g., the SEB) embedded in the CLM, guiding the learning process and improving the accuracy.

4.3.2. Comparison with other advanced ML methods

To further investigate the effectiveness of the proposed ML method, we selected three other advanced ML methods, including an XGBoost model (Chen and Guestrin, 2016) and two deep learning models, i.e., a DBN (Hinton et al., 2006) and a GRNN (Specht, 1991), as the comparative approaches, which are popular and widely used in climate and environment variable estimation (Li et al., 2020b; Shen et al., 2020; Tan et al., 2021b). It is worth noting that the random forest model was not used in this study, because it is time-consuming and inappropriate for tackling massive (e.g., 10 million level) datasets (Breiman, 2001). In addition to the test accuracy, the efficiency is critical when the data size is large. Fig. 10 shows the model accuracy and the time efficiency under the space-based validation for the different ML models using the same input data. The experimental environment is a Linux (Centos 7) server, with an Intel(R) Xeon(R) Bronze 3104 CPU (12 cores) and 256 GB memories. As shown, all the ML methods achieve a reasonable accuracy, with an RMSE (R) of 2.62–3.55 K (0.97–0.99) during the daytime and 1.56–2.25 K (0.98–0.99) during the nighttime. The decision tree based boosting models (i.e., MG-LGBM and MG-XGBoost) perform better than the neural network based models (i.e., MG-DBN and MG-GRNN) on the basis of model accuracy, which could be due to the advantages of boosting methods in countering overfitting problems when solving pixel-pixel regression tasks (Chen et al., 2022). In particular, the MG-LGBM model performs the best among all the mechanism-guided ML methods, with an RMSE of 1.56–2.62 K and R of 0.99. Furthermore, MG-LGBM accelerates the training process mostly, speeding up by an average of 206×, 127× and 5.1× respectively comparing with MG-DBN, MG-GRNN and MG-XGBoost models. The encouraging efficiency was mainly attributable to the EFB and GOSS strategies used in MG-LGBM. Overall, the LGBM-based method outperforms the other popular ML methods in terms of model accuracy and efficiency.

4.3.3. Comparison with other gapless LST data

A number of fusion or reconstruction methods have been proposed to generate gapless LST. To further evaluate the proposed method, we compared the MG-LGBM method with one interpolation method (Zhang et al., 2022) and three fusion-based methods (Wu et al., 2022; Xu and Cheng, 2021; Zhang et al., 2021), which were recently proposed and proven to be effective in gapless LST mapping. Note that the four aforementioned methods have been applied to produce gapless LST data or products covering the HRB-MU, so it is convenient to make such a comparison. Table 5 provides the summary information of the gapless LST data. To prevent any possible confusion, the products produced by Zhang et al. (2022) and Zhang et al. (2021) are referred to as the Zhang I's LST and Zhang II's LST, respectively. Fig. 11 shows a spatial comparison between the MG-LGBM LST and the four sets of gapless LST data with the original MODIS LST for DOY 106 in 2010. It can be seen that, in comparison with the original MODIS LST, all the LST data are spatially complete during both the daytime and nighttime, but show a difference

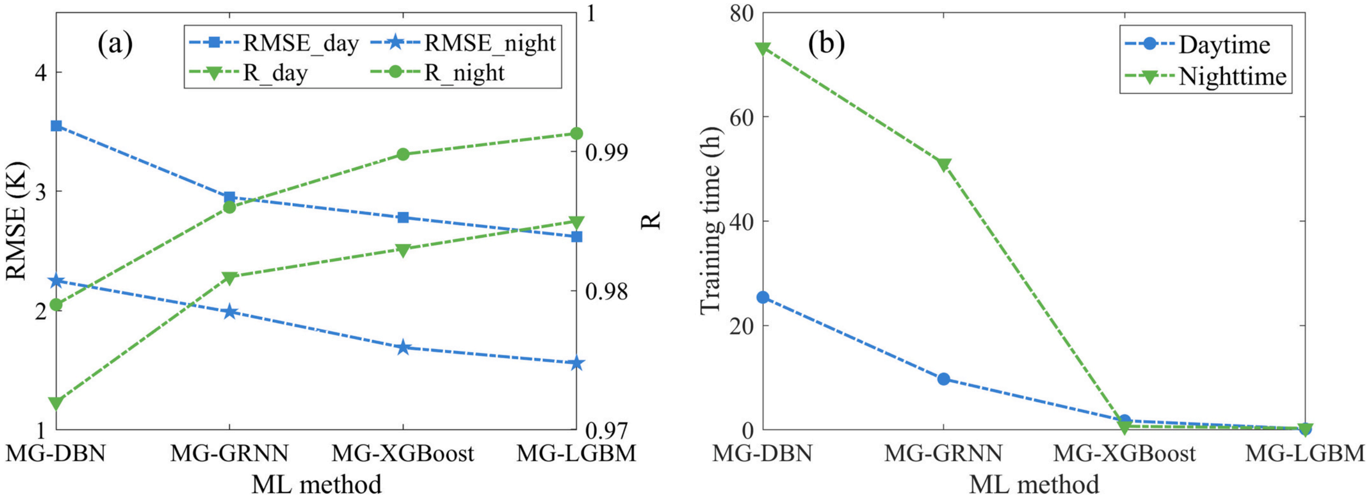


Fig. 10. Space-based (a) model accuracy (RMSE and R) and (b) efficiency (training time) for the different ML models.

Table 5
Summary of the gapless LST data used for a comparison.

Study	Spatiotemporal resolution	Spatial Coverage	Temporal Coverage	Input data	Methodology
Zhang et al. (2022)	1 km / 13:30 and 01:30 local time (LT)	global land	2003–2020	MODIS	spatiotemporal interpolation of MODIS LST
Zhang et al. (2021)	1 km / 13:30 and 01:30 LT	Tibetan Plateau	2000–2021	MODIS GLDAS/ CLDAS	fusion of TIR LST and reanalysis LST by LST time series decomposition
Xu and Cheng (2021)	1 km / 13:30 and 01:30 LT	mainland China	2002–2020	MODIS AMSR-2	fusion of TIR LST and PMW LST by CDF[a] matching and MKF[b]
Wu et al. (2022)	1 km / 13:30 and 01:30 LT	Mainland China	2010	MODIS AMSR-E	fusion of TIR LST and PMW LST by a deep learning method
Our	1 km / 13:30 and 01:30 LT	HRB-MU	2008–2011	MODIS CLDAS GLASS	a mechanism-guided ML method using multi-source remote sensing and model simulated data

[a] Cumulative distribution function.
[b] Multiresolution Kalman filtering.

