

Impacts of land surface temperature and ambient factors on near-surface air temperature estimation: A multisource evaluation using SHAP analysis

Songyang Li ^{a, b, *}, Man Sing Wong ^{a, b, *}, Rui Zhu ^c, Guoqiang Shi ^a, Jinxin Yang ^d

^a Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, PR China

^b Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, PR China

^c Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Singapore 138632, Republic of Singapore

^d School of Geography and Remote Sensing, Guangzhou University, Guangzhou, PR China

ARTICLE INFO

Keywords:

Near-surface air temperature

LST

LightGBM

SHAP

Influential Factor

Remote sensing

ABSTRACT

Near-surface air temperature (Ta) is a vital indicator depicting urban thermal environments and sustainability. Machine learning (ML) models have been increasingly adopted for Ta estimation. However, there is still an urgent need to investigate how daytime and nighttime Ta are impacted by multisource ambient physical and anthropogenic factors across various environments. To this end, geospatial datasets incorporating MODIS-derived land surface temperature and 29 ancillary factors were employed to estimate Ta from 292 stations in China using ML modeling (training: 2017–2020). The optimal LightGBM-based models outperformed and obtained testing RMSEs of 3.03 °C (daytime) and 2.64 °C (nighttime) in 2021. Distinct spatiotemporal patterns in stations' Ta prediction were observed, with coastal areas showing better daytime estimates and northern midtemperate regions exhibiting lower nighttime accuracy. Comprehensive and individual models-based SHapley Additive exPlanations (SHAP) interpretation highlights the importance of incorporating macroscale meteorological backgrounds and terrain-related variables for Ta estimation improvement, as well as the critical impact of local urban morphology and anthropogenic indicators. This study has the potential to offer suggestions on ambient factors for improving Ta modeling and future urban heat island-related planning within specific regional and local climatic contexts.

1. Introduction

1.1. Background

As one of the well-known expressions of climate-related issues, the urban heat island (UHI) effect, characterized by the temperature difference between urban and rural areas, has been extensively researched over the past decades (Oke, 1973). Central to understanding the UHI effect, near-surface air temperature (Ta) serves as a crucial indicator at the urban canopy layer and has been widely investigated in terms of its evolutionary patterns, associations with other physical and anthropogenic factors, and regulation methods (Doick et al., 2013; Khamchiangta & Dhakal, 2019; Stewart et al., 2021; Varentsov et al., 2018). Furthermore, variations in Ta and UHI intensity in urban regions are critically linked to health outcomes, including the incidence of cardiovascular and respiratory disease (Huang et al., 2019, 2020), dengue outbreaks

(Akhtar et al., 2016), mental health issues (Thompson et al., 2018), mortality change (Yadav et al., 2023), and air pollutant concentrations (Wang et al., 2021). Besides, given Ta's interactive correlation with anthropogenic activities, it also plays a vital role in energy consumption and related carbon emissions (Roxon et al., 2020; Santamouris et al., 2015). Consequently, analyzing the spatiotemporal patterns of Ta is a critical step in studying the UHI effect and its associated issues.

1.2. Ta and its influential factors

Currently, most of the standard Ta data are usually collected continuously by weather stations following the standards of the World Meteorological Organization (1996). Besides official weather stations, collaborative automatic weather station networks, such as the Community Weather Information Network (Co-WIN) in Hong Kong (Lee et al., 2018), also serve as additional data sources that can provide massive points-based Ta data for meteorological studies. Moreover, to

* Corresponding author.

E-mail address: Ls.charles@polyu.edu.hk (M.S. Wong).

Nomenclature	
2D	2-Dimensional
3D	3-Dimensional
BH	Building Height
BH _{masked}	Building Height among buildings-covered regions
BSF	Building Surface Fraction
distanceRoad	Distance to road
DOY	Day of the year
era5_albedo	Daily-aggregated local albedo from ERA5 climate reanalysis dataset
era5_net_sr	Net solar radiation from ERA5 climate reanalysis dataset
era5_precipitation	Precipitation from ERA5 climate reanalysis dataset
era5_pressure	Atmospheric pressure from ERA5 climate reanalysis dataset
era5_sr_down	Downwards shortwave radiation from ERA5 climate reanalysis dataset
GEE	Google Earth Engine
GISD30	Global 30m Impervious-Surface Dynamic Dataset
greeneryFraction	Surface fraction of greenery
GRIP4	Global patterns of current and future road infrastructure
ISD	Global Hourly – Integrated Surface Database
LCZ	Local Climate Zone
LST	Land Surface Temperature
LULC	Land Use/Land Cover
MAE	Mean Absolute Error
ML	Machine Learning
MT	Mid-Temperate zone
NDBI	Normalized Difference Built-up Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NT	North Temperate zone
NTL	NightTime Light
ntl_count	Count numbers of pixels with nighttime lights
ntl_mean	Mean value of nighttime light intensity
ntl_sum	Sum value of nighttime light intensity
OLS	Ordinary Least Squares
pop_density	Density of population
pop_sum	Sum value of the population
PT	PlaTeau climate zone
R ²	Coefficient of determination
RF	Random Forest
RMSE	Root Mean Standard Error
roadFraction	Surface fraction of the road
SDI	Shannon Diversity Index
SHAP	SHapley Additive exPlanations
SRTM	Shuttle Radar Topography Mission
ST	Sub-Tropical zone
SVF	Sky View Factor
SVF _{masked}	Sky view factor among non-building areas
Ta	Air temperature
TP	TroPical zone
UHI	Urban Heat Island
WT	Warm-Temperate zone
XGBoost	eXtreme Gradient Boosting

ensure spatial coverage of Ta data products, interpolation methods such as kriging or inverse distance weighting have been developed to model Ta spatially (Benavides et al., 2007; Cho et al., 2020; Kurtzman & Kadmon, 1999). However, these methods rely on the number of ground stations; the accuracy decreases where the weather stations are sparsely distributed (Chen et al., 2021). To tackle this problem, the latest Ta estimation methods are often integrated with remotely sensed or in-situ data to improve the accuracy of estimation by considering the heterogeneity of Ta.

Therefore, for better spatial retrieving of high-resolution Ta, various geospatial datasets had been investigated for Ta modeling. Remotely sensed land surface temperature (LST), the radiative skin temperature of the land surface (Khan et al., 2021), representing one of the most critical energy sources of near-surface Ta in the surface energy balance, was proved to have promising capability for Ta estimation via linear regression-based models (Azevedo et al., 2016; Zhu et al., 2013a) within single cities or small regions. However, accurately estimating Ta on larger scales is relatively more challenging, primarily due to the heterogeneity of numerous physical and anthropogenic factors influencing Ta, such as the atmospheric dynamics (Cho et al., 2020), local land use/land cover (LULC) (Balas et al., 2023), multidimensional urban morphology patterns (Yu et al., 2020), terrain conditions (Mutiibwa et al., 2015), and human activities across diverse social backgrounds (Mokhov et al., 2012; Peng et al., 2017). Consequently, these extensive factors, encompassing meteorological, macroclimatic, natural landscape, urban structural, and anthropogenic features, interact through complex static and dynamic processes (Wang et al., 2023a). However, most of the studies conducted in regional or global scales usually focus more on factors, such as regional meteorological dynamics or indicators representing land surface energy-balance physical processes, while research for city-level often examines features related to local 2D/3D urban landscapes or human activities (Tian et al., 2024; Zhu et al., 2013b). Therefore, to support the analysis of Ta and UHI effect, understanding the contributory performance of various factors under

varying scales and functional categories for Ta estimation is thus necessary.

1.3. Machine learning for Ta estimation

Various multivariate regression models have been applied in previous studies to estimate Ta using inputs of numerous influential factors. Compared to traditional generalized linear models, machine learning (ML) regression models can effectively combine the merits of multiple variables through hyperdimensional nonlinear processes and enhance the overall accuracy of Ta estimation. ML models such as support vector regression (Cho et al., 2020) and random forest (RF, Venter et al., 2020) have been utilized in previous studies to estimate Ta variations in daily averages or daily maxima and minima, as well as using temperature-vegetation index (TVX) method, or Deep Learning (DL) methods like UNet to generate spatially seamless daily minimum/maximum Ta datasets (Su et al., 2023; Zhu et al., 2013b). However, although several studies have been conducted to combine large-scale LULC-related and local urban morphological features into ML models for LST-based UHI intensity modeling (Zekar et al., 2023), more investigation is still required to explore the Ta estimation and its accuracy using ML methods supported by diverse regional and local variables.

The tree-based ML models, owing to their capacity to automatically handle multiple subsets of samples of critical features and tune for optimized predictive performance (Belgiu & Drăguț, 2016; Breiman, 1984), demonstrate higher usability and outperform traditional ML models (Wang et al., 2023a). In particular, the light gradient boosting machine (LightGBM), a tree-based ML model, was successfully employed to predict LST, in-door Ta using diverse variables (Laukkarinen & Vinha, 2024; Liu et al., 2024). The eXtreme Gradient Boosting (XGBoost), one of the cutting-edge trees-based ML models, has also shown outstanding applicability to multivariate data, high accuracy, and rapid modeling speed, indicating significant potential for Ta

estimation (Chen & Guestrin, 2016). Besides applicable predictive accuracy, the ability to calculate the relative importance of features in tree-based ML modeling facilitates the analysis of essential factors in Ta modeling under data-driven processes and can guide the Ta regulation strategies. In recent years, advanced ML and DL methods have been designed to retrieve Ta with high prediction accuracies (Su et al., 2023; Yao et al., 2023), whereas the complex structures make them difficult to be interpreted (Pichler & Hartig, 2023). Tree-based ML models, such as LightGBM, XGBoost, and RF, however, offer feasible solutions between applicable prediction accuracy, efficiency, and interpretability (Deng et al., 2024; Grinsztajn et al., 2022).

1.4. Data-driven interpretation among influential factors for Ta estimation

Ta is often influenced by different dominant processes in individual regions, resulting in significant patterns of importance among influential factors across different samples, indicating the need for innovative evaluation methods among variables' sensitive contributions in subsets. SHapley Additive exPlanations (SHAP), designed by Lundberg and Lee (2017), represents a novel method for explaining individual predictions based on the game theoretically optimal Shapley values. One of the merits of SHAP analysis is its capability to analyze the importance of input features within specific subsets through the combined calculation of SHAP values among specific samples, further revealing potential feature interactions under certain spatiotemporal domains (Wang et al., 2024). SHAP-based interpretable approaches have been utilized in various remote sensing-supported applications, including LULC classification (Temenos et al., 2023), water quality assessment (Zhu et al., 2022), objective detection (Kawauchi & Fuse, 2022), and flood prediction (Wang et al., 2023b), while it was still seldom used for evaluating the influential factors of air thermal environments. One of the studies in this research field was conducted by Zhu et al. (2025), which successfully combined RF and SHAP to interpret how building morphological indicators, street landscapes, vegetation and waterbodies impact the street LST in Nanjing. On the other hand, the patterns of daytime and nighttime Ta variations are different; since more studies are focusing on estimative analyses of daily maximum and minimum temperatures (Nikolaou et al., 2023), there is still a strong demand to analyze the impacts of variables in different scales or categories on Ta estimation at daytime and nighttime, respectively.

Moreover, climate zones, regions categorized by distinct climates using different climatic parameters, can serve as evaluation tools for Ta estimation (Zhang & Yan, 2014). Besides, the local climate zone (LCZ) system, which considers both 2-dimensional (2D) LULC information and 3-dimensional (3D) local landscape patterns, was designed to analyze urban climate issues (Stewart & Oke, 2012). LCZs can be mapped using remote sensing or GIS methods (Wang et al., 2018) and have been widely employed in UHI-related studies to investigate differences in thermal environments and their related factors across various built-up (LCZ-1 to LCZ-10) and land-covered regions (LCZ-A to LCZ-G). Therefore, integrating SHAP values into evaluating Ta estimation across different sample subsets, classified by climate zones and LCZs, could help unveil the underlying feature contributions and the interactions among various physical and anthropogenic factors within specific regions.

1.5. Contributions

Overall, for evaluating the diverse impacts of influential factors on Ta estimation, remotely sensed LST and a series of ancillary factors around the weather stations are selected as explanatory variables in this study. The study's contributions include: 1) developing ML regression models using LST and multi-scale and multi-category factors for Ta estimation in China during daytime and nighttime, respectively; 2) evaluating the patterns and accuracy of Ta estimation under specific

regional and local climate backgrounds; 3) revealing the varying impacts of influential factors across different spatial and temporal domains for Ta estimation using SHAP-based assessment.

2. Data

2.1. Dataset of near-surface air temperature

In this study, near-surface Ta data was collected from the Global Hourly – Integrated Surface Database (ISD) from the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration, the U.S. (formally the National Climatic Data Center). The ISD consists of hourly and synoptic surface observations derived from numerous sources (Smith et al., 2011). In this study, the ISD Ta data from 2017 to 2021 was collected hourly (three-hour for some of the stations) by the weather stations located in various regional climate zones and LCZs in China (Fig. 1).

2.2. Datasets of ambient influential factors

To analyze the estimation of Ta based on LST and other ambient factors, multi-source data were collected and processed to represent potential factors of Ta for subsequent analyses. Considering the spatial resolution of ambient influential factors data, buffers with 500-m radii are expected to extract information of the explanatory variables with similar spatial coverage. Therefore, although some of the influential factors are at coarse resolution, to ensure a fair data collection, all the imagery-derived explanatory variables were extracted using 500-m buffers around the stations, based on their best temporal resolutions in data acquisition (zonal-calculated into mean values unless otherwise specified) during 2017 – 2021 (Fig. 1). Adjusted from the framework proposed by Lowry (1977), Ta can be decomposed as:

$$Ta = T_w + T_b + T_l + T_u + T_h \quad (1)$$

where, as the original three key components suggested by Lowry, T_b , T_l , and T_u stand for the components of the temperature decided by the background macroclimate, the temperature perturbation controlled by the factors of local natural landscape and urban structure, respectively. T_w and T_h represent additional components influenced by short-term weather information and human and anthropogenic activities, respectively. The influential factors of Ta classified into these five categories introduced in the following subsections and summarized in Appendix A were inputted as explanatory variables for Ta estimation.