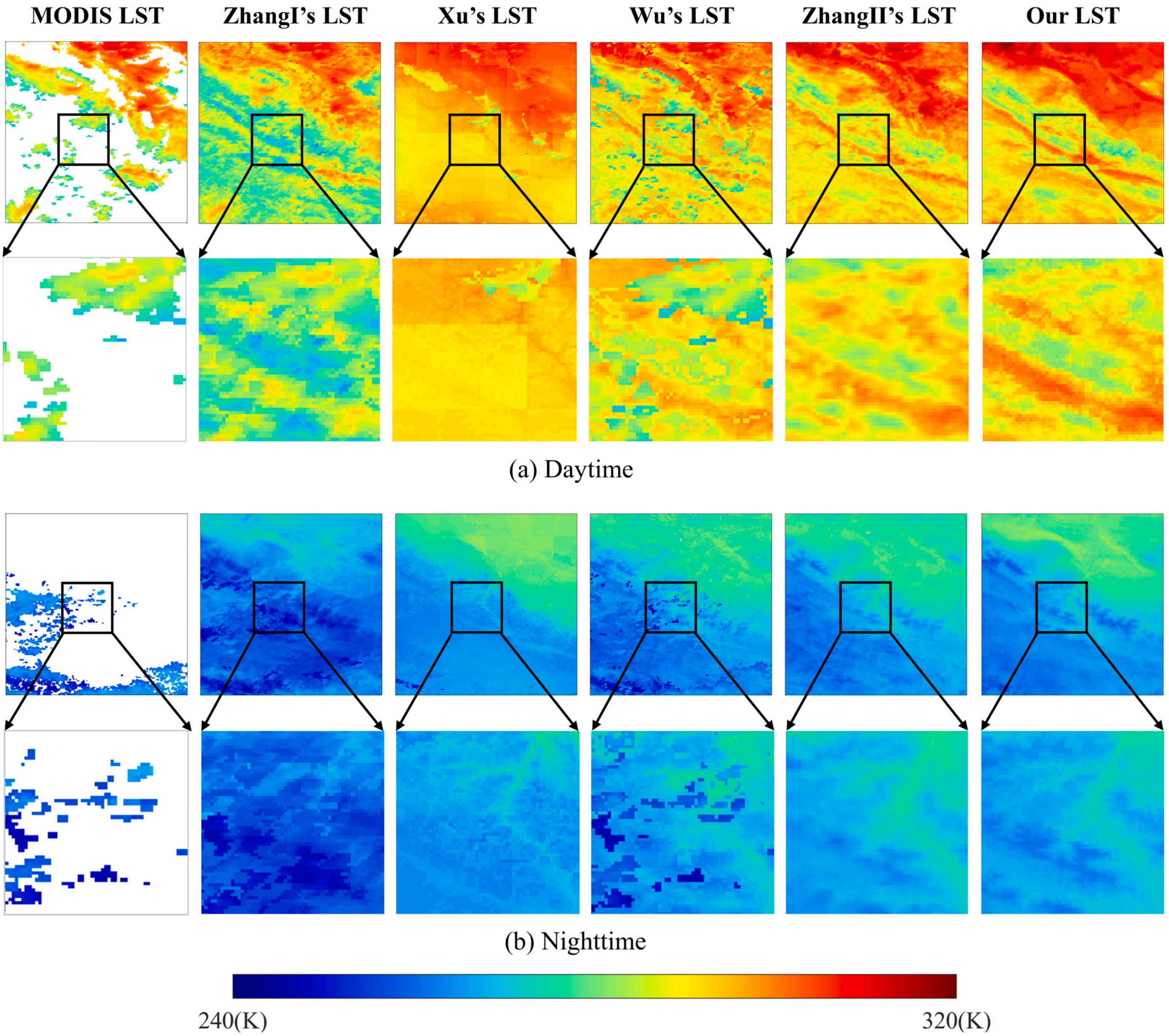


Fig. 11. Spatial patterns of the original MODIS LST and the gapless LST from the ZhangI's LST (Zhang et al., 2022), Xu's LST (Xu and Cheng, 2021), Wu's LST (Wu et al., 2022), ZhangII's LST (Zhang et al., 2021) and our LST for DOY 106 in 2010 during (a) the daytime and (b) the nighttime. The black boxes denote certain regions for a detailed comparison.

in spatial patterns and LST magnitude. Obvious block effects and "blurred" effects are found in the Xu's LST (Fig. 11a). This phenomenon can be attributed to the coarse resolution of the AMSR-2 data (i.e., 0.1° × 0.1°) used for multiresolution Kalman filtering in LST fusion (Xu and Cheng, 2021). Since only spatial and temporal neighboring information of clear-sky MODIS LST is utilized in the reconstruction process proposed by Zhang et al. (2022), the reconstructed LST is a hypothetical "clear-sky" LST, resulting in a spatial difference between the Zhang I's LST and other LSTs in cloudy conditions. Some artifacts can be observed in the Wu's LST, because the original MODIS LST is retained. However, such a broken effect may not represent the real LST pattern due to the uncertainties in identifying the clear-sky pixels in MODIS LST (Wang et al., 2019). The uncertainties can lead to abnormal retrievals of partially cloud-contaminated pixels, particularly near the clustered cloud pixels (see Section 4.4) (Jia et al., 2022; Ma et al., 2020). Our LST shares a similar LST pattern with the Zhang II's LST, enhancing the spatial details and appearing more "natural", indicating the robustness of the proposed method in obtaining a more real LST.

To quantitatively compare the accuracy of the developed gapless LST with the four sets of LST data, the same in-situ measurements from the AR, HZZ, and YK sites were used. The validation results under all-weather conditions during the daytime and nighttime are shown in Fig. 12. In general, all the gapless LST data have a relatively high accuracy, with $R > 0.92$ (0.97) and the RMSE < 3.9 K (3.1 K) during the daytime (nighttime). Although the RMSE of our LST is slightly higher than that of some of the LST data (e.g., at the daytime AR site and the nighttime YK site), the mean RMSE remains the lowest, and the R of the MG-LGBM LST is higher than that of all the other LST data. The average values of R (RMSE) for the ZhangI's LST, Xu's LST, Wu's LST, ZhangII's LST and our LST from the three sites during both the daytime and nighttime are 0.96 (3.23 K), 0.96 (3.21 K), 0.96 (3.2 K), 0.96 (3.22 K), and 0.98 (3.01 K), respectively. Overall, the proposed method is superior to the above-mentioned fusion-based and reconstruction methods, based on both the mapping effect and validation accuracy over the HRB-MU region.

4.4. *Comparison between clear-sky LST and the original MODIS LST*

Clouds are prevalent over the land surface, and can be especially heavy in regions with complex terrain. To obtain more clear-sky MODIS LST samples for LST fusion (e.g., a similar pixel selection) or reconstruction, a number of studies have selected clear pixels with the "LST produced" and "average LST error <=3 K" according to the MODIS QC flags (hereafter termed "clear LST") (Duan et al., 2017; Gong et al., 2023; Long et al., 2019; Wu et al., 2022). However, owing to the solar-cloud-satellite geometry (SCSG) effect existing in the MODIS LST (Wang et al., 2019), LST detection based on poor QC flags may not represent the true situation. In this case, such a selection could weaken the accuracy of the fusion or reconstruction results.

In this study, the clear-sky MODIS LST pixels with the highest quality ("good quality" and "average LST error <= 1 K", hereafter termed "fully clear LST") were selected as the label data, and the likely-cloudy LST pixels ("other quality" and "average LST error > 1 K & <=3 K", hereafter termed "partially cloudy LST") were then recovered by MG-LGBM. Taking 2010 as an example, Fig. 13a shows the cumulative distribution plot of the spatially averaged percentages of the clear-sky samples for the daytime and nighttime over the HRB-MU. Overall, clear-sky LST pixels accounts for 55.3% over the HRB-MU, of which 38.4% are fully clear and a non-ignorable proportion of 16.9% are partially cloudless. Fig. 13b–c shows the scatter plots of the official MODIS LST under different conditions and the corresponding MG-LGBM LST against in-situ LST measurements in 2010. Under fully clear-sky conditions, the MG-LGBM LST closely approximates the MODIS LST, with an RMSE difference of 0.02 K, illustrating the excellent performance of the MG-LGBM model in reproducing the magnitude of clear-sky MODIS LST. Meanwhile, under partially cloudy conditions, the MODIS LST shows an unsatisfactory performance, followed by an RMSE of 5.31 K, MAE of 3.77 K, and a considerable negative BIAS of 2.58 K. It has been demonstrated that partially cloud-contaminated areas generally block the surface heating processes, leading to abnormally cool LSTs in the images (Jia et al., 2022). Correspondingly, the MG-LGBM LST largely makes up for this underestimation, with an RMSE of 3.72 K, MAE of 2.79 K, and a small BIAS of −0.45 K. In summary, MG-LGBM model can generate LST that is comparable to the original MODIS LST under fully clear-sky conditions and can correct for the likely-cloudy LST under partially cloudy conditions.

5. Discussion

5.1. *Impacts of time, land-cover type and elevation on clear-sky LST estimation*

LST exhibits significant temporal variations. Previous studies have indicated that the accuracy of LST estimation is affected by the season (Li et al., 2021; Wu et al., 2022) and month (Xiao et al., 2023) within the year. It is therefore of special significance to evaluate the accuracies of

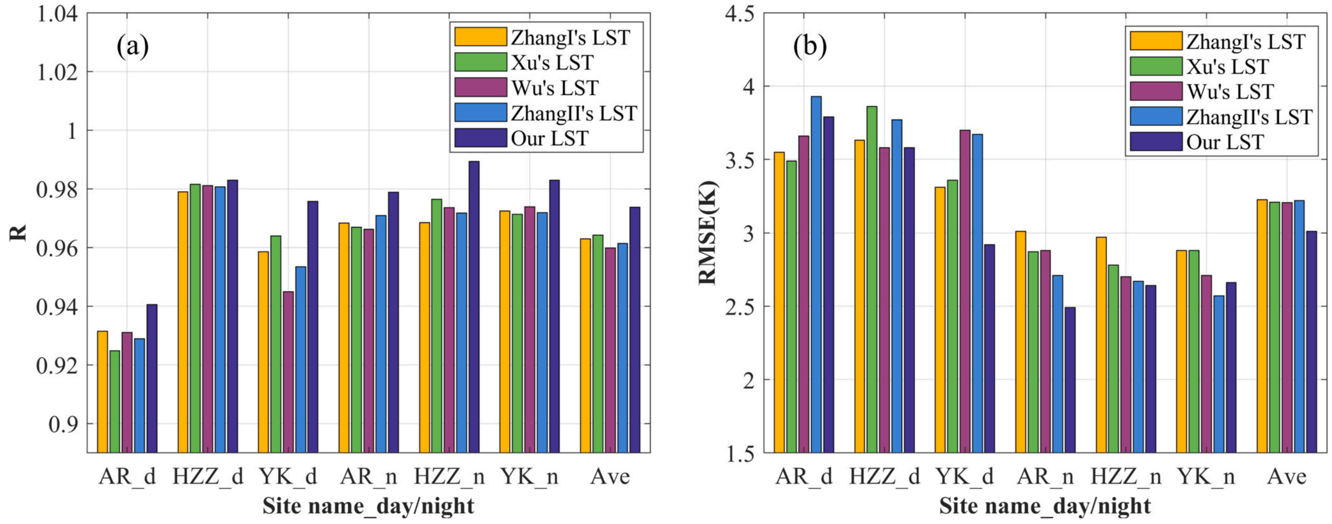


Fig. 12. Comparison between the gapless MG-LGBM LST and the LST obtained by Zhang et al. (2022), Xu and Cheng (2021), Wu et al. (2022) and Zhang et al. (2021) against in-situ LST under all-weather conditions at the AR, HZZ, and YK sites in 2010: (a) R, (b) RMSE.

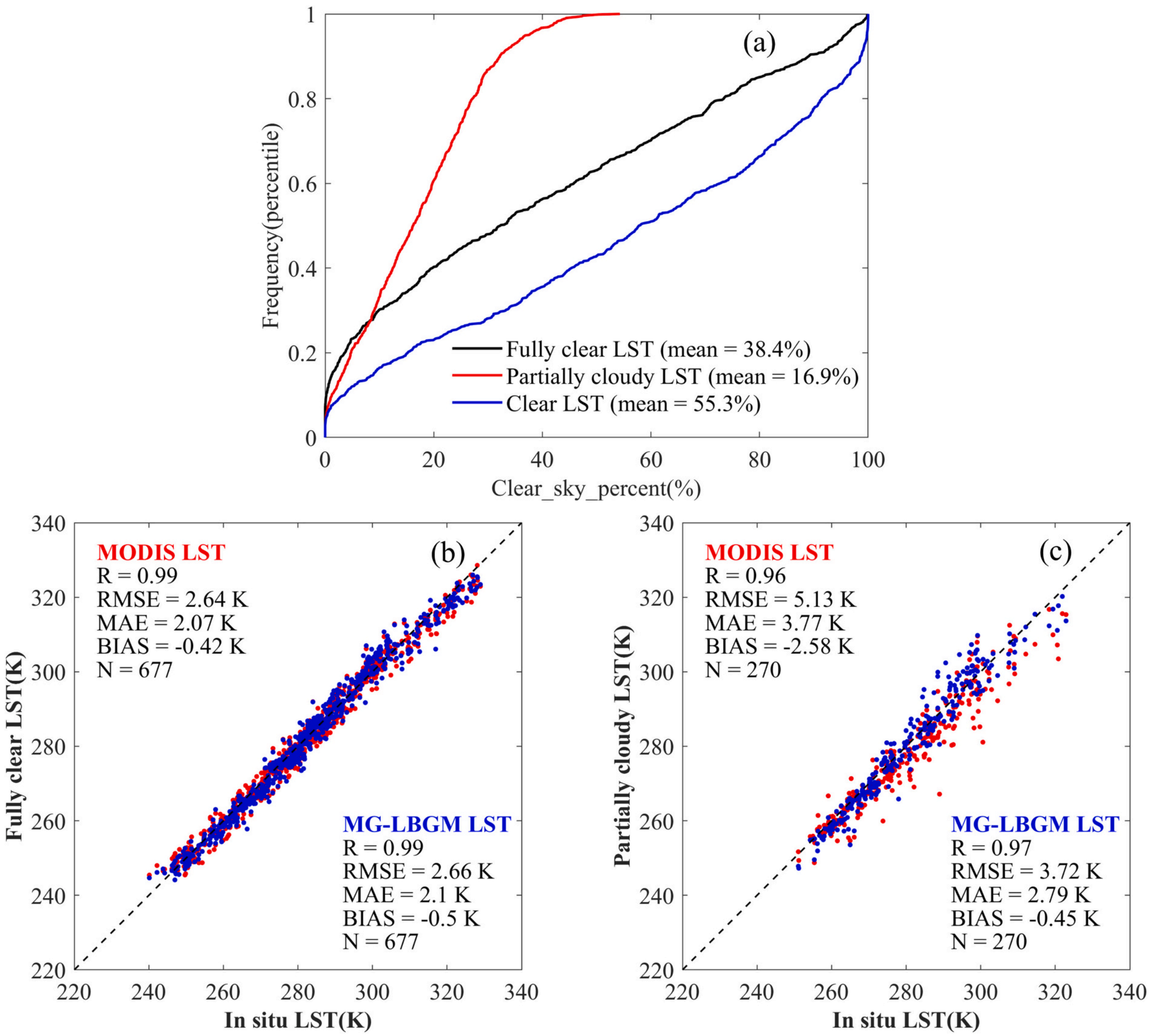


Fig. 13. (a) Cumulative distribution plots of the spatially averaged percentages of clear sky samples over the HRB-MU, and scatter plots of the official MODIS LST and MG-LGBM LST against in-situ LST measurements at the AR, HZZ, and YK sites for the daytime and nighttime in 2010 from (b) fully clear-sky pixels, and (c) partially cloud-covered pixels. Clear_sky percent denotes the proportion of clear-sky samples among all samples over the HRB-MU region in 2010.

Table 6
Sample-based validation results in the different seasons for the daytime and nighttime, respectively.

Time	Season	RMSE (K)	MAE (K)	R	N
Daytime	spring	2.3	1.71	0.98	4,021,766
	summer	2.1	1.53	0.98	3,907,006
	autumn	1.79	1.32	0.99	4,840,660
	winter	1.95	1.44	0.97	3,633,792
Nighttime	spring	1.25	0.93	0.99	6,967,453
	summer	1.03	0.73	0.99	6,400,269
	autumn	1.17	0.88	0.99	8,191,298
	winter	1.41	1.08	0.97	8,097,223

N denotes the sample size.

the estimated LST at different time scales (e.g., seasonal and monthly scales). Table 6 lists the sample-based validation results in the different seasons during the daytime and nighttime. The performance of the MG-LGBM model varies in the different seasons, with a relatively low accuracy observed in the spring and summer during the daytime. The reason for this may be that the vegetation grows and peaks in these two seasons, which brings uncertainty to the LST estimation (Li et al., 2021). A relatively high accuracy is obtained in autumn, with the RMSE <1.8 K and R of 0.99. In addition, the season impact during the daytime is slightly higher than that during the nighttime. The difference between the maximum and minimum RMSE reaches 0.51 K and 0.38 K for the daytime and nighttime, respectively.

To further evaluate the intraannual variation of MG-LGBM LST, Fig. 14 shows the model accuracies at monthly scale. Similar to the seasonal validation results, the MG-LGBM model's performance fluctuates across various months. During the daytime and nighttime, the smallest RMSEs of 1.73 K and 0.98 K were obtained in October and August, respectively, whereas the largest RMSEs of 2.33 K and 1.42 K were obtained in April and February, respectively. It was also observed that MG-LGBM model tends to slightly overestimate LST during the day, while underestimating it slightly at night. Overall, the impacts of season and month on the model accuracy exist, but are not significant, demonstrating the robust performance of the MG-LGBM model.

The heterogeneity of LST is significantly affected by different land-cover types (Holmes et al., 2009; Xu et al., 2022) and elevations (He et al., 2019; Zhao et al., 2019). Fig. 15 shows the performance of the MG-LGBM model by calculating the RMSE for different land-cover types and elevations. As shown in Fig. 15a, all the RMSE values are <2.27 K but varies across the different land-cover types, with the lowest value occurring in cropland and the highest value occurring in grassland

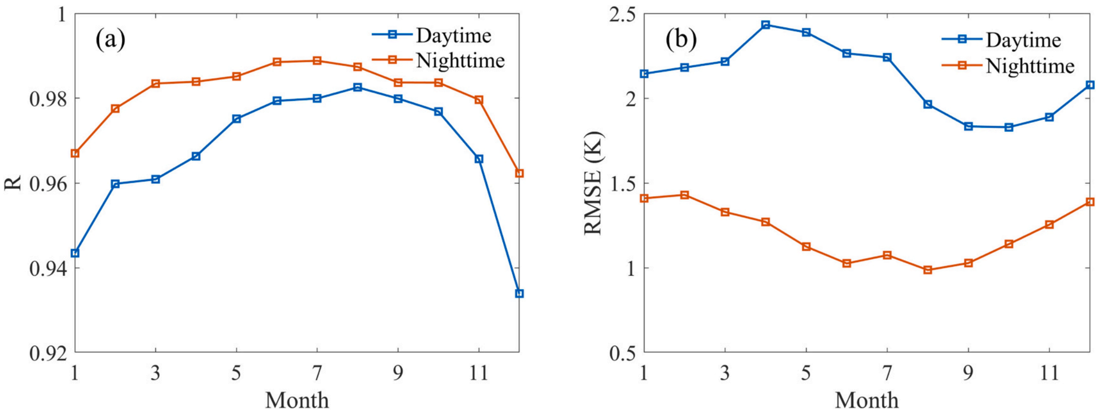


Fig. 14. Comparison of the MG-LGBM LST with MODIS LST at monthly scale: (a) RMSE (K) and (b) BIAS (K).