2.2.1. LST and other variables for weather information (T_w)

As one of the most important explanatory variables for weather information, LST is closely related to the heat-transferring and energy-exchanging processes between the land surface and the atmosphere. Daytime and nighttime LST data were collected from MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global (MOD11A1.061) on the Google Earth Engine (GEE) platform (Wan et al., 2021). To ensure data quality and avoid cloud effects and emissivity / LST errors, all LST data within the stations' buffer regions were filtered based on the product's quality control records at daytime (QC_Day) and nighttime (QC_Night). Imagery masking by bitmask (Gandhi, 2021) was conducted to make sure the pixels were at good quality both in mandatory quality assessment flags (Bits 0–1), data quality flag (Bits 2–3), average emissivity error no more than 0.02 (via Bits 4–5: Emissivity error flag), and average LST error no more than 2K (via Bits 6–7: LST error flag). All the available LST during the study periods were sampled by the stations in this study. In addition, as the representatives of local meteorological situation, daily-aggregated local albedo ("era5_albedo"), net solar radiation ("era5_net_sr"), downwards shortwave radiation ("era5_sr_down"), precipitation ("era5_precipitation"), and atmospheric pressure ("era5_pressure") were collected from the ECMWF ERA5 climate

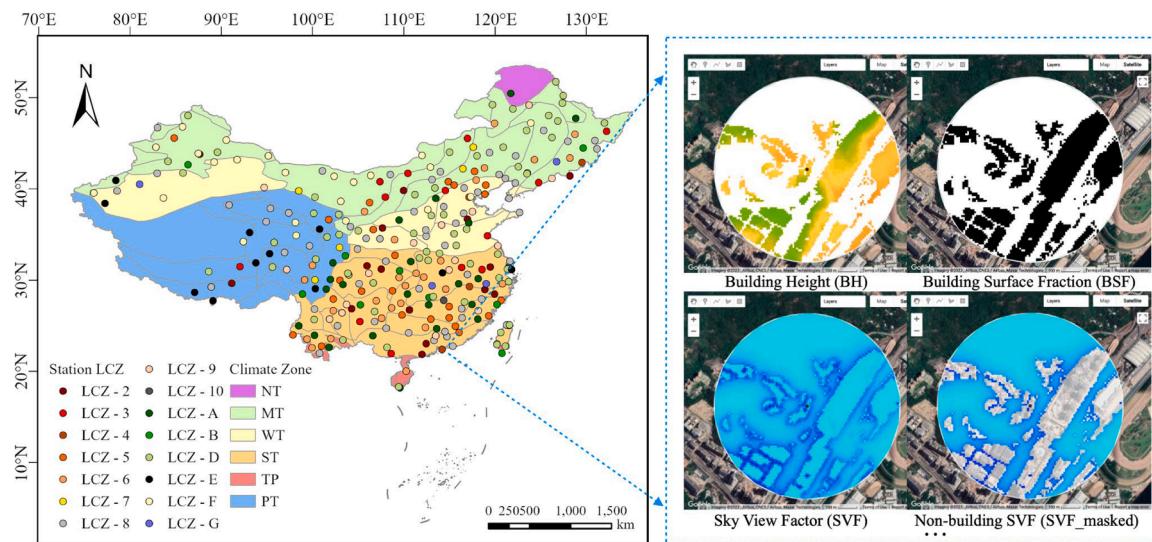


Fig. 1. Left: Weather stations used in this study under different regional climate zones and LCZ types. Right: the schematic diagram showing the data extraction for explanatory variables by stations' 500-m buffers.

reanalysis dataset (Copernicus Climate Change Service, 2019) and inputted as explanatory variables.

2.2.2. Variables for background macroclimate (T_B)

The estimation of Ta is influenced by macroclimate conditions in large-scale spatial and temporal domains. The latitude of the weather station ("Latitude") is crucial in determining the intensity of overall solar radiation received during the Earth's rotation and revolution, leading to significant spatial differences in thermal distribution under macroclimates. Additionally, the day of the year ("DOY") was also recorded as a dynamic variable representing the temporal background information for Ta estimation.

2.2.3. Variables for local natural landscape (T_L)

Regarding the indicators for the local natural landscapes, several variables representing surface and LULC compositions were collected from remotely sensed and geo-calculated datasets. The normalized difference vegetation index ("NDVI", Eq. 2), normalized difference built-up index ("NDBI", Eq. 3), and normalized difference water index ("NDWI", Eq. 4), widely used spectral indices representing LULC, coverage about vegetation, built-up regions, and water bodies, were collected from the MOD09GA.061 Terra Surface Reflectance Daily Global 500m dataset (Vermote & Wolfe, 2021) in GEE with the bands of near-infrared (NIR, band 2), red (R, band 1), short-wave infrared (SWIR, band 6), and green (Green, band 4). Before data extraction, NDVI, NDBI, and NDWI were filtered based on the product's quality control records (QC_500m) via a bitmask-based imagery masking process to reduce abnormal pixels with low data quality or cloud-covered situation, then the maximum composite images were calculated for each specific date with Ta records using the images collected within the periods ± 7 days, and were further used for data extraction into mean values using the buffers.

$$NDVI = \frac{NIR - R}{NIR + R} \quad (2)$$

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \quad (3)$$

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (4)$$

The percentage of impervious surface (%) ("Impervious") of the buffer regions was derived from the Global 30m Impervious-Surface Dynamic Dataset (GISD30) created by Zhang et al. (2022b). The fraction of the

road ("roadFraction") was calculated as the ratio value of the distance of the total road distance (m) and area of the buffer region (m^2) using the road networks from the global patterns of current and future road infrastructure (GRIP4) dataset, which includes multi-resource road datasets such as OpenStreetMap (Meijer et al., 2018). The yearly surface fractions of greenery ("greeneryFraction") of the buffer regions were extracted from Dynamic World, a 10m near-real-time Land Use/Land Cover dataset (Brown et al., 2022) in GEE. Moreover, the Shannon diversity index ("SDI") was calculated using the same LULC dataset based on Eq. 5:

$$SDI = - \sum_{i=0}^n P_i \log_2(P_i) \quad (5)$$

where P_i represents the proportion of each LULC type in the specific buffer region. The "Aspect", "Elevation", "Slope", and "Hillshade", as crucial terrain indicators showing the information of local landscape, were calculated from the digital elevation model data from the Shuttle Radar Topography Mission (SRTM) V3 product (30 m) from NASA's Jet Propulsion Laboratory (Farr et al., 2007).

2.2.4. Variables for urban structure indicators (T_U)

To represent the factors related to urban structure and building morphology, the dataset of Chinese building height estimate at 10-m resolution (CNBH-10m) was utilized to calculate building height ("BH", Eq. 6) for the stations' buffer regions (Wu et al., 2023). Then, related urban morphology indicators, including building surface fraction ("BSF", Eq. 7) and sky view factor ("SVF", Eq. 8), were further derived using GEE and SAGA GIS (Häntzschel et al., 2005; Oke, 2002). In addition to the mean values directly calculated within the buffer regions of the stations, the mean values of building height ("BH_masked") among buildings-covered regions ($BH > 0$) and sky view factor ("SVF_masked") among non-building areas were further calculated and utilized as individual variables in subsequent analyses.

$$BH = \frac{\sum_{i=1}^n H_i}{n} \quad (6)$$

$$BSF = \frac{\sum_{i=1}^n A_i}{A_{all}} \quad (7)$$

$$SVF = 1 - \frac{\sum_{i=1}^n \sin \gamma_i}{n} \quad (8)$$

where n refers to the number of the buildings in the pixel, H_i represents the height of each building, A_i represents the area of each building footprint, and A_{all} represents the area of the pixel, γ_i represents the angle between the line of sight of each pixel cell upward to the visible sky in different directions in the ambient environment.

2.2.5. Variables for human and anthropogenic indicators (T_H)

Variables related to population, distance to road, and nighttime light were calculated to represent human and anthropogenic indicators. The sum ("pop_sum") and density ("pop_density") of the population in the buffer regions were calculated yearly from the WorldPop Global Project Population Data by 100 × 100m grids (Sorichetta et al., 2015). Distance to the road ("distanceRoad") was calculated as the mean distance to the nearest road (m) of all the pixels in the buffered region using the road network data from the GRIP4 dataset (Meijer et al., 2018). Besides, to investigate the potential impact of urbanization and human activity on air temperature as indicated by nighttime light (NTL) conditions, we calculated the mean ("ntl_mean") and sum ("ntl_sum") intensity values in the buffered regions, as well as the count of pixels with lights ("ntl_count") using the nighttime light remote sensing data from the extended NPP-VIIRS-like NTL dataset by Chen et al. (2021) in GEE, respectively.

2.3. Datasets of regional and local climate zones

To analyze the Ta – LST relationships and evaluate the accuracy of Ta estimation, the regional climate zone types and major LCZ types were collected for each weather station (Fig. 1). The primary climate zone types for all stations were derived from the climate zoning map of China (downloaded from the IGSNRR resource and environmental science data platform of the Chinese Academy of Science (CAS, n.d.)), including types of the sub-tropical zone (ST), mid-temperate zone (MT), plateau climate zone (PT), warm-temperate zone (WT), tropical zone (TP), north temperate zone (NT). Besides, the major LCZ type ("major_LCZ"), which accounts for the largest proportion of area within each buffer, was calculated from the global 100 m-resolution LCZ map created by Dermuzere et al. (2022). These climate zone types and major_LCZs were subsequently utilized as key indicators to evaluate the accuracy of Ta estimation across different spatial regions.

2.4. Data integration and preprocessing

To align the response variable Ta data with the primary explanatory variable LST, the hourly Ta values were paired with corresponding LST records during daytime and nighttime, respectively. Given that the LST data from MODIS images has a relatively slight difference in hour time stamps of data acquisition and some weather stations recorded Ta at 3-hour intervals instead of hourly, the integration process involved aligning LST records as closely as possible with Ta records, specifically within a 2-hour margin. Pairs with a discrepancy greater than 2 hours were excluded. Subsequently, data for the remaining 29 explanatory variables were associated with these aligned pairs according to their respective temporal resolutions and recording time. To promote uniform data distribution across different stations and periods, enhancing the robustness of model training and analysis, only stations that consistently provided data for all explanatory variables were included in this study. Therefore, from 2017 to 2021, 53,094 records in daytime periods and 54,594 records in nighttime periods from 292 stations were utilized (Fig. 1). The stations are extensively distributed in most of the regions in China, which underpins a comprehensive analysis of Ta estimation and its ambient influential factors in different areas.

3. Methodologies

3.1. Analytical framework

As depicted in Fig. 2, after comprehensively analyzing the relationship between Ta and various explanatory factors, different ML methods were employed to develop predictive models for Ta estimation using daytime and nighttime datasets, respectively. Based on the predictive accuracy in testing processes, the optimal ML models were selected for daytime and nighttime datasets, respectively. The performances of optimal models were evaluated across the entire dataset and within specific subsets. Subsequently, the feature importance evaluation and model interpretability analyses would be further conducted and investigated using SHAP methods (Lundberg & Lee, 2017) and across different climatic zoning and LCZ backgrounds using comprehensive optimal models and individual models trained by records under diverse regional and local climate types.

3.2. Linear relationship analysis among Ta and LST

To ascertain the extent to which variations in Ta can be estimated from LST, the coefficient of determination (R^2) was computed between Ta and LST for daytime and nighttime datasets, respectively. The Ta – LST linear relationship was further examined across different spatial (by LCZs) and temporal (by years and DOY) domains. To reduce bias due to sample volume in correlation evaluations, both adjusted R^2 and the Mean Absolute Error (MAE) were calculated using ordinary least squares regressions (80 % of the records: fitting, the remaining 20 %: testing) for comparison among different subsets of Ta – LST pairs.

3.3. Estimation of Ta using machine learning models

In this study, three tree-based ML models, including RF, XGBoost, and LightGBM, were utilized to train the models to estimate Ta using LST and 29 other ancillary variables from datasets of daytime and nighttime, respectively. To address potential overfitting, data from 2017 to 2020 were used for training, while independent data from 2021 served as the testing set.

RF is an ensemble learning method that can be used for classification, regression, and various tasks by constructing numerous decision trees during training, with each tree built by a random subset of the training set and features to enhance diversity and avoid overfitting (Breiman, 2001). By combining the predictions of these trees, the RF model improves accuracy and generalizability, and it is robust to noisy data and capable of dealing with large, high-dimensional datasets. The RandomForestRegressor function from the *sklearn* package in Python was used to train the RF models (RandomForestRegressor, n.d.).

XGBoost, a tree-based ensemble ML model, is known for its exceptional predictive performance and efficiency, facilitated by optimized distributed gradient boosting algorithms. It operates by amalgamating multiple weak models to enhance overall prediction robustness and is particularly adept at handling diverse data types, including both numerical and categorical variables, also with resistance to nonlinearity, inherent feature selection, and discernibility. The XGBRegressor function from the *xgboost* package in Python was used to train the models (Python Package Introduction — Xgboost 1.7.5 Documentation, n.d.).

LightGBM is a gradient boosting-based model focusing on high performance and efficiency in model training (Ke et al., 2017). A novel technique named Gradient-based One-Side Sampling is integrated into LightGBM to filter out the data instances that contribute less to the gradient updates during training, improving efficiency and accuracy simultaneously. The LGBMRRegressor function from the *lightgbm* package in Python was used to train the LightGBM models in this study (Shi et al., 2024).