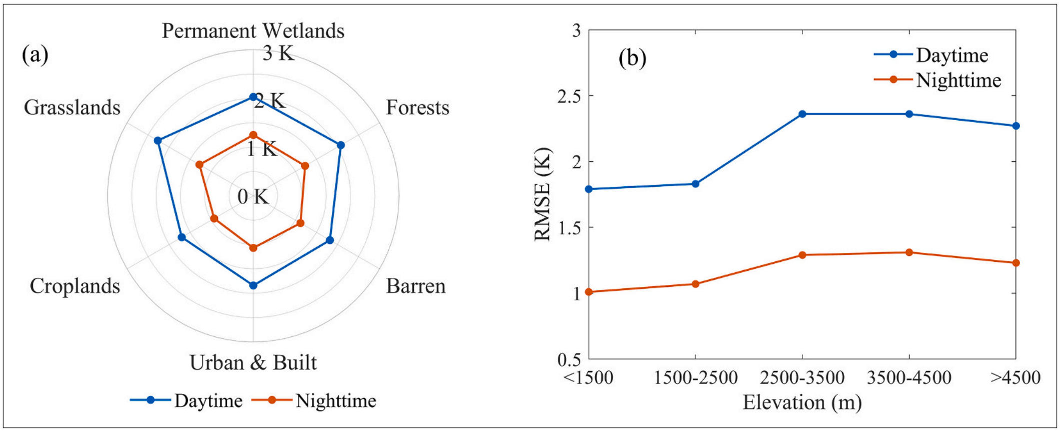


Fig. 15. Validation results (RMSE) of the MG-LGBM LST against MODIS LST in clear-sky conditions for specific (a) land-cover types and (b) elevation ranges.

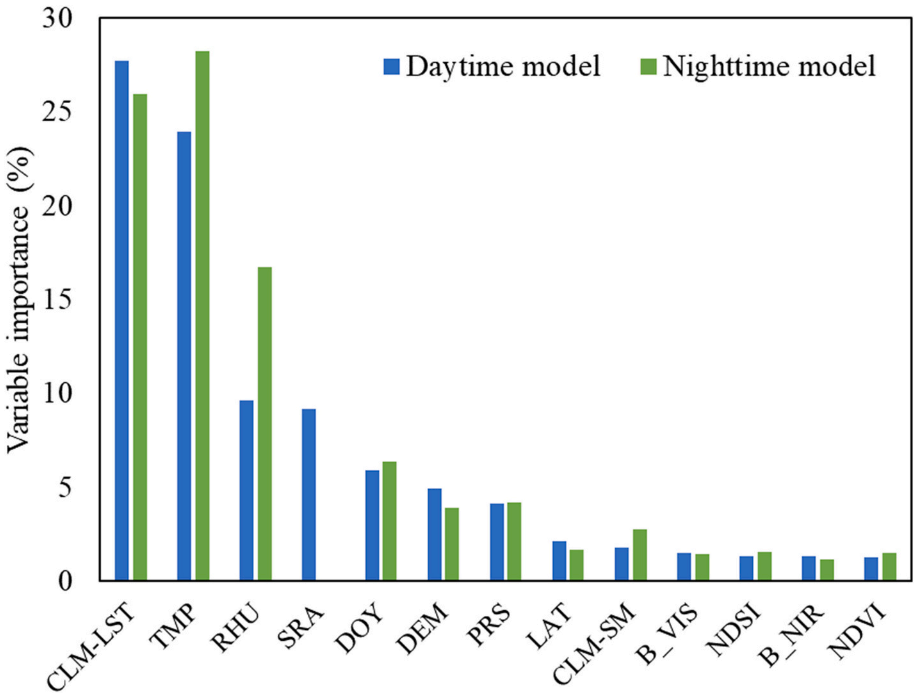


Fig. 16. Mean importance values of the input variables of the daytime model and nighttime model.

during both the daytime and nighttime. The difference between the maximum and minimum RMSE for the MG-LGBM model is 0.56 K and 0.35 K for the daytime and nighttime, respectively. These results suggest that the proposed method demonstrates both a high accuracy and stability for different land-cover types.

To evaluate the impact of topography on the model performance, the model accuracies for different elevation ranges are presented in Fig. 15b. In general, as the elevation increases, the accuracy of the model tends to decrease (Xu et al., 2022; Zhao et al., 2019), while relatively little change can be observed during the nighttime. However, there is a reverse trend when the elevation is above 3500–4500 m, which indicates that LST is comprehensively affected by different factors (Xu et al., 2022; Zhao and Duan, 2020). Nonetheless, the RMSE values remain stable, at <2.4 K and 1.3 K during the daytime and nighttime, respectively. This indicates that the MG-LGBM model shows a stable accuracy in estimating clear-sky LST across different elevation ranges.

5.2. Variable importance analysis

LST varies rapidly in space and time. The LST variation is affected by the surface conditions, topographic conditions, and atmospheric conditions (Crosson et al., 2012; Prata et al., 1995). In this study, we selected four sets of CLM meteorological forcing data (SRA, TMP, RHU, PRS), two sets of CLM simulation data (CLM-LST and CLM-SM), six sets of remote sensing data (DEM, LAT, NDVI, NDSI, B_VIS, B_NIR), and DOY, giving a total of 13 variables as the predictors. During the construction of the decision tree in the LGBM model, the collective benefits gained from the splits that utilize the feature of interest were used to measure the importance of the predictors. Fig. 16 displays the mean importance values of the predictors over the four seasons. The results reveal that CLM-LST and the four sets of meteorological forcing data are the main contributors to the model during both the daytime and nighttime, which confirms the SEB guidance theory in MG-LGBM. DOY and DEM also have important impacts on LST estimation, owing to their contributions in representing LST temporal variability and the terrain effect. In addition, SM is an essential factor contributing to the spatial heterogeneity of LST, due to its controlling impact on surface thermal inertia and evapotranspiration (Sandholt et al., 2002; Tang et al., 2010).

5.3. Further validation of the developed method in non-vegetated regions

As shown in Fig. 8b, CLM model exhibits uncertainty in simulating LST at a site dominated by sparse vegetation (i.e., HZZ). This uncertainty may be attributed to the incomplete parameterization schemes for turbulent flux calculation, which has been reported by other studies (Ma et al., 2021; Trigo et al., 2015; Zheng et al., 2014). Obtaining gapless LST with high accuracy in non-vegetated regions poses a challenge. Although the HZZ site is used for validation, such comparisons may not be comprehensive enough. To further validate the transferability of the proposed method in non-vegetated regions, we selected two typical barren-dominated sites, i.e., the Bajitan Gobi station (GB) and the Shenshawo desert station (SSW), from the Heihe Watershed Allied Telemetry Experimental Research (HiWATER)(Li et al., 2013a). The generated gapless LST from 2013 to 2015 over the HRB-MU region was then validated against in-situ LST.

Fig. 17 shows the comparison results between the estimated LST and in-situ LST measurements under all-weather conditions. The MG-LGBM estimated LST agree adequately with the in-situ LST under clear-sky conditions, with an RMSE of 2.14–2.43 K and R of 1, respectively. By extracting complex features from the space of input data, MG-LGBM can effectively recover LST under cloudy-sky conditions, with an RMSE ranging from 2.98 to 3.25 K and R ranging from 0.98 to 0.99, respectively. This indicates the effectiveness of transferring the proposed model to the non-vegetated regions.

5.4. Advantages and limitations of the proposed method

The advantages of the proposed method are reflected in the following aspects. Firstly, the previous studies usually applied the relationship of the clear-sky model to restore cloudy LST, without sufficient consideration of the cloud effects on LST (Buo et al., 2021; Xiao et al., 2021). Meanwhile, the linking predictors were typically temporally sparse (e.g., 3-h to 16-day), which resulted in uncertainty in presenting the high temporal variability of LST (Cho et al., 2022; Li et al., 2021). In response to this issue, the MG-LGBM model fully maximizes the potential of proxy data for clouds with a high spatio-temporal resolution (e.g., reanalysis data and model-simulated data with ~7 km and 1-h resolutions), which can discriminate between clear-sky and cloudy-sky conditions, resulting in a good and comparable accuracy in the estimated gapless LST under both clear-sky and cloudy-sky conditions (see Section 4.2).

Secondly, it is acknowledged that ML models suffer from "black box" problems, which has led to an absence of physical meaning in the reconstrued cloudy LST. Inspired by the advances in coupling process understanding and ML in ESS (Reichstein et al., 2019), we investigated the rationality and feasibility of incorporating explicit mechanistic guidance into an implicit ML model to improve LST prediction accuracy

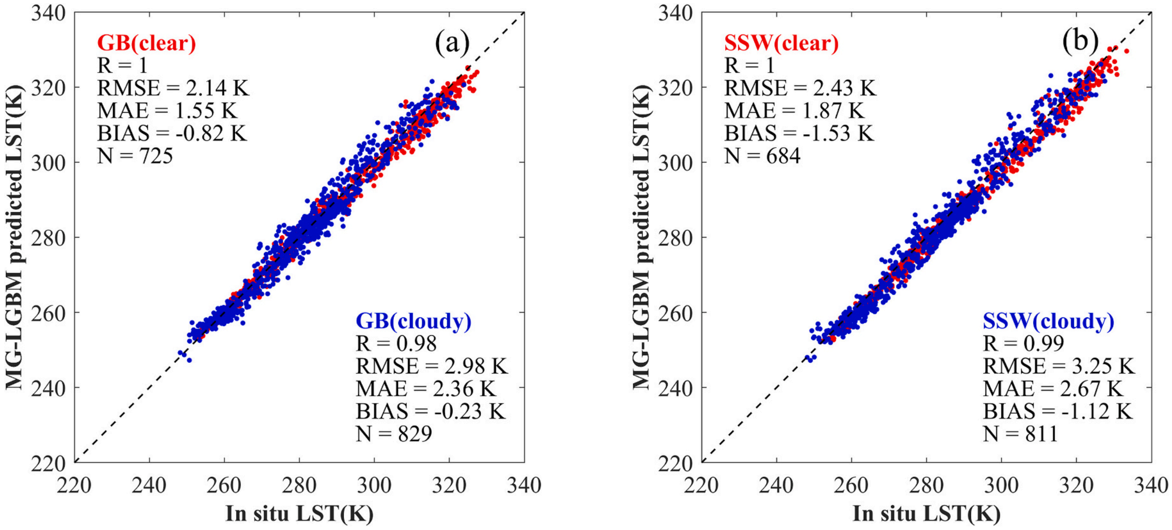


Fig. 17. Scatter plots of the estimated LST against in-situ LST measurements under clear-sky and cloudy-sky conditions during January 2013 to April 2015 at the two sites: (a) GB (gobi, 100.3042°E, 38.915°N), and (b) SSW (desert, 100.4933°E, 38.7892°N).

and mechanistic interpretability. Special attention was paid to the evaluation of uncertainty in different data combinations and the effects of mechanistic intervention. Compared with the pure ML models (e.g., RS, RS + CF, and RS + CS in Section 4.3.1), the MG-LGBM model (i.e., RS + CF + CS) achieved a superior performance. It was demonstrated that the underlying process knowledge (e.g., the SEB and the law of heat conduction, see Eq. (12)) between the key LSM forcing data and LSM simulation data in the CLM was mined and applied in the learning process, so that the learning model has relatively high interpretability. Meanwhile, ML models are known to have extrapolation problems, and coupling such mechanistic guidance helped to capture the LST values in regions without valid samples, thereby enhancing the generalization ability of the ML model (see Section 4.1).

Thirdly, benefiting from the utilization of the highest-quality MODIS LST and the powerful data mining ability of ML, the likely cloudy MODIS LST was corrected. Theoretically and practically, the LST obtained in this study achieved a comparable accuracy to the original MODIS LST in general clear-sky conditions (see Section 4.4). Lastly, the proposed method is currently applicable to the HRB-MU with CLDAS data. Given that all the remote sensing data used are gapless and worldwide, and provided that the LSM forcing data and simulation data are updated on a larger scale, or replaced with other global-scale data (e.g., GLDAS and European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5)), this method could be easily expanded to other regions. Moreover, the MG-LGBM model has excellent efficiency, so once the model is trained, it could be readily applied for generating gapless and long time-series LST products.

Nevertheless, there are also some uncertainties and limitations to this method. We compared our method with other excellent LST reconstruction methods, demonstrating a superior performance of the MG-LGBM model (see Section 4.3.3). However, this comparison was conducted at a regional scale, and further investigation is needed for a larger-scale (e.g., national or global scale) comparison. It is acknowledged that due to the relatively coarse spatial resolution of CLM-LST (approximately 5 km), the variability of CLM-LST may be less pronounced than that in MODIS LST, which may introduce some uncertainty in the estimation of 1-km LST under clouds. Future studies could consider improving the grid resolution of the CLM, to address the uncertainties arising from spatial resolution differences between CLM-LST and MODIS LST. Another uncertainty is the spatial representativeness of the ground sites. Although previous studies have evaluated the spatial representativeness of the sites in the WATER and HiWATER observatory networks and demonstrated that the sites used in the study were qualified for homogeneity validation (Yu et al., 2017; Yu et al., 2014a), the scaling effects induced by spatial mismatch cannot be totally neglected, which is a typical difficulty in remote sensing product validation (Duan et al., 2019; Li et al., 2020a).

In this study, LSM-related data were used as a proxy to reflect cloud effects. However, the true “cloudy” label data were still unavailable. Further studies are advocated to incorporate additional ground-based measurements or PMW information into the learning-based model. Additionally, alternative radiative transfer and energy balance models, such as Soil Canopy Observation, Photochemistry and Energy fluxes model (SCOPE)(Van der Tol et al., 2009), Discrete Anisotropic Radiative Transfer energy budget model(DART-EB)(Gastellu-Etchegorry, 2008), and Thermal Radiosity-Graphics Combined energy budget model (TRGM-EB)(Bian et al., 2017) could be further explored to provide extra mechanism-guided simulation data over urban and natural areas. This consideration could extend the proposed method to finer-scale LST data, such as Landsat and the ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS). Lastly, we mainly verified the effectiveness of coupling the input variable mechanistic guidance into the ML model in this study (Shen and Zhang, 2023). In future work, more attention should be paid to further investigating the different guidance construction methods, such as objective function guidance (e.g., imposing model penalties in the cost function for violating the SEB) and model structure guidance (e.g., adding a mechanistic relationship expressing the SEB to the middle or back of the neural network).

6. Conclusion

Cloud contamination has hindered the access to gapless LST for potential applications. In this paper, we have proposed a mechanism-guided ML model (MG-LGBM) by integrating key CLM forcing data, CLM simulation data and remote sensing data into an ML model. The MG-LGBM model combines the advantages of process knowledge (SEB mechanistic guidance, interpretability) and ML (data adaptability). The well-trained model was applied to instantaneous observations from Aqua MODIS to generate a daily $0.01^\circ \times 0.01^\circ$ gapless LST dataset from 2008 to 2011 over the HRB-MU. The generated LST showed excellent spatio-temporal continuity and retained high consistency with the MODIS LST in spatial details. In terms of temporal variability, the generated LST can basically capture the seasonal and daily changes of LST. In addition, the MG-LGBM model showed a robust performance across different time scales, land-cover types, and elevation ranges. The accuracy of the clear-sky LST generated by the MG-LGBM model was similar to that of the original MODIS LST under fully clear-sky conditions. Moreover, MG-LGBM can correct for the likely-cloudy LST pixels under partially cloudy conditions. The reconstructed cloudy LST against in-situ measurements showed a satisfactory accuracy (RMSE = 2.91–3.66 K, MAE = 2.35–3 K, $R = 0.97$–0.98), which was generally comparable to that obtained under clear-sky conditions. Among the four popular ML methods, the LGBM-based model performs the best in model accuracy and time efficiency. Compared with a pure mechanistic method and pure ML methods, the MG-LGBM model improved the LST prediction accuracy, mechanistic interpretability, and generalization ability of the LST estimation. When compared with other gapless LST data, it was found that the MG-LGBM LST was more “real” and “natural” and had a higher accuracy over the HRB-MU region. In our future work, it is anticipated to expand the proposed model to a larger-scale study. It should further address the spatial transferability issue in ML. One feasible solution is to partition the study area into multiple subregions based on geographical features and then perform the reconstructions separately in each subregion, finally merge them and generate fused estimates at a national or global scale, as in the research into Shen et al. (2020).

CRediT authorship contribution statement

Jun Ma: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Huanfeng Shen:** Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Menghui Jiang:** Writing – review & editing, Visualization. **Liupeng Lin:** Writing – review & editing, Visualization. **Chunlei Meng:** Writing – review & editing, Visualization. **Chao Zeng:** Writing – review & editing, Visualization. **Huifang Li:** Writing – review & editing, Visualization. **Penghai Wu:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgements

This study was jointly supported by the Key Projects of the National Natural Science Foundation of China under grant 42130108, the Special Fund of Hubei Luojia Laboratory under grant 220100041, and the

National Natural Science Foundation of China under grant 42271381. The anonymous reviewers' and editors' comments were highly appreciated. The MODIS products were accessed from the Earthdata website (https://search.earthdata.nasa.gov/search). The GLASS products were downloaded from http://www.glass.umd.edu/. The soil data were downloaded at http://globalchange.bnu.edu.cn/. The in-situ data were provided by the National Tibetan Plateau Data Center (http://data.tpdc.ac.cn). The CLDAS forcing data were provided by the China Meteorological Administration (http://tipex.data.cma.cn/tipex). The Zhang I's LST was obtained from https://doi.org/10.25380/iastate.c.5078492. The Zhang II's LST was obtained from https://data.tpdc.ac.cn. The Xu's LST was obtained from http://www.geodata.cn. The Wu's LST was provided by Wu et al. (2022). The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