For model hyperparameter tuning, the RandomizedSearchCV function from *sklearn* (Sklearn.Model_selection.RandomizedSearchCV, n.d.)

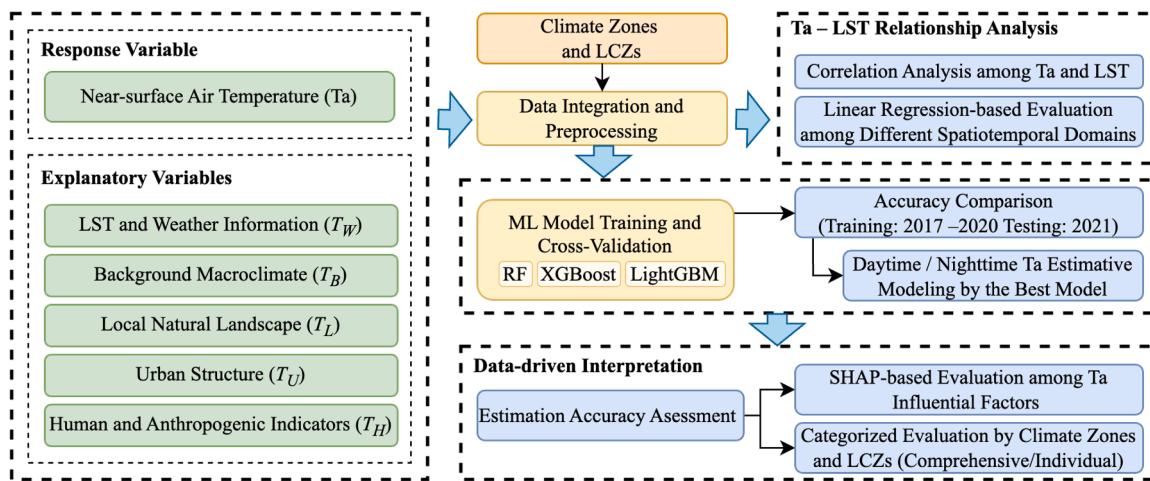


Fig. 2. Analytical framework in this study.

was applied with 5-fold cross-validation to select optimized hyperparameters that ensure the model is well-trained. RMSE was used to compare the performance of the testing processes of different models (Chai & Draxler, 2014). The models with the smallest RMSEs within the daytime and nighttime datasets were selected as the optimal models for the following analyses. Mean absolute error (MAE) was employed to assess the accuracy of Ta estimation under different spatiotemporal contexts.

3.4. SHAP-based interpretation analyses

To elucidate the contribution of different variables in estimating Ta, the relative feature importance scores were first calculated, which facilitated the assessment of each variable's significance in model tuning. Then, SHAP, a powerful approach for interpreting ML models and examining feature interactions, was applied to all explanatory variables within the trained XGBoost model (Lundberg et al., 2020). Deeming all the features as “contributors”, the SHAP approach employs an additive explanatory model which assigns a Shapley value for each predicted feature in a prediction, indicating how much each feature contributes to the final prediction (Lundberg & Lee, 2017; Wang et al., 2023b). The explanation can be specified as:

$$g(\mathbf{z}') = \emptyset_0 + \sum_{j=1}^M \emptyset_j z'_j \quad (9)$$

where g is the explanation model, $\mathbf{z}' \in \{0, 1\}^M$ is the coalition vector, M

is the maximum coalition size, and \emptyset_j represents the feature attribution for a feature j , i.e., the Shapley value. In recent years, SHAP has emerged as an essential tool for interpreting ML models, especially tree-based ones (Angelov et al., 2021). These analyses can be used to interpret critical variables and their interactive effects across varying datasets and specific spatiotemporal conditions.

4. Results

4.1. Descriptive analysis of the datasets

In this study, data were collected from 292 weather stations, mainly distributed across four main climate zone types, including ST, MT, PT, and WT (Fig. 3), and there are 10 and 1 station(s) located in other climate zones of TP and NT, respectively. Among the major LCZ types within the buffer regions, LCZ-8 (built-up types of “large low-rise”), LCZ-D (land-covered types of “low plants”), and LCZ-6 (built-up types of “open low-rise”) were the most prevalent. During the study periods, the dataset exhibited a consistent distribution in the numbers of records of each station in both daytime and nighttime datasets among different years (Fig. 4). This consistency ensures a representative division of the records into training (2017–2020) and testing sets (2021), thereby enhancing the reliability of the modeling and evaluation processes.

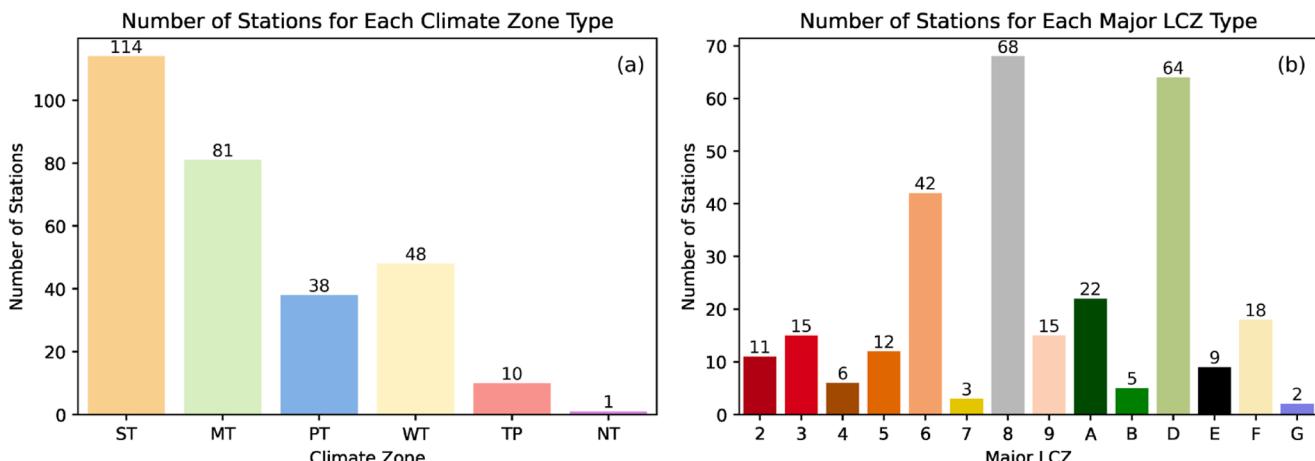


Fig. 3. Distribution of the weather stations by different climate zone types (a) and major LCZ types (b).

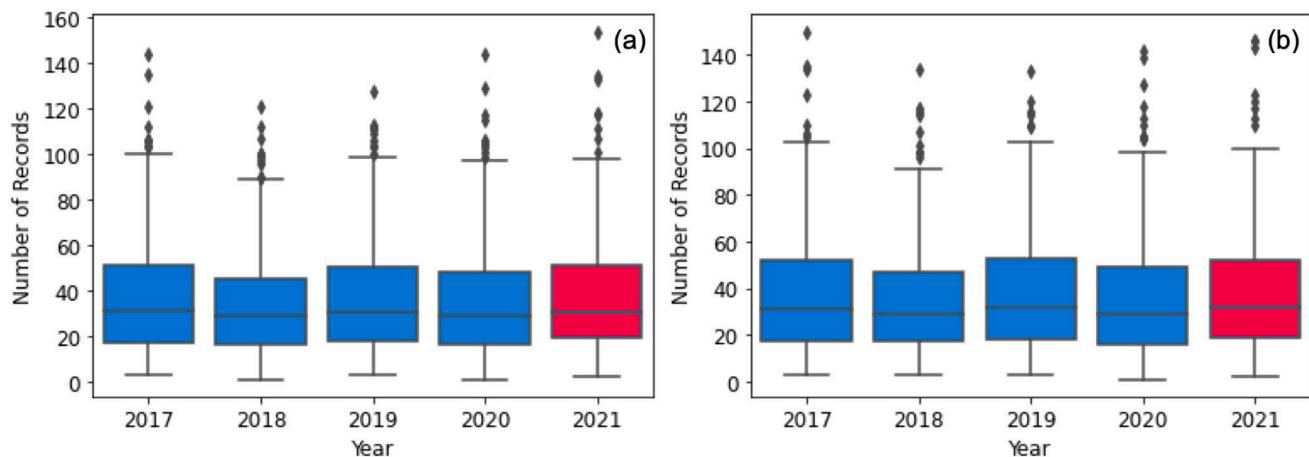


Fig. 4. Distribution of the usable record numbers of each station by years in the daytime (a) and nighttime (b) datasets, respectively (blue: training; red: testing).

4.2. The $T_a - LST$ relationship

In the dataset used in this study, T_a and LST exhibited a significant positive correlation in a linear relationship with an R^2 of 0.87 during the daytime and with a better linear relationship ($R^2 = 0.92$) at nighttime (Fig. 5). These indicate the potential utility of LST-based regression models for estimating T_a .

Further, the correlation relationship between T_a and LST varied by major LCZ types and years (Fig. 6). The records exhibited relatively weak $T_a - LST$ correlation under LCZ-7 (built-up types of "Lightweight low-rise") and LCZ-E (land-covered types of "Bare rock or paved") across both daytime and nighttime. Notably, despite having similar data volumes, the nighttime dataset generally presented better correlations between T_a and LST among major LCZs and years than the daytime dataset. However, even though the R^2 exhibits a linear fitting potential between the two variables, its validity might be influenced by the number of samples for each analysis. Therefore, the T_a predictive performance was also assessed using the MAEs calculated between T_a of each station and predicted T_a by LST-based ordinary least squares (OLS) regressions fitted under different LCZ types (Fig. 7). Records in the nighttime dataset also indicated superior predictive capability than daytime datasets in linear regression-based T_a estimation among most LCZ types. Besides, records in urban regions, particularly in LCZ-2, 4,

and 5, exhibited more consistent correlations between T_a and LST than most non-built-up types.

On the other hand, the relationships between T_a and LST also vary under different periods, which exhibited seasonal variations. As depicted by the adjusted R^2 and RMSE on OLS linear regressions in Fig. 8, both daytime and nighttime datasets demonstrated more robust $T_a - LST$ linear relationships in autumn and winter compared to spring and summer, and the linear relationships between T_a and LST at nighttime are more robust than that at daytime during most of the period.

Overall, the analytical results indicated that although LST can reflect T_a variations under linear regression, its linear-based predictive performance is subject to variation across different spatiotemporal contexts. To enhance the accuracy of T_a estimation, it is critical to integrate additional ancillary natural and social variables into advanced multivariate regression methods. This integration is promising to improve the robustness and applicability of the models across diverse environmental and temporal settings.

4.3. Model training and SHAP-based interpretation

Prior to model training, an analysis of the correlation matrixes (Fig. 9) was conducted to assess the relationships among the variables in this study. Despite some variables of urban morphology situation

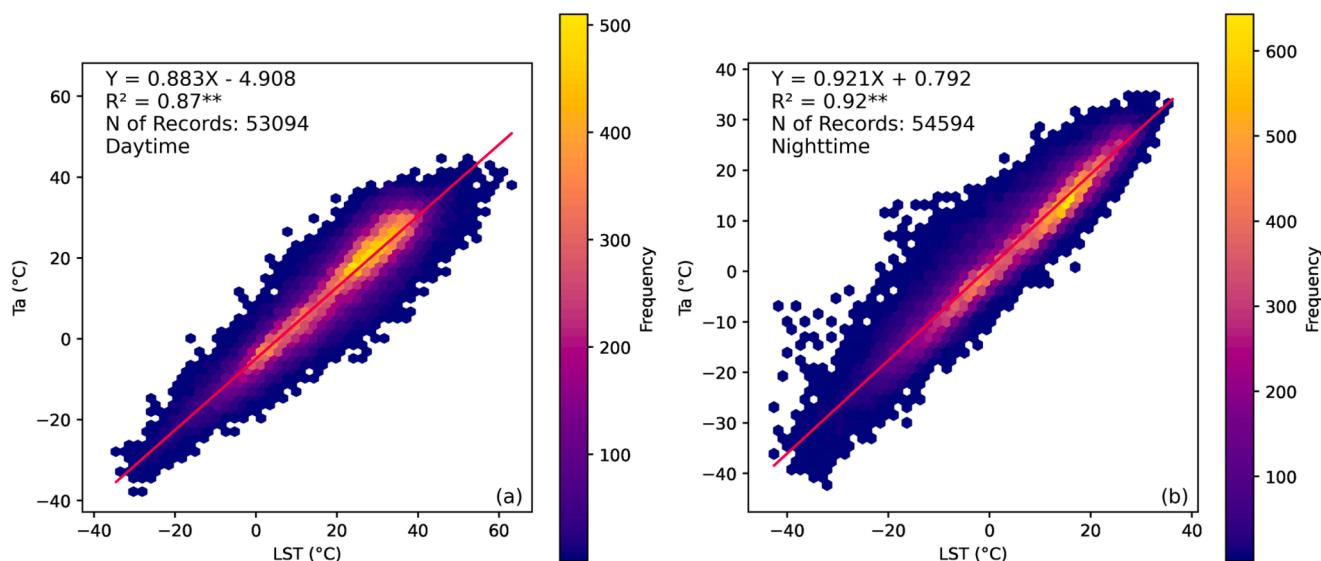


Fig. 5. Correlation analysis among T_a and LST based on records in the daytime (a) and nighttime (b) datasets, respectively (** represents significance levels with p-values of 0.01).