Appendix A. Symbols and corresponding meanings in the methodology

Symbol (in the order of appearance)	Meaning
LST	Label data (i.e., MODIS LST) with the highest quality
ML	Pure ML model
RS	Remote sensing data, including DEM, LAT, NDVI, NDSI, B_VIS, B_NIR, DOY
$L^{(t)}$	Objective function of the $t-th$ iteration solution in the LGBM model
x_i	Sample in the LGBM model
n	Number of samples in the LGBM model
l	Loss function in the LGBM model
y_i	Actual value in the LGBM model
$\hat{y}$	Predicted value of the LGBM model
∂	First-order gradient of l in the LGBM model
∂^2	Second-order gradient of l in the LGBM model
f_t	Independent regression tree in the LGBM model
$f_t(x)$	Corresponding increment of f_t in the LGBM model
$\Omega(f)$	Regularization term in the LGBM model
T	Number of leaf nodes in f_t in the LGBM model
w	Leaf weights of f_t in the LGBM model
γ	Hyper-parameter of $\Omega(f)$ in the LGBM model
λ	Hyper-parameter of $\Omega(f)$ in the LGBM model
$MG-LGBM$	Mechanism-guided LGBM model
CF	CLM forcing data that are mechanistically dominant in CLM-LST simulation, including SRA, TMP, RHU, PRS
CS	CLM simulation data, including CLM-LST, CLM-SM
Δz	Layer thickness between two layers in the CLM model
T_g	Soil temperature at the first layer in the CLM model
T_2	Soil temperature at the second layer in the CLM model
$\overrightarrow{S}_g$	Net solar radiation absorbed by the ground
$\overrightarrow{L}_g$	Net longwave radiation absorbed by the ground
H_g	Sensible heat flux
λE_g	Latent heat flux
T_s	Skin temperature (LST)
$L\uparrow$	Surface upward longwave radiation
L	Exposed leaf area index
S	Exposed stem area index
T_v	Vegetation temperature
σ	Stefan-Boltzmann constant
ε_g	Ground emissivity
ε_v	Vegetation emissivity
$L_{atm}\downarrow$	Downward atmospheric longwave radiation
e_{atm}	Atmospheric vapor pressure
P_{atm}	Air pressure (PRS)
q_{atm}	Air specific humidity (RHU)
T_{atm}	Air temperature (TMP)
G	Ground heat flux
E_g	Water vapor flux
λ	Multiplier for converting the water vapor flux to an energy flux
$R_{n,g}$	Surface net radiation
$S\downarrow$	Incident solar shortwave radiation
α	Ground albedo
c	Volumetric heat capacity of snow/soil
t	Time step
t_k	Soil property
z	Vertical direction depth

Appendix B. Abbreviation index

Term (in alphabetical order)	Abbreviation
Advanced Microwave Scanning Radiometer-2	AMSR-2
Advanced Microwave Scanning Radiometer for EOS	AMSR-E
Advanced Very High-resolution Radiometer	AVHRR
Air pressure	PRS
Air temperature	TMP
Annual temperature cycle	ATC
Black-sky albedo in near-infrared spectral range	B_NIR
Black-sky albedo in visible spectral range	B_VIS
broadband emissivity	BBE
China Meteorological Administration Land Data Assimilation System	CLDAS
Community Earth System Model	CESM
Community Land Model	CLM
Community Land Model simulated land surface temperature	CLM-LST
Community Land Model simulated soil moisture	CLM-SM
Cross-validation	CV
Cumulative distribution function	CDF
Discrete Anisotropic Radiative Transfer energy budget model	DART-EB
Day of year	DOY
Deep belief network	DBN
Digital elevation model	DEM
Diurnal temperature cycle	DTC
Downwelling longwave radiation	DLW
Earth system science	ESS
ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station	ECOSTRESS
European Centre for Medium-Range Weather Forecasts Reanalysis v5	ERA5
Exclusive Feature Bundling	EFB
Extreme gradient boosting	XGBoost
Generalized regression neural network	GRNN
Global Land Data Assimilation System	GLDAS
Global LAnd Surface Satellite	GLASS
Gradient-based One-Side Sampling	GOSS
Land surface model	LSM
Land surface temperature	LST
Latitude	LAT
Machine learning	ML
Mean absolute error	MAE
Middle and upper reaches of Heihe River basin	HRB-MU
Moderate Resolution Imaging Spectroradiometer	MODIS
Multiresolution Kalman filtering	MKF
Normalized difference snow index	NDSI
Normalized difference vegetation index	NDVI
Overall bias	BIAS
Passive microwave	PMW
Pearson correlation coefficient	R
Mechanism-guided Light gradient boosting model	MG-LGBM
Precipitation rate	PRE
Quality control	QC
Root-mean-square error	RMSE
Soil Canopy Observation, Photochemistry and Energy fluxes model	SCOPE
Soil moisture	SM
Solar cloud satellite geometry	SCSG
Solar shortwave radiation	SRA
Specific humidity	RHU
Spinning Enhanced Visible and Infrared Imager	SEVIRI
Surface energy balance	SEB
Thermal Radiosity-Graphics Combined and energy budget model	TRGM-EB
Thermal infrared	TIR
Upwelling longwave radiation	ULW
Visible Infrared Imaging Radiometer	VIIRS
Wind speed	WIN

References

Aldrich, J., 1995. Correlations genuine and spurious in Pearson and Yule. Stat. Sci. 364–376.

Altman, N., Krzywinski, M., 2015. Points of significance: association, correlation and causation. Nat. Methods 12.

Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. Remote Sens. Environ. 122, 50–65.

Bian, Z., Du, Y., Li, H., Cao, B., Huang, H., Xiao, Q., Liu, Q., 2017. Modeling the temporal variability of thermal emissions from row-planted scenes using a radiosity and energy budget method. IEEE Trans. Geosci. Remote Sens. 55, 6010–6026.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Buo, I., Sagris, V., Jaagus, J., 2021. Gap-filling satellite land surface temperature over heatwave periods with machine learning. IEEE Geosci. Remote Sens. Lett. 19, 1–5.

Cao, B., Liu, Q., Du, Y., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Trigo, I.F., Zhan, W., Yu, Y., Cheng, J., Jacob, F., 2019. A review of earth surface thermal radiation directionality observing and modeling: historical development, current status and perspectives. Remote Sens. Environ. 232, 111304.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

Chen, J., Shen, H., Li, X., Li, T., Wei, Y., 2022. Ground-level ozone estimation based on geo-intelligent machine learning by fusing in-situ observations, remote sensing data, and model simulation data. Int. J. Appl. Earth Obs. Geoinf. 112, 102955.

Cheng, J., Liang, S., Verhoef, W., Shi, L., Liu, Q., 2015. Estimating the hemispherical broadband longwave emissivity of global vegetated surfaces using a radiative transfer model. IEEE Trans. Geosci. Remote Sens. 54, 905–917.

Cho, D., Bae, D., Yoo, C., Im, J., Lee, Y., Lee, S., 2022. All-sky 1 km MODIS land surface temperature reconstruction considering cloud effects based on machine learning. Remote Sens. 14, 1815.

Connors, J.P., Galletti, C.S., Chow, W.T., 2013. Landscape configuration and urban heat island effects: assessing the relationship between landscape characteristics and land surface temperature in Phoenix, Arizona. Landsc. Ecol. 28, 271–283.

Crosson, W.L., Al-Hamdan, M.Z., Hemmings, S.N., Wade, G.M., 2012. A daily merged MODIS Aqua–Terra land surface temperature data set for the conterminous United States. Remote Sens. Environ. 119, 315–324.

De Bézenac, E., Pajot, A., Gallinari, P., 2019. Deep learning for physical processes: incorporating prior scientific knowledge. J. Stat. Mech. 2019, 124009.

Deng, M., Meng, X., Lyv, Y., Zhao, L., Li, Z., Hu, Z., Jing, H., 2020. Comparison of soil water and heat transfer modeling over the Tibetan plateau using two community land surface model (CLM) versions. J. Adv. Model. Earth Syst. 12 (e2020MS002189).

Duan, S.-B., Li, Z.-L., Leng, P., 2017. A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. Remote Sens. Environ. 195, 107–117.

Duan, S.-B., Li, Z.-L., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., Coll, C., 2019. Validation of collection 6 MODIS land surface temperature product using in situ measurements. Remote Sens. Environ. 225, 16–29.

Fisher, R.A., Koven, C.D., 2020. Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. J. Adv. Model. Earth Syst. 12 (e2018MS001453).

Fu, P., Weng, Q., 2016. Consistent land surface temperature data generation from irregularly spaced Landsat imagery. Remote Sens. Environ. 184, 175–187.

Fu, P., Xie, Y., Weng, Q., Myint, S., Meacham-Hensold, K., Bernacchi, C., 2019. A physical model-based method for retrieving urban land surface temperatures under cloudy conditions. Remote Sens. Environ. 230, 111191.

Gastellu-Etchegorry, J.-P., 2008. 3D modeling of satellite spectral images, radiation budget and energy budget of urban landscapes. Meteorog. Atmos. Phys. 102, 187–207.

Gong, Y., Li, H., Shen, H., Meng, C., Wu, P., 2023. Cloud-covered MODIS LST reconstruction by combining assimilation data and remote sensing data through a nonlocality-reinforced network. Int. J. Appl. Earth Obs. Geoinf. 117, 103195.

Göttsche, F.-M., Olesen, F.S., 2001. Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data. Remote Sens. Environ. 76, 337–348.

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. Rev. Geophys. 48.

He, J., Zhao, W., Li, A., Wen, F., Yu, D., 2019. The impact of the terrain effect on land surface temperature variation based on Landsat-8 observations in mountainous areas. Int. J. Remote Sens. 40, 1808–1827.

Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput. 18, 1527–1554.

Holmes, T., De Jeu, R., Owe, M., Dolman, A., 2009. Land surface temperature from Ka band (37 GHz) passive microwave observations. J. Geophys. Res. Atmos. 114.

Hong, F., Zhan, W., Goettsche, F.-M., Lai, J., Liu, Z., Hu, L., Fu, P., Huang, F., Li, J., Li, H., 2021. A simple yet robust framework to estimate accurate daily mean land surface temperature from thermal observations of tandem polar orbiters. Remote Sens. Environ. 264, 112612.

Idso, S.B., 1981. A set of equations for full spectrum and 8-to 14-μm and 10.5-to 12.5-μm thermal radiation from cloudless skies. Water Resour. Res. 17, 295–304.

Jia, A., Ma, H., Liang, S., Wang, D., 2021. Cloudy-sky land surface temperature from VIIRS and MODIS satellite data using a surface energy balance-based method. Remote Sens. Environ. 263, 112566.

Jia, A., Liang, S., Wang, D., 2022. Generating a 2-km, all-sky, hourly land surface temperature product from advanced baseline imager data. Remote Sens. Environ. 278, 113105.

Jin, M., 2000. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 2. Cloudy-pixel treatment. J. Geophys. Res. Atmos. 105, 4061–4076.

Jing, Y., Li, X., Shen, H., 2022. STAR NDSI collection: a cloud-free MODIS NDSI dataset (2001−2020) for China. Earth Syst. Sci. Data 14, 3137–3156.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. 3, 422–440.

Karnieli, A., Agam, N., Pinker, R.T., Anderson, M., Imhoff, M.L., Gutman, G.G., Panov, N., Goldberg, A., 2010. Use of NDVI and land surface temperature for drought assessment: merits and limitations. J. Clim. 23, 618–633.

Karpatne, A., Watkins, W., Read, J., Kumar, V., . Physics-Guided Neural Networks (pgnn): An Application in Lake Temperature Modeling. *arXiv preprint. arXiv:1710.11431*, *2*.

Ke, L., Ding, X., Song, C., 2013. Reconstruction of time-series MODIS LST in Central Qinghai-Tibet plateau using geostatistical approach. IEEE Geosci. Remote Sens. Lett. 10, 1602–1606.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural Inf. Proces. Syst. 30.

Kilibarda, M., Hengl, T., Heuvelink, G.B., Gräler, B., Pebesma, E., Perčec Tadić, M., Bajat, B., 2014. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. J. Geophys. Res. Atmos. 119, 2294–2313.

Koppa, A., Rains, D., Hulsman, P., Poyatos, R., Miralles, D.G., 2022. A deep learning-based hybrid model of global terrestrial evaporation. Nat. Commun. 13, 1912.

Lawrence, D.M., Fisher, R.A., Koven, C.D., Oleson, K.W., Swenson, S.C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., 2019. The community land model version 5: description of new features, benchmarking, and impact of forcing uncertainty. J. Adv. Model. Earth Syst. 11, 4245–4287.

Li, X., Li, X., Li, Z., Ma, M., Wang, J., Xiao, Q., Liu, Q., Che, T., Chen, E., Yan, G., 2009. Watershed allied telemetry experimental research. J. Geophys. Res. Atmos. 114.

Li, X., Cheng, G., Liu, S., Xiao, Q., Ma, M., Jin, R., Che, T., Liu, Q., Wang, W., Qi, Y., 2013a. Heihe watershed allied telemetry experimental research (HiWATER): scientific objectives and experimental design. Bull. Am. Meteorol. Soc. 94, 1145–1160.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., 2013b. Satellite-derived land surface temperature: current status and perspectives. Remote Sens. Environ. 131, 14–37.

Li, X., Zhou, Y., Asrar, G.R., Zhu, Z., 2018. Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous United States. Remote Sens. Environ. 206, 84–97.

Li, H., Yang, Y., Li, R., Wang, H., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., Liu, Q., 2019. 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 Sens. 57, 8081–8094.

Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., Liu, Q., 2020a. 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 Sens. 59, 1794–1807.

Li, T., Shen, H., Yuan, Q., Zhang, L., 2020b. Geographically and temporally weighted neural networks for satellite-based mapping of ground-level PM2. 5. ISPRS J. Photogramm. Remote Sens. 167, 178–188.

Li, T., Shen, H., Zeng, C., Yuan, Q., 2020c. A validation approach considering the uneven distribution of ground stations for satellite-based PM 2.5 estimation. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 13, 1312–1321.

Li, B., Liang, S., Liu, X., Ma, H., Chen, Y., Liang, T., He, T., 2021. Estimation of all-sky 1 km land surface temperature over the conterminous United States. Remote Sens. Environ. 266, 112707.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., 2021. The global land surface satellite (GLASS) product suite. Bull. Am. Meteorol. Soc. 102, E323–E337.

Lin, L., Shen, Y., Wu, J., Nan, F., 2023. CAFE: a cross-attention based adaptive weighting fusion network for MODIS and landsat spatiotemporal fusion. IEEE Geosci. Remote Sens. Lett.

Liu, S., Li, X., Xu, Z., Che, T., Xiao, Q., Ma, M., Liu, Q., Jin, R., Guo, J., Wang, L., 2018. The Heihe integrated observatory network: a basin-scale land surface processes observatory in China. Vadose Zone J. 17, 1–21.

Liu, J., Shi, C., Sun, S., Liang, J., Yang, Z.-L., 2019. Improving land surface hydrological simulations in China using CLDAS meteorological forcing data. J. Meteorol. Res. 33, 1194–1206.

Long, D., Bai, L., Yan, L., Zhang, C., Yang, W., Lei, H., Quan, J., Meng, X., Shi, C., 2019. Generation of spatially complete and daily continuous surface soil moisture of high spatial resolution. Remote Sens. Environ. 233, 111364.

Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., Yang, H., Tian, F., Zeng, C., Meng, X., 2020. Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach. Remote Sens. Environ. 246, 111863.

Lu, L., Venus, V., Skidmore, A., Wang, T., Luo, G., 2011. Estimating land-surface temperature under clouds using MSG/SEVIRI observations. Int. J. Appl. Earth Obs. Geoinf. 13, 265–276.

Luo, Q., Wen, J., Hu, Z., Lu, Y., Yang, X., 2020. Parameter sensitivities of the community land model at two alpine sites in the Three-River source region. J. Meteorol. Res. 34.

Ma, J., Zhou, J., Göttsche, F.-M., Liang, S., Wang, S., Li, M., 2020. A global long-term (1981–2000) land surface temperature product for NOAA AVHRR. Earth Syst. Sci. Data 12, 3247–3268.

Ma, X., Jin, J., Zhu, L., Liu, J., 2021. Evaluating and improving simulations of diurnal variation in land surface temperature with the community land model for the Tibetan plateau. PeerJ 9, e11040.

Ma, J., Shen, H., Wu, P., Wu, J., Gao, M., Meng, C., 2022. Generating gapless land surface temperature with a high spatio-temporal resolution by fusing multi-source satellite-observed and model-simulated data. Remote Sens. Environ. 278, 113083.

Minder, J.R., Mote, P.W., Lundquist, J.D., 2010. Surface temperature lapse rates over complex terrain: lessons from the Cascade Mountains. J. Geophys. Res. Atmos. 115.

Monin, A.S., Obukhov, A.M., 1954. Basic laws of turbulent mixing in the surface layer of the atmosphere. Contrib. Geophys. Inst. Acad. Sci. USSR 151, e187.

Neteler, M., 2010. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. Remote Sens. 2, 333–351.

Oleson, K.W., Bonan, G.B., Feddema, J., Vertenstein, M., Kluzek, E., 2010. Technical Description of an Urban Parameterization for the Community Land Model (CLMU). NCAR, Boulder.

Pearson, R.K., 2002. Outliers in process modeling and identification. IEEE Trans. Control Syst. Technol. 10, 55–63.

Prata, A., Caselles, V., Coll, C., Sobrino, J., Ottle, C., 1995. Thermal remote sensing of land surface temperature from satellites: current status and future prospects. Remote Sens. Rev. 12, 175–224.