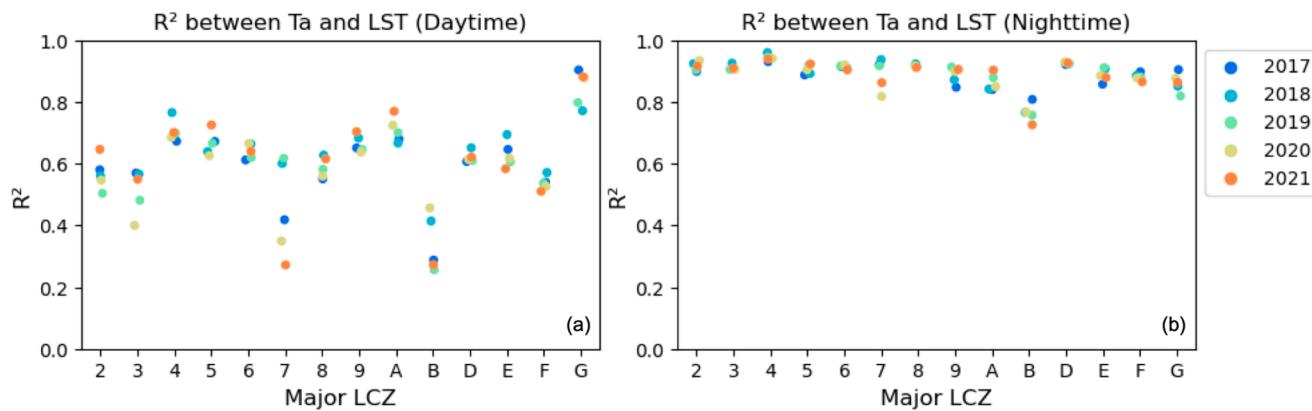


Fig. 6. Coefficient of Determination (R^2) between Ta and LST based on records in the daytime (a) and nighttime (b) datasets under different major LCZs and in different years.

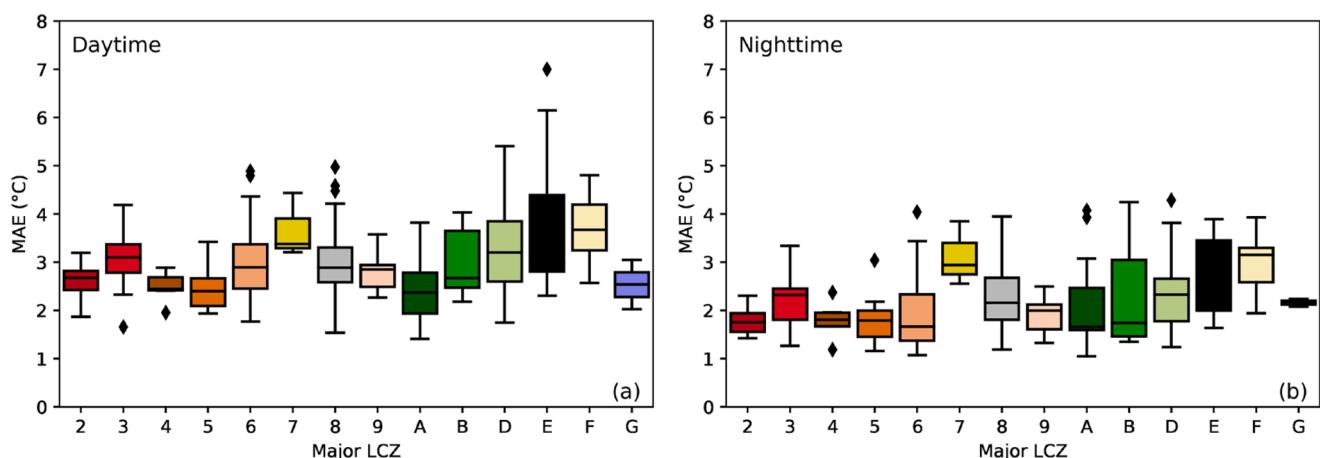


Fig. 7. MAE for Ta estimation based on LST-based OLS linear regressions with records in the daytime (a) and nighttime (b) datasets under different major LCZs.

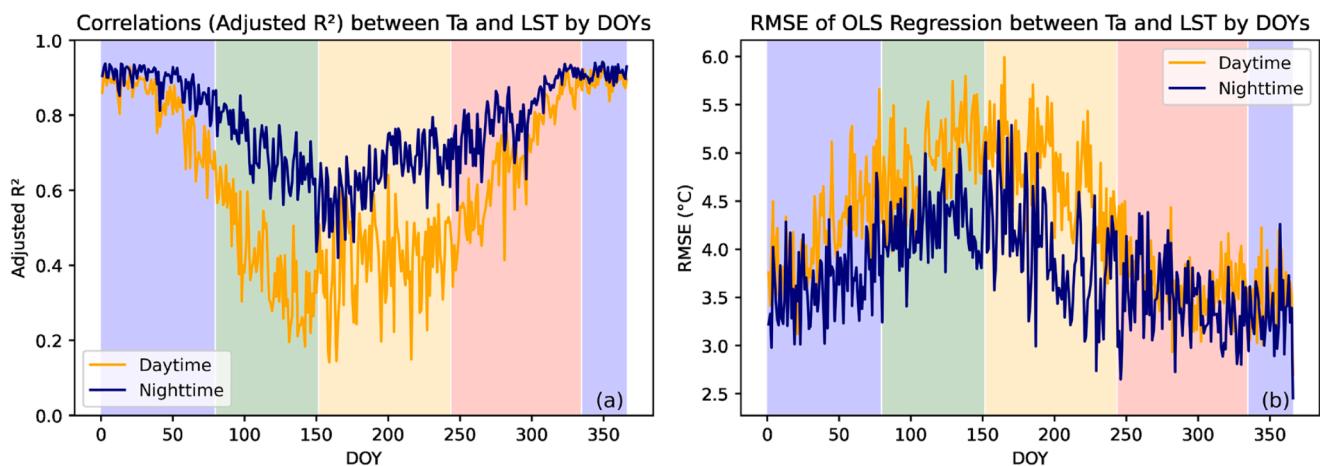


Fig. 8. Adjusted R^2 and RMSE for Ta based on LST-based linear regressions by DOYs.

showing significant positive or negative relationships with each other, most explanatory variables exhibited relatively low correlations across different categories of influential factors. Even so, as different variables might offer unique contributions to the predictive capabilities, and the models selected in this study can cope with overfitting issues caused by intercorrelations to some extent, all the variables were used in the following model training and SHAP-based analyses.

Data for Ta and all explanatory variables were divided into training and testing sets based on the year of data collection (training: 2017 – 2020, testing: 2021). The tuning processes of RF, XGBoost, and LightGBM were conducted using cross-validations to determine the optimal parameter settings (Appendix B). The LightGBM-based models achieved the optimal predictive performance both in the testing datasets of daytime and nighttime, with RMSE values of about 3.03 °C and 2.64

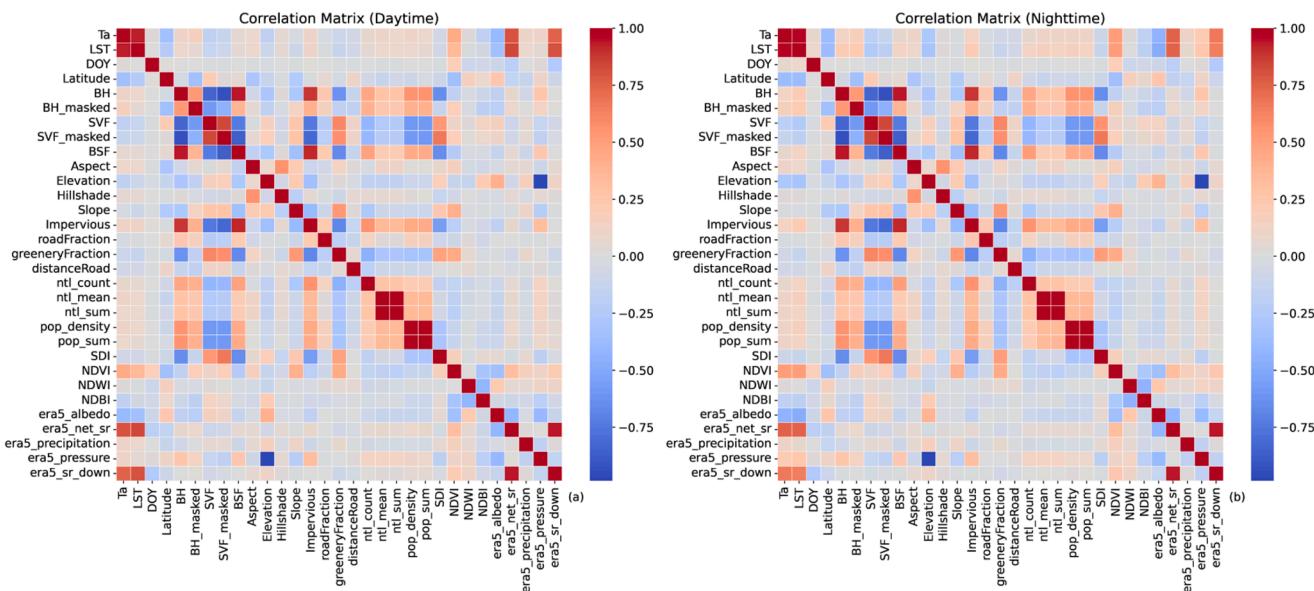


Fig. 9. Correlation matrix (r) among the variables used in this study at daytime (a) and nighttime (b).

°C, respectively (Fig. 10).

After the LightGBM model training, the relative feature importances of all explanatory variables were calculated using a gradient-boosting approach. As shown in Fig. 11, LST was identified as the most influential variable across both datasets, significantly impacting Ta estimation. In addition to LST, terrain and location-related variables such as elevation, slope, and latitude and ERA5-related climatological variables were highly ranked and accounted for more importance in modeling.

However, the relative importance in the modeling process provides only global rankings of feature importance based on aggregated

importance metrics rather than individualized ones, offering a general view that might not be feasible for detailed analysis of subset importance (Hastie et al., 2009). To address these limitations, the SHAP value-based feature importance analysis among all the records was employed for further investigation (Fig. 12). Apart from LST and DOY, which were already emphasized by relative importance (Fig. 11), Latitude and elevation also demonstrated vital significance by SHAP values and were ranked as key factors. Besides, the SHAP value rankings among explanatory variables highlighted the significant role of the background meteorological information from ERA5 in Ta estimation, even though

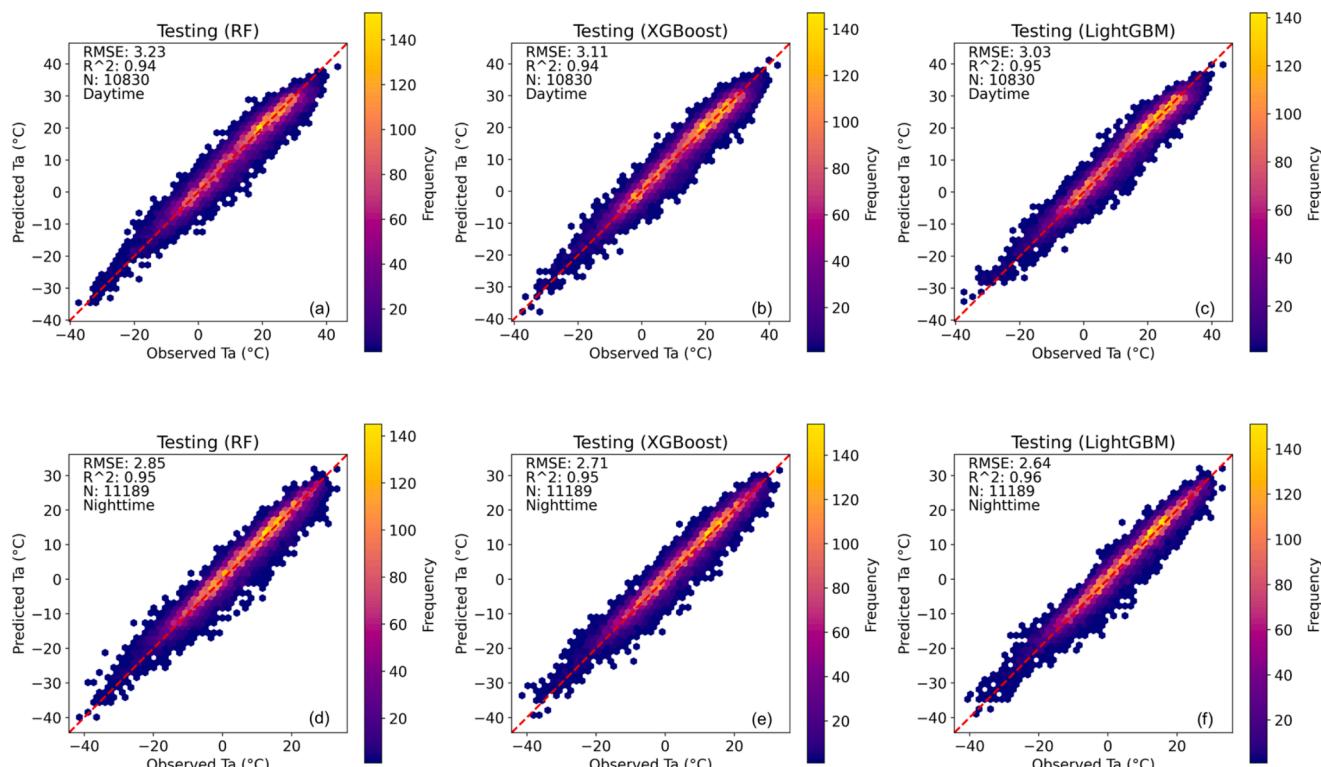


Fig. 10. Testing performance of Ta estimation (Hexbin Plots) based on the datasets of daytime (a, b, c) and nighttime (d, e, f) in 2021 using models of RF, XGBoost, and LightGBM, respectively. (frequency: number of records presented in each hexagon).

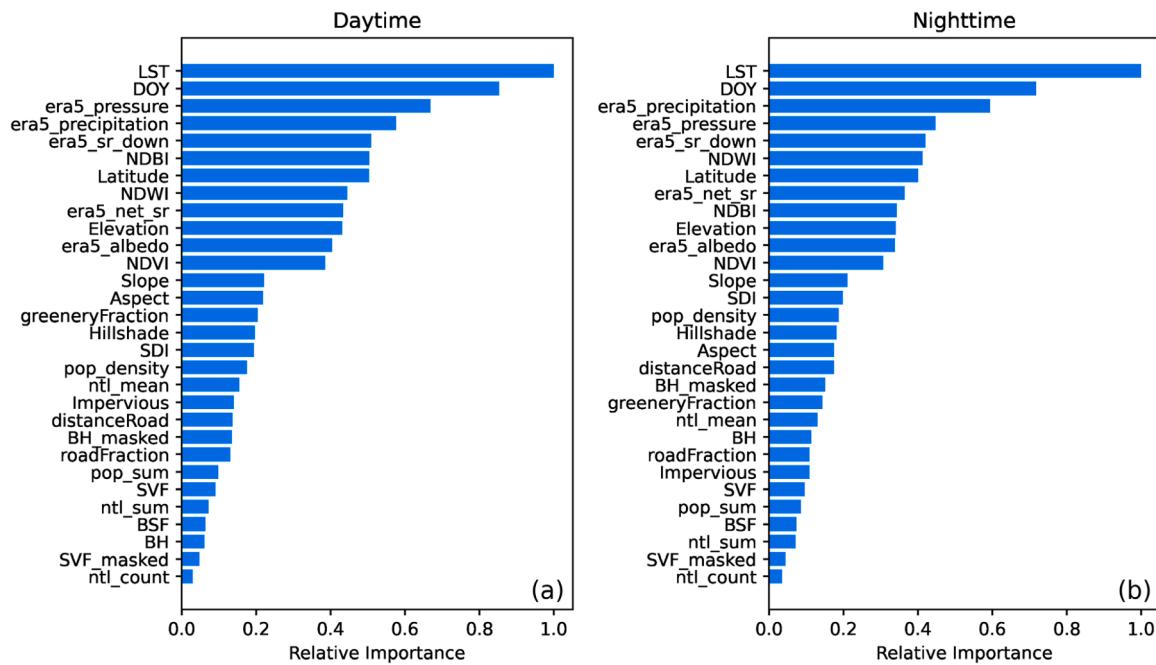


Fig. 11. Relative feature importance of influential factors in the optimal models' training processes based on the datasets of daytime (a) and nighttime (b).

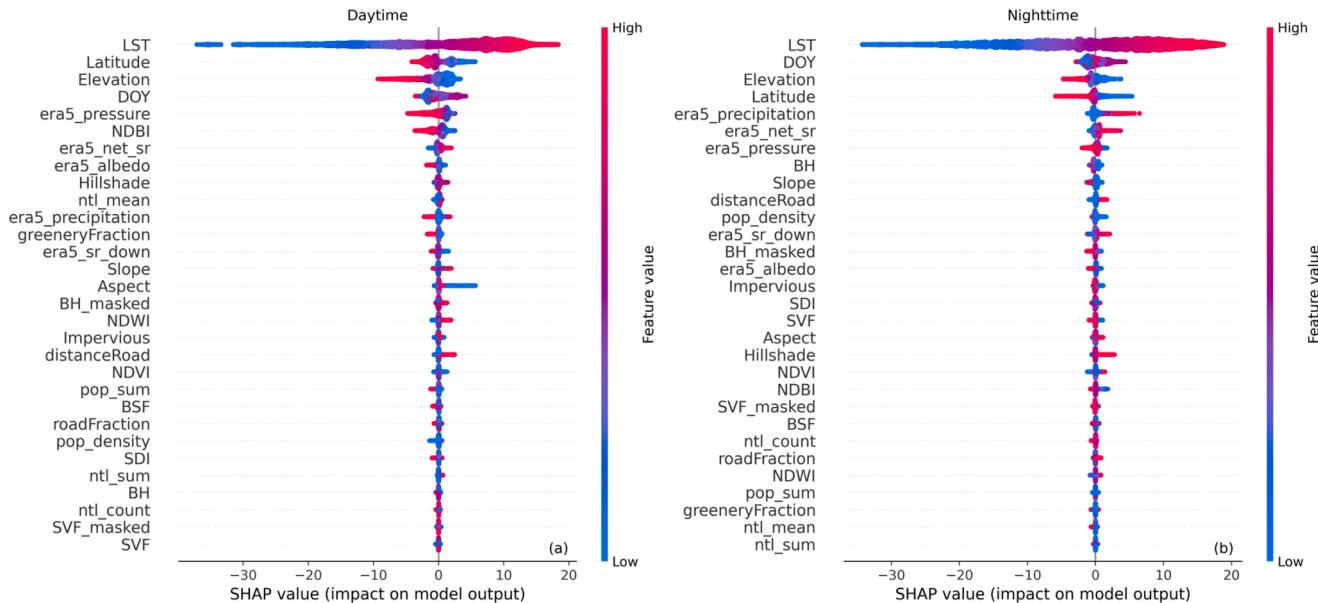


Fig. 12. Distribution and rankings of SHAP values (top: most vital impact) for influential factors in the optimal models for full daytime (a) and nighttime (b) datasets.

most of those variables were collected with coarse resolutions. In daytime datasets, NDBI, the variable representing the coverage and urbanization situation, was more critical with a relatively high ranking among variables, while BH accounts for much importance at nighttime. Furthermore, the distribution of SHAP values across variables revealed diverse impacts on Ta estimation performance corresponding to the daytime and nighttime datasets. It should be noted that some human activity-induced variables, such as distanceRoad and pop_density, tended to have more complex influence patterns at nighttime.

To examine the consistency of the SHAP-based feature importance ranking under different temporal contexts, the LightGBM models were further trained using data from different years (Fig. 13), in which data from each 4 years' selection were used to train the LightGBM model, the data from the remaining one year was used for testing and SHAP

analysis. The separated five models received relatively similar accuracies in testing (RMSE in models for the year for testing: 2017: 3.00 °C, 2018: 3.10 °C, 2019: 3.10 °C, 2020: 3.00 °C, and 2021: 3.03 °C in daytime dataset; 2017: 2.70 °C, 2018: 2.76 °C, 2019: 2.66 °C, 2020: 2.64 °C, and 2021: 2.64 °C in nighttime dataset). The variations of SHAP-based feature importance rankings were presented in Fig. 13 and summarized in Appendix C, where most of the influential factors, especially top features, remained stable rankings both in daytime and nighttime datasets. All the variables exhibited low coefficient of variation (CV) of the rankings (maximum: 17 % in the daytime dataset and 13 % in the nighttime dataset).

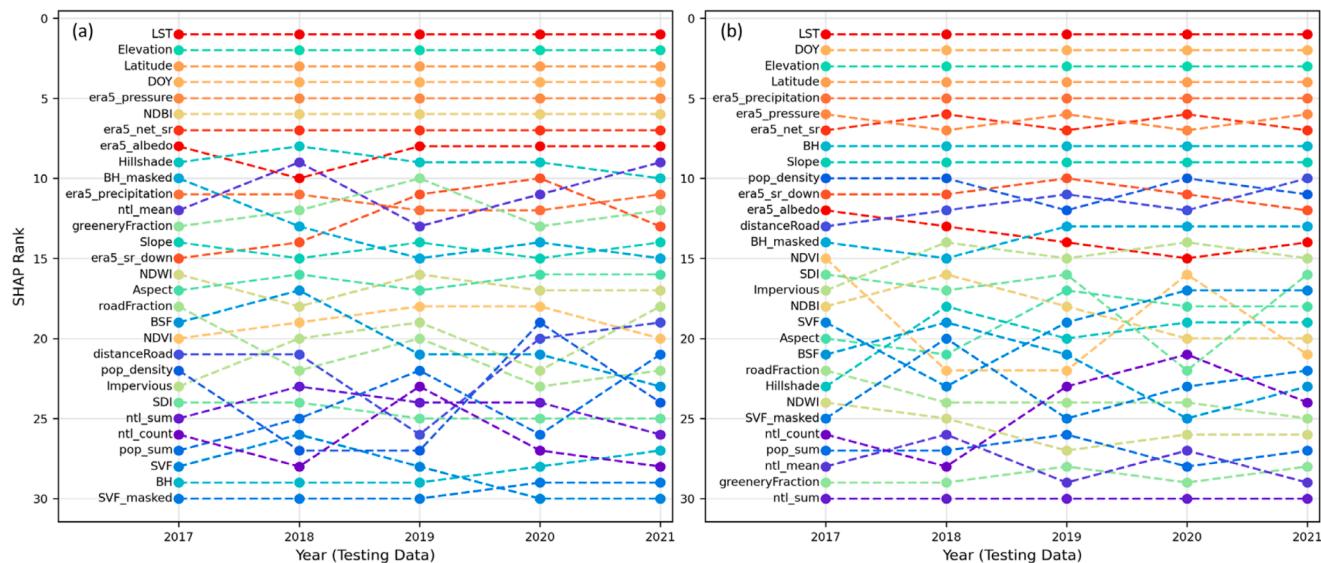


Fig. 13. Variation of the SHAP-based feature importance rankings based on LightGBM models for testing data in different years (records in each year for testing and remaining records for model training) for daytime (a) and nighttime (b) datasets.

4.4. Spatiotemporal evaluation by climate zones and LCZs

The predictive errors of the trained LightGBM-based models for Ta estimation were analyzed using MAE across different weather stations. As depicted in Fig. 14, most stations in the nighttime dataset generally had superior Ta estimative performance compared to the daytime dataset. Stations located in southeastern near-coastal regions exhibited relatively better accuracy for daytime datasets. Besides, weather stations in Inner Mongolia and the north-eastern regions, roughly in the mid-temperate zones, showed relatively lower accuracy at nighttime. This suggests the need for further investigation among stations in different regional climate zones and LCZs.

Therefore, all the stations in this study were categorized based on their climate zone types and were employed for further analysis of MAE distribution by climate zones (Fig. 15). Unlike the other climate zones, the stations in the mid-temperate zone did not exhibit improved nighttime accuracy and recorded the highest median MAE among all the climate zones during nighttime. On the other hand, as shown by the

MAE values across different major LCZ types in Fig. 16, most LCZ types generally achieved better Ta estimation accuracy at night, except LCZ-7 (“lightweight lowrise”), which might be due to low data record volume from only three stations. Besides, stations in low-rise and sparsely built LCZ types (LCZ-3, 6, 7, 8, 9) were characterized by lower accuracy than more urbanized built-up areas.

To identify the most influential explanatory variables under different regional climate types and LCZs within diverse urban/rural landscapes, SHAP values were calculated from the comprehensive LightGBM trained in the aforementioned steps (“global model”) for the features of stations in four main climate zones and various grouped LCZ categories (Appendix D), which were further ranked and summarized in Fig. 17. Considering LST is always the most vital feature among all the subsets, only the other ambient influential factors were analyzed and compared. Among all the influential factors, most variables of T_B , T_W and elevation presented vital contributions with high SHAP rankings among various regional and local climate zones during both daytime and nighttime periods. NDBI, Hillshade, and ntl_mean exhibited obvious importance

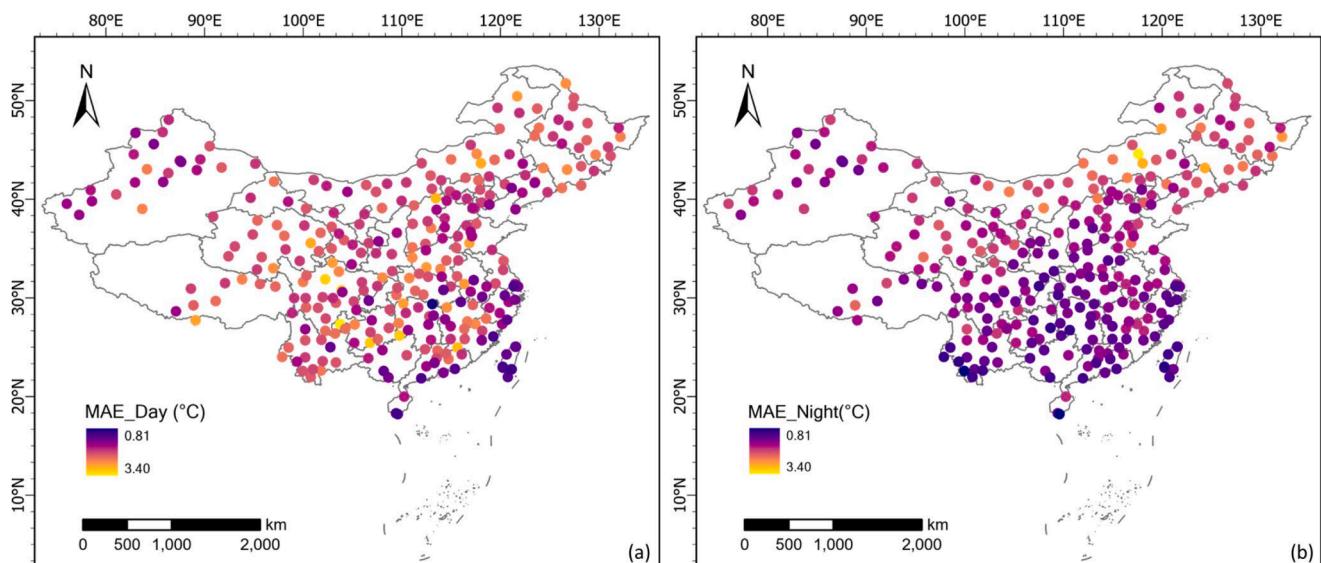


Fig. 14. Mean predictive error (MAE) for Ta estimation among different weather stations based on the datasets of daytime (a) and nighttime (b).