Read, J.S., Jia, X., Willard, J., Appling, A.P., Zwart, J.A., Oliver, S.K., Karpatne, A., Hansen, G.J., Hanson, P.C., Watkins, W., 2019. Process-guided deep learning predictions of lake water temperature. Water Resour. Res. 55, 9173–9190.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven earth system science. Nature 566, 195–204.

Rodriguez, J.D., Perez, A., Lozano, J.A., 2009. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32, 569–575.

Sandholt, I., Rasmussen, K., Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens. Environ. 79, 213–224.

Shang, K., Yao, Y., Di, Z., Jia, K., Zhang, X., Fisher, J.B., Chen, J., Guo, X., Yang, J., Yu, R., 2023. Coupling physical constraints with machine learning for satellite-derived evapotranspiration of the Tibetan plateau. Remote Sens. Environ. 289, 113519.

Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., Ji, D., Ye, A., Yuan, H., Zhang, Q., 2013. A China data set of soil properties for land surface modeling. J. Adv. Model. Earth Syst. 5, 212–224.

Shen, H., Zhang, L., 2023. Mechanism-learning coupling paradigms for parameter inversion and simulation in earth surface systems. Sci. China Earth Sci. 66, 568–582.

Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L., 2015. Missing information reconstruction of remote sensing data: a technical review. IEEE Geosci. Remote Sens. Mag. 3, 61–85.

Shen, H., Huang, L., Zhang, L., Wu, P., Zeng, C., 2016. Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: a 26-year case study of the city of Wuhan in China. Remote Sens. Environ. 172, 109–125.

Shen, H., Jiang, Y., Li, T., Cheng, Q., Zeng, C., Zhang, L., 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens. Environ. 240, 111692.

Shen, H., Wang, Y., Guan, X., Huang, W., Chen, J., Lin, D., Gan, W., 2022. A spatiotemporal constrained machine learning method for OCO-2 solar-induced chlorophyll fluorescence (SIF) reconstruction. IEEE Trans. Geosci. Remote Sens. 60, 1–17.

Shi, C., Xie, Z., Qian, H., Liang, M., Yang, X., 2011. China land soil moisture EnKF data assimilation based on satellite remote sensing data. Sci. China Earth Sci. 54, 1430–1440.

Shwetha, H., Kumar, D.N., 2016. Prediction of high spatio-temporal resolution land surface temperature under cloudy conditions using microwave vegetation index and ANN. ISPRS J. Photogramm. Remote Sens. 117, 40–55.

Specht, D.F., 1991. A general regression neural network. IEEE Trans. Neural Netw. 2, 568–576.

Tan, J., Che, T., Wang, J., Liang, J., Zhang, Y., Ren, Z., 2021a. Reconstruction of the daily MODIS land surface temperature product using the two-step improved similar pixels method. Remote Sens. 13, 1671.

Tan, W., Wei, C., Lu, Y., Xue, D., 2021b. Reconstruction of all-weather daytime and nighttime MODIS aqua-terra land surface temperature products using an XGBoost approach. Remote Sens. 13, 4723.

Tang, R., Li, Z.-L., Tang, B., 2010. An application of the Ts–VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: implementation and validation. Remote Sens. Environ. 114, 540–551.

Trigo, I., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., Sandu, I., 2015. Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling. J. Geophys. Res. Atmos. 120 (12,096-012,111).

Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., Su, Z., 2009. An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. Biogeosciences 6, 3109–3129.

Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45.

Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 34, 892–905.

Wang, K., Liang, S., 2009. Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. Remote Sens. Environ. 113, 1556–1565.

Wang, T., Shi, J., Ma, Y., Husi, L., Comyn-Platt, E., Ji, D., Zhao, T., Xiong, C., 2019. Recovering land surface temperature under cloudy skies considering the solar-cloud-satellite geometry: application to MODIS and Landsat-8 data. J. Geophys. Res. Atmos. 124, 3401–3416.

Wang, H., Mao, K., Yuan, Z., Shi, J., Cao, M., Qin, Z., Duan, S., Tang, B., 2021. A method for land surface temperature retrieval based on model-data-knowledge-driven and deep learning. Remote Sens. Environ. 265, 112665.

Wang, Y., Shi, L., Hu, X., Song, W., Wang, L., 2023. Multiphysics-informed neural networks for coupled soil hydrothermal modeling. Water Resour. Res. (e2022WR031960).

Weiss, D.J., Atkinson, P.M., Bhatt, S., Mappin, B., Hay, S.I., Gething, P.W., 2014. An effective approach for gap-filling continental scale remotely sensed time-series. ISPRS J. Photogramm. Remote Sens. 98, 106–118.

Wu, P., Yin, Z., Yang, H., Wu, Y., Ma, X., 2019. Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network. Remote Sens. 11, 300.

Wu, P., Yin, Z., Zeng, C., Duan, S.-B., Göttsche, F.-M., Ma, X., Li, X., Yang, H., Shen, H., 2021. Spatially continuous and high-resolution land surface temperature product generation: a review of reconstruction and spatiotemporal fusion techniques. IEEE Geosci. Remote Sens. Mag. 9, 112–137.

Wu, P., Su, Y., Duan, S.-B., Li, X., Yang, H., Zeng, C., Ma, X., Wu, Y., Shen, H., 2022. A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data. Remote Sens. Environ. 277, 113070.

Xia, H., Chen, Y., Gong, A., Li, K., Liang, L., Guo, Z., 2021. Modeling daily temperatures via a phenology-based annual temperature cycle model. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 14, 6219–6229.

Xiao, Y., Zhao, W., Ma, M., He, K., 2021. Gap-free LST generation for MODIS/Terra LST product using a random forest-based reconstruction method. Remote Sens. 13, 2828.

Xiao, Y., Zhao, W., Ma, M., Yu, W., Fan, L., Huang, Y., Sun, X., Lang, Q., 2023. An integrated method for the generation of Spatio-temporally continuous LST product with MODIS/Terra observations. IEEE Trans. Geosci. Remote Sens. 61, 1–14.

Xu, S., Cheng, J., 2021. A new land surface temperature fusion strategy based on cumulative distribution function matching and multiresolution Kalman filtering. Remote Sens. Environ. 254, 112256.

Xu, Y., Shen, Y., 2013. Reconstruction of the land surface temperature time series using harmonic analysis. Comput. Geosci. 61, 126–132.

Xu, F., Fan, J., Yang, C., Liu, J., Zhang, X., 2022. Reconstructing all-weather daytime land surface temperature based on energy balance considering the cloud radiative effect. Atmos. Res. 279, 106397.

Yang, F., Lu, H., Yang, K., Wang, W., Li, C., Han, M., Li, Y., 2017. Evaluation and comparison among multiple forcing data sets for precipitation and shortwave radiation over mainland China. Hydrol. Earth Syst. Sci. Discuss. 21, 1–32.

Yu, W., Ma, M., Wang, X., Geng, L., Tan, J., Shi, J., 2014a. Evaluation of MODIS LST products using longwave radiation ground measurements in the northern arid region of China. Remote Sens. 6, 11494–11517.

Yu, W., Ma, M., Wang, X., Tan, J., 2014b. Estimating the land-surface temperature of pixels covered by clouds in MODIS products. J. Appl. Remote. Sens. 8 (083525–083525).

Yu, W., Ma, M., Li, Z., Tan, J., Wu, A., 2017. New scheme for validating remote-sensing land surface temperature products with station observations. Remote Sens. 9, 1210.

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., Hong, Y., 2018. A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. ISPRS J. Photogramm. Remote Sens. 141, 30–45.

Zhan, W., Zhou, J., Ju, W., Li, M., Sandholt, I., Voogt, J., Yu, C., 2014. Remotely sensed soil temperatures beneath snow-free skin-surface using thermal observations from tandem polar-orbiting satellites: an analytical three-time-scale model. Remote Sens. Environ. 143, 1–14.

Zhang, X., Zhou, J., Liang, S., Chai, L., Wang, D., Liu, J., 2020. Estimation of 1-km all-weather remotely sensed land surface temperature based on reconstructed spatial-seamless satellite passive microwave brightness temperature and thermal infrared data. ISPRS J. Photogramm. Remote Sens. 167, 321–344.

Zhang, X., Zhou, J., Liang, S., Wang, D., 2021. A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature. Remote Sens. Environ. 260, 112437.

Zhang, T., Zhou, Y., Zhu, Z., Li, X., Asrar, G.R., 2022. A global seamless 1 km resolution daily land surface temperature dataset (2003−2020). Earth Syst. Sci. Data 14, 651–664.

Zhao, W., Duan, S.-B., 2020. Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data. Remote Sens. Environ. 247, 111931.

Zhao, W., Duan, S.-B., Li, A., Yin, G., 2019. A practical method for reducing terrain effect on land surface temperature using random forest regression. Remote Sens. Environ. 221, 635–649.

Zheng, D., van der Velde, R., Su, Z., Booij, M.J., Hoekstra, A.Y., Wen, J., 2014. Assessment of roughness length schemes implemented within the Noah land surface model for high-altitude regions. J. Hydrometeorol. 15, 921–937.