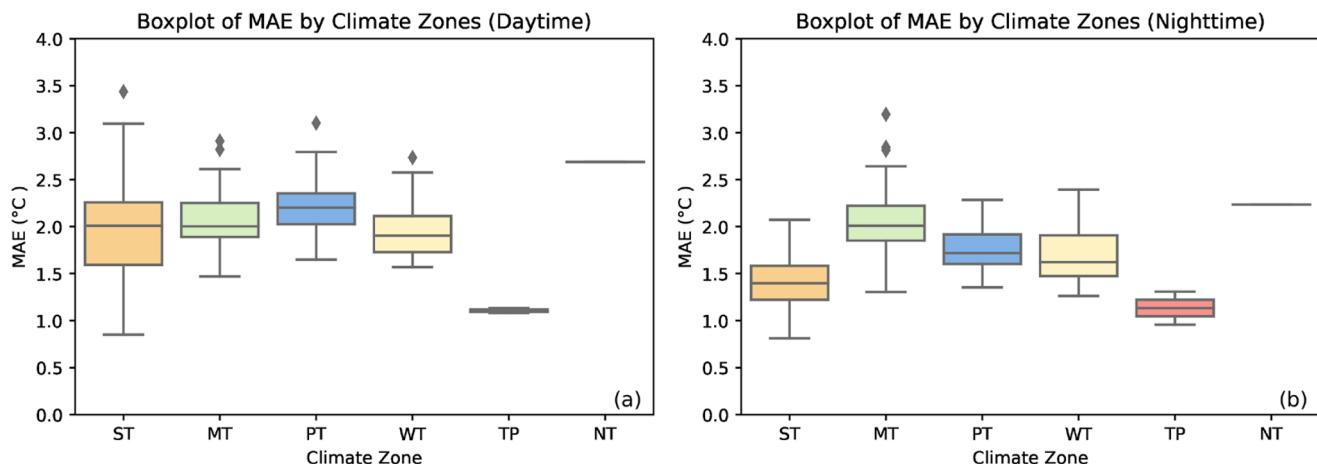


Fig. 15. Predictive accuracy (MAE) for Ta estimation among stations by climate zone types based on the datasets of daytime (a) and nighttime (b).

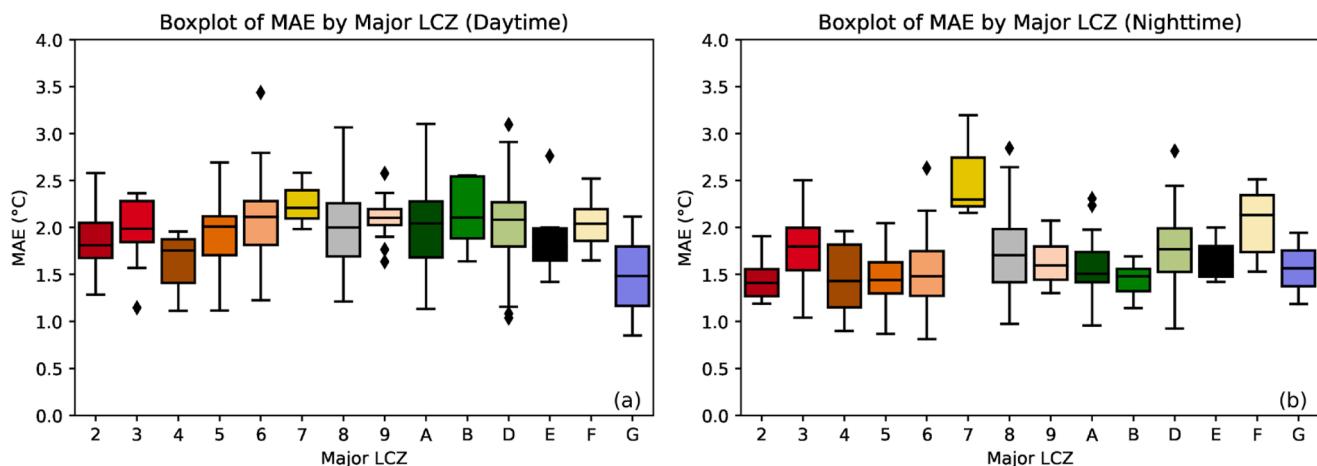


Fig. 16. Predictive accuracy (MAE) for Ta estimation among stations by major LCZ types based on the datasets of daytime (a) and nighttime (b).

for Ta estimation at daytime, while BH and some human activity-related variables such as pop_density and distanceRoad showed strong SHAP rankings, specifically at nighttime. Compared to the distribution under regional climate zones, the SHAP values of factors under different LCZs indicated more evident and distinct ranking patterns. It can be noted that SVF exhibited relatively high importance in high- and mid-rise built-up regions (LCZ – 2, 4, 5). During daytime periods, the greeneryFraction also reflects critical influence with higher SHAP rankings in these higher-rise urban regions, as well as vegetation-covered regions (LCZ – A, B, D). This analysis indicates that influential factors assume diverse and different levels of functional roles in Ta estimation under varying spatiotemporal domains.

To further investigate the robustness of SHAP analysis and address potential regional biases from the global models, records from diverse categories of the regional climate zones and LCZs were utilized to train the separated LightGBM models (“local model”), resulting in diverse predictive accuracies (Fig. 18), respectively. Similar to the results in global models, separated local models trained in each climate zone type and LCZ category presented better accuracies during nighttime than daytime, where ST performed the best among climate zones, and records with LCZ types of high- and mid-rise built-up regions (LCZ – 2, 4, 5) indicated better testing accuracies than the other LCZ types.

The SHAP value-based feature rankings were also derived based on testing data and individual local models for each climate zone and LCZ category (Fig. 19). Elevation and most of the T_B , T_W variables still presented similar high rankings, which are consistent with the global

models’ results, denoting the impacts of these factors are under global patterns among all the stations rather than local dynamics. However, some of the variables exhibited obviously varied SHAP rankings compared to the results of the global models. Out of the complex ranking distributions from the local models, NDBI still presented outstanding importance in daytime datasets for Ta estimation, and BH also showed strong SHAP rankings at nighttime but only in built-up LCZ regions. It is also worth mentioning that SVF still displayed considerable importance in high- and mid-rise built-up regions (LCZ – 2, 4, 5).

5. Discussion

In this study, we conducted and evaluated estimative modeling of insitu daytime/nighttime Ta using trained ML models with inputs of LST and various ambient factors collected by buffers for weather stations across China. This research, encompassing multisource modeling and assessments, provides a valuable data-driven evaluation of how LST and ambient factors influence Ta under diverse regional and local climatical contexts, as well as at daytime and nighttime, respectively.

Given the massive concerns about LST’s representativeness in quantifying UHI effects, the Ta has gradually received more investigation along with LST via estimative analyses (Agathangelidis et al., 2016). Previous studies predominantly analyzed the relationships between Ta and LST by stations within a single city or small regions (Azevedo et al., 2016; Oswald et al., 2012). So, in this study, preliminary analyses of linear relationships between Ta and LST were applied for all the records

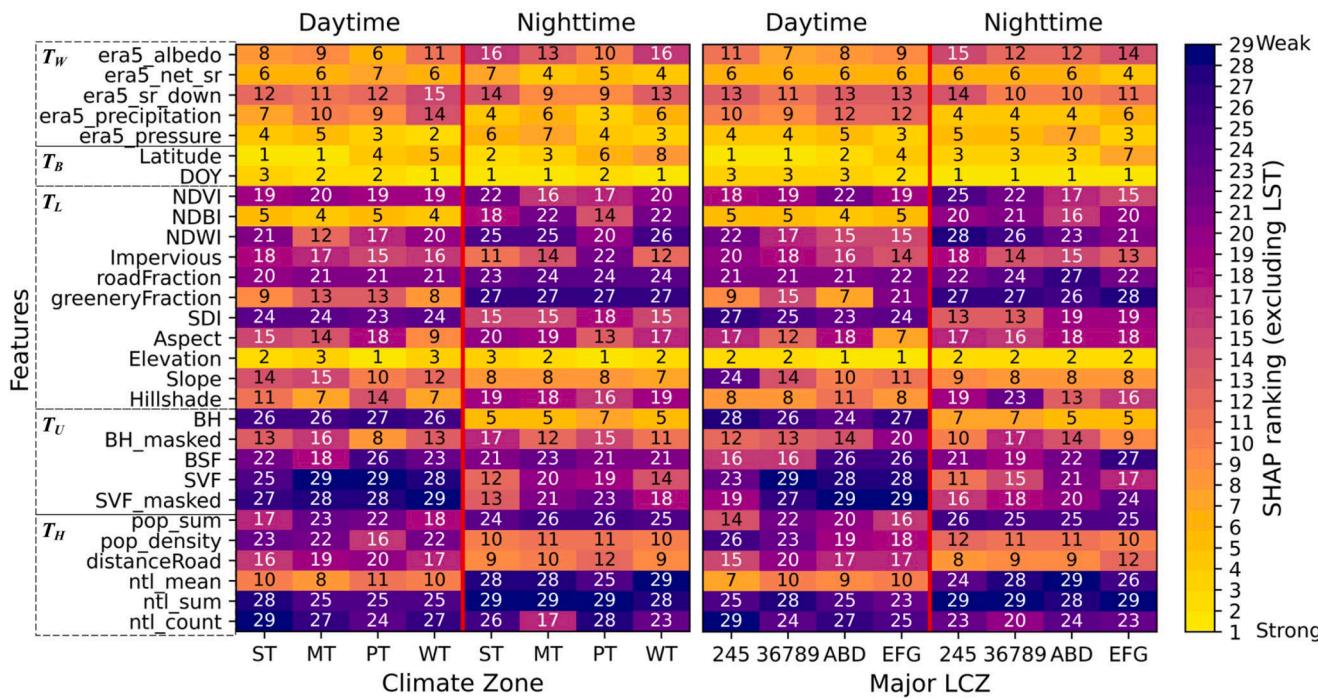


Fig. 17. Rankings of the SHAP values among the influential factors (excluding LST) based on the comprehensive LightGBM model for the datasets within different climate zones (ST, MT, PT, and WT) and major LCZ categories.

during the 5-year period, which confirmed LST's strong capability to capture Ta variations, albeit with varying accuracy across different spatial and temporal domains. Besides, with similar landscapes in urban regions, the predictive Ta – LST linear relationships among built-up LCZ types are overall more stable and better than those of the LCZs in land-cover types. Regarding temporal perspectives, the nighttime Ta – LST correlations proved generally more robust than daytime, while instability was still observed in some land-covered LCZ types. Seasonal variations also influenced the Ta – LST correlations, with better results shown in autumn and winter. These findings align with [Burnett and Chen \(2021\)](#), who reported stronger Ta – LST relationships in winter than summer across all land cover types in Southern Ontario, Canada. [Peng et al. \(2018\)](#) noted that LULC changes, particularly vegetation coverage, undergo apparent seasonal variations, affecting the surface solar energy balance and resulting in heterogeneous thermal distributions on the land surface and near-surface air. These insights underscore the need to examine the diverse impacts of the ambient influential factors with multidimensional spatiotemporal information for in-situ Ta estimation using advanced machine learning methods.

Utilizing data from 2017 to 2020 for cross-validation-supported tuning and 2021 data for independent testing, three ML models were trained to estimate daytime and nighttime Ta with inputs of LST and various auxiliary variables. For 10,830 daytime records and 11,189 nighttime independent records in 2021, the LightGBM-based models achieved the optimal testing accuracies in both daytime and nighttime datasets, with RMSE of 3.03 °C and 2.62 °C and R^2 of 0.95 and 0.96, respectively. Although some related studies report similar or better RMSEs, they often rely only on cross-validation results or random dataset splits without independent testing, potentially leading to overfitting caused by the high similarity among training and testing subsets, limiting the robustness of the models for extended applications ([Wang et al., 2023a; Wang & Wu, 2016; Zhang et al., 2022a](#)). In contrast, accuracy in this study was derived from near-real-time (following LST's time of retrieval) Ta estimation using a year's worth of individual testing data from 292 separately located weather stations across China. These accuracy results are similar to the study of [Hrisko et al. \(2020\)](#), which reported an RMSE of 2.6 °C in Ta estimation using a satellite

LST-supported regression neural network trained with data from 150 weather stations over five months in the continental U.S. It was also reported that [Su et al. \(2023\)](#) achieved similar predictive performance in validating spatially seamless Ta estimation with RMSEs of 2.376 °C, 2.808 °C and 2.823 °C for mean, minimum and maximum Ta using observed in-situ records, respectively. Overall, the applicable accuracies of the optimal models (LightGBM) trained in this study can be employed to evaluate the impacts of the influential factors for in-situ Ta estimation.

On the other side, SHAP value rankings indicated the diverse contributions of different ancillary variables across various geospatial regions. In nighttime models, variables related to urban morphology and human activities, such as the BH, pop_density, distanceRoad, and SVF ranked relatively higher than in daytime models. This potentially suggests that after sunset, the 3D structures of different landscapes would influence heat dissipation to various degrees. At the same time, anthropogenic heat emitted from human activities could become a more significant heat source affecting Ta's development ([Kotharkar et al., 2023](#)). The analyses among variations of SHAP rankings based on yearly separated-trained LightGBM showed consistent overall SHAP rankings, which further supported the robustness of SHAP-based influential factors evaluations.

Moreover, meteorological variables from ERA5, despite being derived from the coarse-resolution data sources, still achieved very high SHAP-based contributory values in Ta estimation. This finding was not only indicated by SHAP results over all the stations, but was also validated via the region-specific SHAP assessments by single comprehensive LightGBM models and individual local models trained in different climate zones and LCZ contexts. This supports the suggestions of [Wang et al. \(2023a\)](#) that simulated spatially varying meteorological variables are critical to be incorporated into ML models for bias corrections as well as to improve Ta prediction ([Cho et al., 2020](#)). On the other side, the SHAP analyses based on individual local models reveal that despite some factors (e.g., elevation and the T_B T_w factors) exhibiting global impacts across the stations, most of the impacts from the ambient influential factors are spatiotemporal context-dependent. Several factors stand out from certain contexts, such as NDBI in the daytime, BH for

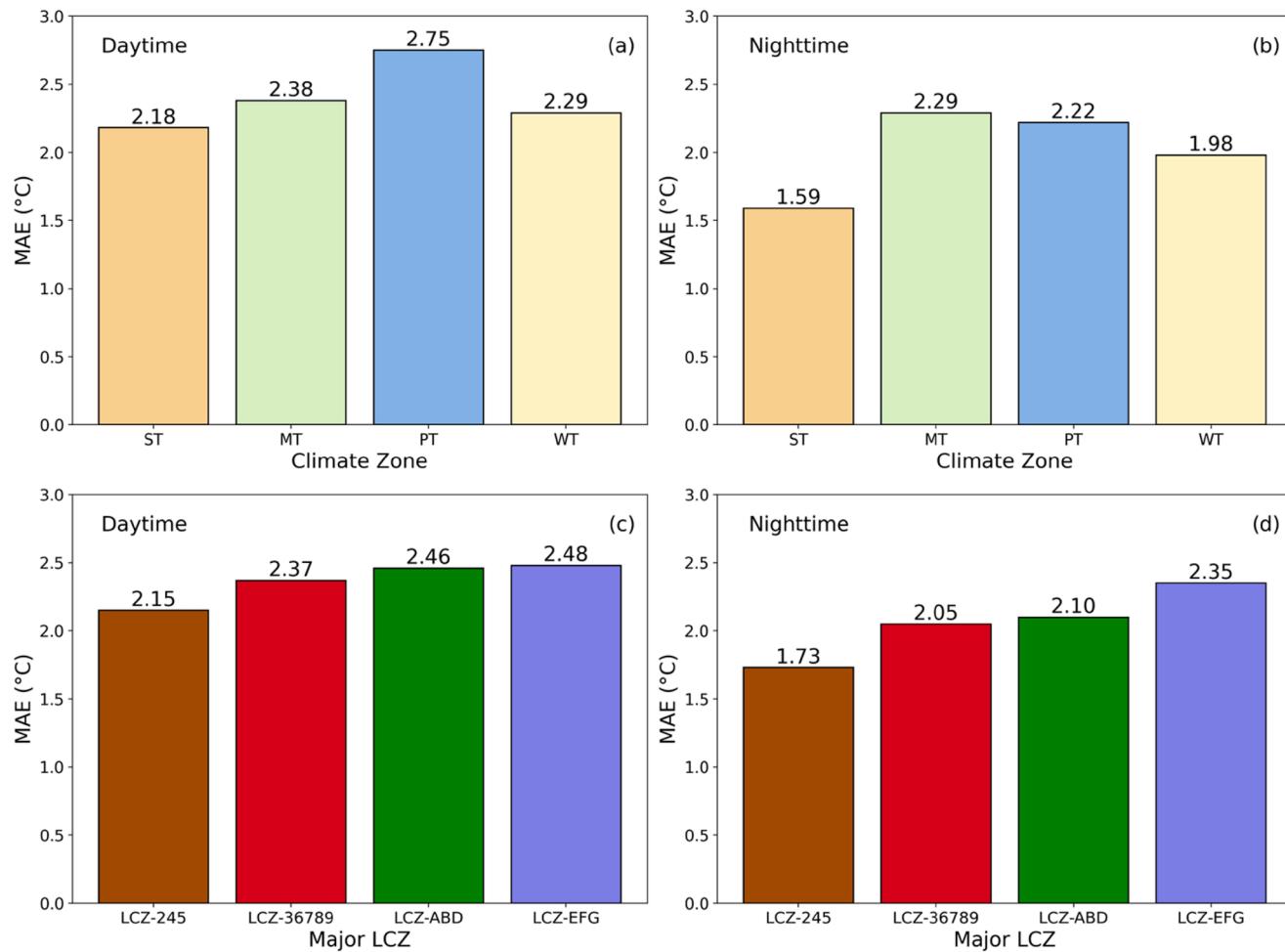


Fig. 18. Predictive accuracy (MAE) for testing data in T_a estimation by individual LightGBM models trained by records within different climate zones (ST, MT, PT, and WT) and major LCZ types.

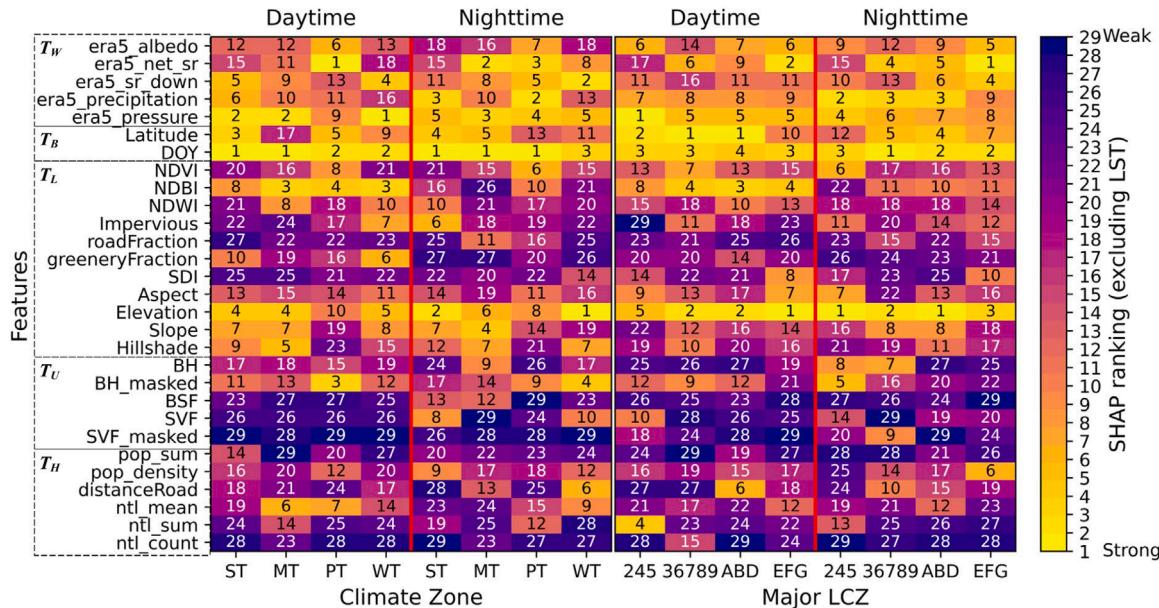


Fig. 19. Rankings of the SHAP values among the influential factors (excluding LST) based on the individual LightGBM models trained by the datasets within different climate zones (ST, MT, PT, and WT) and major LCZ categories.

built-up LCZ types during nighttime, and SVF for high-rise built-up regions, reflecting critical impacts. Overall, the SHAP values in Ta modeling processes reveal the diverse functional influences of different factors on the spatiotemporal distribution of Ta, highlighting the importance of evaluating variables across various regional and local climate conditions. Consequently, predictive performance evaluations and feature assessments were carried out across different weather stations, categorized by various types of climate zones and LCZs. The results of SHAP values in different regions demonstrate that LCZ can serve as a promising tool for guiding essential feature selections under specific spatiotemporal contexts, better exhibiting the actual heterogeneity of Ta, potentially improving Ta estimation via LCZ-specific models, and offering insights for UHI-related studies.

6. Conclusion and future works

This study conducts an evaluative analysis of the impacts of LST and multi-source ambient factors on the ML-based estimation of Ta in 292 weather stations in China. It demonstrates the capacity of SHAP value-based interpretation to identify sensitive ambient factors for Ta estimation and inconsistencies across different regional and local climate types.

The LightGBM-based models exhibited the optimal and applicable predictive accuracies for daytime and nighttime datasets. Several spatial and temporal patterns were observed: nighttime datasets outperformed daytime ones in Ta estimation performances; coastal weather stations achieved better estimation accuracy during the daytime, while stations in the northern parts achieved relatively low predictive accuracy at nighttime. Nevertheless, despite LightGBM's robustness in multivariate ML predictive modeling, it is still a data-driven approach that heavily depends on the data quality. Therefore, it is strongly recommended that more standardized data sources with high spatiotemporal resolution be incorporated into future research to investigate the diverse impacts of factors across the macro and local contexts. Additionally, model stacking might serve as a viable solution for better integrating the strength of various local models to improve overall modeling accuracy. While the current model training approach adequately meets the objectives of this study, diverse accuracy assessment solutions still need to be further explored in future studies to evaluate the generalizability of advanced models for regions lacking station coverage. More importantly, our findings indicate that Ta estimation under different LCZs, seasons, and landscapes follows diverse patterns and processes, highlighting the necessity of uncovering their underlying theoretical framework and interactive mechanisms among both macroclimates and local ambient influential factors for future improvements.

Furthermore, the SHAP-based interpretation analyses pointed out the significance of incorporating dynamic meteorological and terrain-related variables to enhance estimation accuracy, as well as building landscape-related variables such as NDBI, BH and SVF for more accurate Ta estimation in urban regions. As shown by diverse SHAP value distribution patterns, regional climate types and LCZs should be regarded as valuable tools for key feature selection, region-specific modeling and multi-scale dynamic Ta mapping in future studies. Thus, leveraging this study as a foundation, more promising ML and physical-based models should be further explored and analyzed with multisource geospatial data of influential factors in future studies. This will help causatively disclose Ta variation under the mixed effects of physical and anthropogenic processes, thereby enhancing Ta estimation and mapping for better UHI mitigation-oriented studies and strategy planning.

Data availability: Data will be made available on reasonable request.

CRediT authorship contribution statement

Songyang Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Man Sing Wong:** Writing – review & editing,

Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Rui Zhu:** Writing – review & editing, Supervision, Methodology, Investigation. **Guoqiang Shi:** Writing – review & editing, Validation. **Jinxin Yang:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was supported by the General Research Fund [Grant No. 15603920, 15609421, 15603923], and the Collaborative Research Fund [Grant No. C5062-21GF and C6003-22Y] from the Research Grants Council, Hong Kong, as well as the funding support from the Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hong Kong, China (Grant No. 1-BBG2). The authors would like to thank the National Oceanic and Atmospheric Administration, the U.S., for providing the dataset of near-surface air temperature by weather stations. We also thank the National Aeronautics and Space Administration (NASA) of the United States, the European Centre for Medium-Range Weather Forecasts (ECMWF), World Resources Institute, WorldPop, Global Roads Inventory Project (GRIP), and the other data source as cited in the Appendix A for their collection and free distribution of the geospatial dataset used in this study. We would also like to thank the editors and anonymous reviewers who provided constructive comments on the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.scs.2025.106257](https://doi.org/10.1016/j.scs.2025.106257).

Data availability

Data will be made available on request.

References

- Agathangelidis, I., Cartalis, C., & Santamouris, M. (2016). Estimation of air temperatures for the urban agglomeration of athens with the use of satellite data. *Geoinformatics & Geostatistics: An Overview*, 2016. <https://doi.org/10.4172/2327-4581.1000139>
- Akhtar, R., Gupta, P.T., & Srivastava, A.K. (2016). Urbanization, urban heat island effects and dengue outbreak in Delhi. In R. Akhtar (Ed.), *Climate change and human health scenario in south and southeast asia* (pp. 99–111). Springer International Publishing. https://doi.org/10.1007/978-3-319-23684-1_7.
- Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: An analytical review. *WIREs Data Mining and Knowledge Discovery*, 11(5), e1424. <https://doi.org/10.1002/widm.1424>
- Azevedo, J. A., Chapman, L., & Muller, C. L. (2016). Quantifying the daytime and nighttime urban heat island in Birmingham, UK: A comparison of satellite derived land surface temperature and high resolution air temperature observations. *Remote Sensing*, 8(2), Article 2. <https://doi.org/10.3390/rs8020153>. Article.
- Balas, D. B., Tiwari, M. K., Trivedi, M., & Patel, G. R. (2023). Impact of land surface temperature (LST) and ground air temperature (Ta) on land use and land cover (LULC): An investigative study. *International Journal of Environment and Climate Change*, 13(10), 3117–3130. <https://doi.org/10.9734/ijecc/2023/v13i102980>
- Belgiu, M., & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *Isprs Journal of Photogrammetry & Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.05.001>
- Benavides, R., Montes, F., Rubio, A., & Osoro, K. (2007). Geostatistical modelling of air temperature in a mountainous region of Northern Spain. *Agricultural and Forest Meteorology*, 146(3–4), 173–188. <https://doi.org/10.1016/j.agrformet.2007.05.014>
- Breiman, L. (1984). *Nonlinear discriminant analysis via scaling and ace*. Department of Statistics, University of California.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., & Tait, A. M. (2022).

- Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*, 9(1), 251. <https://doi.org/10.1038/s41597-022-01307-4>
- Burnett, M., & Chen, D. (2021). The impact of seasonality and land cover on the consistency of relationship between air temperature and LST derived from landsat 7 and MODIS at a local scale: A case study in Southern Ontario. *Land*, 10(7), Article 7. <https://doi.org/10.3390/land10070672>. Article.
- CAS. (n.d.). Resource and environmental science data platform. Retrieved April 9, 2024, from <https://www.resdc.cn/data.aspx?DATAID=243>.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- 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). <https://doi.org/10.1145/2939672.2939785>
- Chen, Y., Liang, S., Ma, H., Li, B., He, T., & Wang, Q. (2021a). An all-sky 1 km daily land surface air temperature product over mainland China for 2003–2019 from MODIS and ancillary data. *Earth System Science Data*, 13(8), 4241–4261. <https://doi.org/10.5194/essd-13-4241-2021>
- Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wang, C., Wu, B., & Wu, J. (2021b). An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth System Science Data*, 13(3), 889–906. <https://doi.org/10.5194/essd-13-889-2021>
- Cho, D., Yoo, C., Im, J., & Cha, D.-H. (2020). Comparative assessment of various machine learning-based bias correction methods for numerical weather prediction model forecasts of extreme air temperatures in Urban Areas. *Earth and Space Science*, 7(4), Article e2019EA000740. <https://doi.org/10.1029/2019EA000740>
- Cho, D., Yoo, C., Im, J., Lee, Y., & Lee, J. (2020). Improvement of spatial interpolation accuracy of daily maximum air temperature in urban areas using a stacking ensemble technique. *GIScience & Remote Sensing*, 57(5), 633–649. <https://doi.org/10.1080/15481603.2020.1766768>
- Copernicus Climate Change Service. (2019). ERA5-Land monthly averaged data from 2001 to present. *ECMWF*. [</Dataset>](https://doi.org/10.24381/CDS.68D2BB30)
- Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., & Bechtel, B. (2022). A global map of local climate zones to support earth system modelling and urban-scale environmental science. *Earth System Science Data*, 14(8), 3835–3873. <https://doi.org/10.5194/essd-14-3835-2022>
- Deng, S., Su, J., Zhu, Y., Yu, Y., & Xiao, C. (2024). Forecasting carbon price trends based on an interpretable light gradient boosting machine and Bayesian optimization. *Expert Systems with Applications*, 242, Article 122502. <https://doi.org/10.1016/j.eswa.2023.122502>
- Doick, K., Hutchings, T., & others. (2013). *Air temperature regulation by urban trees and green infrastructure*. Forestry Commission.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2), Article 2005RG000183. <https://doi.org/10.1029/2005RG000183>
- Gandhi, U. (2021). Working with qa bands and bitmasks in google earth engine. August 18. Spatial Thoughts <https://spatialthoughts.com/2021/08/19/qa-bands-bitmasks-gee/>.
- Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data?. In *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. https://openreview.net/forum?id=Fp7_phQszn.
- Häntzschel, J., Goldberg, V., & Bernhofer, C. (2005). GIS-based regionalisation of radiation, temperature and coupling measures in complex terrain for low mountain ranges. *Meteorological Applications*, 12(1), 33–42. <https://doi.org/10.1017/S1350482705001489>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Irisko, J., Ramamurthy, P., Yu, Y., Yu, P., & Melecio-Vázquez, D. (2020). Urban air temperature model using GOES-16 LST and a diurnal regressive neural network algorithm. *Remote Sensing of Environment*, 237, Article 111495. <https://doi.org/10.1016/j.rse.2019.111495>
- Huang, H., Deng, X., Yang, H., Zhou, X., & Jia, Q. (2020). Spatio-temporal mechanism underlying the effect of urban heat island on cardiovascular diseases. *Iranian Journal of Public Health*, 49(8), 1455–1466. <https://doi.org/10.18502/ijph.v49i8.3889>
- Huang, H., Yang, H., Deng, X., Zeng, P., Li, Y., Zhang, L., & Zhu, L. (2019). Influencing Mechanisms of Urban Heat Island on Respiratory Diseases. *Iranian Journal of Public Health*, 48(9), 1636–1646.
- Kawauchi, H., & Fuse, T. (2022). SHAP-Based Interpretable Object Detection Method for Satellite Imagery. *Remote Sensing*, 14(9), 1970. <https://doi.org/10.3390/rs14091970>
- 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. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 3149–3157).
- Khamchiangta, D., & Dhakal, S. (2019). Physical and non-physical factors driving urban heat island: Case of Bangkok Metropolitan Administration, Thailand. *Journal of Environmental Management*, 248, Article 109285. <https://doi.org/10.1016/j.jenvman.2019.109285>
- Khan, A., Chatterjee, S., & Weng, Y. (2021). 2—Characterizing thermal fields and evaluating UHI effects. In A. Khan, S. Chatterjee, & Y. Weng (Eds.), *Urban heat island modeling for tropical climates* (pp. 37–67). Elsevier. <https://doi.org/10.1016/B978-0-12-819669-4.00002-7>
- Kotharkar, R., Dongarsane, P., & Keskar, R. (2023). Determining influence of urban morphology on air temperature and heat index with hourly emphasis. *Building and Environment*, 233, Article 110044. <https://doi.org/10.1016/j.buildenv.2023.110044>
- Kurtzman, D., & Kadmon, R. (1999). Mapping of temperature variables in Israel: Sa comparison of different interpolation methods. *Climate Research*, 13(1), 33–43. <https://doi.org/10.3354/cr013033>
- Laukkarinen, A., & Vinha, J. (2024). Long-term prediction of hourly indoor air temperature using machine learning. *Energy and Buildings*, 325, Article 114972. <https://doi.org/10.1016/j.enbuild.2024.114972>
- Lee, T., Wong, W., & Tam, K. (2018). Urban-focused weather and climate services in Hong Kong. *Geoscience Letters*, 5(1), 18. <https://doi.org/10.1186/s40562-018-0119-6>
- Liu, Z.-H., Weng, S.-S., Zeng, Z.-L., Ding, M.-H., Wang, Y.-Q., & Liang, Z. (2024). Hourly land surface temperature retrieval over the Tibetan Plateau using Geo-LightGBM framework: Fusion of Himawari-8 satellite, ERA5 and site observations. *Advances in Climate Change Research*, 15(4), 623–635. <https://doi.org/10.1016/j.accre.2024.06.007>
- Lowry, W. P. (1977). Empirical estimation of urban effects on climate: A problem analysis. *Journal of Applied Meteorology and Climatology*, 16(2), 129–135. [https://doi.org/10.1175/1520-0450\(1977\)016<0129:EEUEO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1977)016<0129:EEUEO>2.0.CO;2)
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, 13(6), Article 064006. <https://doi.org/10.1088/1748-9326/aabdf4>
- Mokhov, I. I., Smirnov, D. A., & Karpenko, A. A. (2012). Assessments of the relationship of changes of the global surface air temperature with different natural and anthropogenic factors based on observations. *Doklady Earth Sciences*, 443(1), 381–387. <https://doi.org/10.1134/S1028334X12030178>
- Mutibwa, D., Strachan, S., & Albright, T. (2015). Land surface temperature and surface air temperature in Complex Terrain. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(10), 4762–4774. <https://doi.org/10.1109/JSTARS.2015.2468594>. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- Nikolaou, N., Dallavalle, M., Stafoggia, M., Bouwer, L. M., Peters, A., Chen, K., Wolf, K., & Schneider, A. (2023). High-resolution spatiotemporal modeling of daily near-surface air temperature in Germany over the period 2000–2020. *Environmental Research*, 219, Article 115062. <https://doi.org/10.1016/j.envres.2022.115062>
- Oke, T. R. (1973). City size and the urban heat island. *Atmospheric Environment*, 7(8), 769–779. [https://doi.org/10.1016/0004-6981\(73\)90140-6](https://doi.org/10.1016/0004-6981(73)90140-6)
- Oke, T. R. (2002). *Boundary layer climates*. Routledge.
- Oswald, E. M., Rood, R. B., Zhang, K., Gronlund, C. J., O'Neill, M. S., White-Newsome, J. L., Brines, S. J., & Brown, D. G. (2012). An investigation into the spatial variability of near-surface air temperatures in the Detroit, Michigan, Metropolitan Region. *Journal of Applied Meteorology and Climatology*, 51(7), 1290–1304. <https://doi.org/10.1175/JAMC-D-11-0127.1>
- Peng, J., Jia, J., Liu, Y., Li, H., & Wu, J. (2018). Seasonal contrast of the dominant factors for spatial distribution of land surface temperature in urban areas. *Remote Sensing of Environment*, 215, 255–267. <https://doi.org/10.1016/j.rse.2018.06.010>
- Peng, X., She, Q., Long, L., Liu, M., Xu, Q., Zhang, J., & Xiang, W. (2017). Long-term trend in ground-based air temperature and its responses to atmospheric circulation and anthropogenic activity in the Yangtze River Delta, China. *Atmospheric Research*, 195, 20–30. <https://doi.org/10.1016/j.atmosres.2017.05.013>
- Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, 14(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- Python Package Introduction—Xgboost 1.7.5 documentation*. (n.d.). Retrieved April 18, 2023, from https://xgboost.readthedocs.io/en/stable/python/python_intro.html
- RandomForestRegressor. (n.d.). Scikit-Learn. Retrieved November 29, 2024, from <https://scikit-learn/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>
- Roxon, J., Ulm, F.-J., & Pellenq, R. J.-M. (2020). Urban heat island impact on state residential energy cost and CO₂ emissions in the United States. *Urban Climate*, 31, Article 100546. <https://doi.org/10.1016/j.ulclim.2019.100546>
- Santamouris, M., Cartalis, C., Synnefa, A., & Kolokotsa, D. (2015). On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—A review. *Energy and Buildings*, 98, 119–124. <https://doi.org/10.1016/j.enbuild.2014.09.052>
- Shi, Y., Ke, G., Soukhavong, D., Lamb, J., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., Titov, N., & Cortes, D. (2024). *lightgbm: Light Gradient Boosting Machine*. <https://github.com/Microsoft/LightGBM>
- Sklearn.model_selection.RandomizedSearchCV. (n.d.). Scikit-Learn. Retrieved February 24, 2024, from https://scikit-learn/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html
- Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent developments and partnerships. *Bulletin of the American Meteorological Society*, 92(6), 704–708. <https://doi.org/10.1175/2011BAMS3015.1>
- Sorichetta, A., Hornby, G. M., Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). High-resolution gridded population datasets for Latin America and the

- Caribbean in 2010, 2015, and 2020. *Scientific Data*, 2(1), Article 150045. <https://doi.org/10.1038/sdata.2015.45>
- Stewart, I. D., Krayenhoff, E. S., Voogt, J. A., Lachapelle, J. A., Allen, M. A., & Broadbent, A. M. (2021). Time evolution of the surface urban heat island. *Earth's Future*, 9(10), Article e2021EF002178. <https://doi.org/10.1029/2021EF002178>
- Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879–1900. <https://doi.org/10.1175/BAMS-D-11-00019.1>
- Su, P., Abera, T., Guan, Y., & Pellikka, P. (2023). Image-to-image training for spatially seamless air temperature estimation with satellite images and station data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 3353–3363. <https://doi.org/10.1109/JSTARS.2023.3256363>. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A., & Doulamis, N. (2023). Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP. *IEEE Geoscience and Remote Sensing Letters*, 20, 1–5. <https://doi.org/10.1109/LGRS.2023.3251652>. IEEE Geoscience and Remote Sensing Letters.
- Thompson, R., Hornigold, R., Page, L., & Waite, T. (2018). Associations between high ambient temperatures and heat waves with mental health outcomes: A systematic review. *Public Health*, 161, 171–191. <https://doi.org/10.1016/j.puhe.2018.06.008>
- Tian, L., Hao, T., He, X., Chan, I., Niu, J., Chan, P. W., Ng, W. Y., & Huang, J. (2024). Examining the non-linear relationship between urban form and air temperature at street level: A case of Hong Kong. *Building and Environment*, 264, Article 111884. <https://doi.org/10.1016/j.buildenv.2024.111884>
- Varentsov, M., Konstantinov, P., Baklanov, A., Esau, I., Miles, V., & Davy, R. (2018). Anthropogenic and natural drivers of a strong winter urban heat island in a typical Arctic city. *Atmospheric Chemistry and Physics*, 18(23), 17573–17587. <https://doi.org/10.5194/acp-18-17573-2018>
- Venter, Z. S., Brousse, O., Esau, I., & Meier, F. (2020). Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sensing of Environment*, 242, Article 111791. <https://doi.org/10.1016/j.rse.2020.111791>
- Vermote, E., & Wolfe, R. (2021). MODIS/Terra surface reflectance daily L2G global 1km and 500m sin grid V061. NASA EOSDIS Land Processes Distributed Active Archive Center. [</Dataset>](https://doi.org/10.5067/MODIS/MOD09GA.061)
- Wan, Z., Hook, S., & Hulley, G. (2021). MODIS/Terra land surface temperature/emissivity daily L3 global 1km sin grid V061. NASA EOSDIS Land Processes Distributed Active Archive Center. [</Dataset>](https://doi.org/10.5067/MODIS/MOD11A1.061)
- Wang, H., Liang, Q., Hancock, J. T., & Khoshgoftaar, T. M. (2024). Feature selection strategies: A comparative analysis of SHAP-value and importance-based methods. *Journal of Big Data*, 11(1), 44. <https://doi.org/10.1186/s40537-024-00905-w>
- Wang, H., Yang, J., Chen, G., Ren, C., & Zhang, J. (2023a). Machine learning applications on air temperature prediction in the urban canopy layer: A critical review of 2011–2022. *Urban Climate*, 49, Article 101499. <https://doi.org/10.1016/j.ulclim.2023.101499>
- Wang, M., Li, Y., Yuan, H., Zhou, S., Wang, Y., Adnan Ikram, R. M., & Li, J. (2023b). An XGBoost-SHAP approach to quantifying morphological impact on urban flooding susceptibility. *Ecological Indicators*, 156, Article 111137. <https://doi.org/10.1016/j.ecolind.2023.111137>
- Wang, R., Ren, C., Xu, Y., Lau, K. K.-L., & Shi, Y. (2018). Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: A case study of Hong Kong. *Urban Climate*, 24, 567–576. <https://doi.org/10.1016/j.ulclim.2017.10.001>
- Wang, Y., Guo, Z., & Han, J. (2021). The relationship between urban heat island and air pollutants and them with influencing factors in the Yangtze River Delta, China. *Ecological Indicators*, 129, Article 107976. <https://doi.org/10.1016/j.ecolind.2021.107976>
- Wang, Y., & Wu, L. (2016). On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. *Energy*, 112, 208–220. <https://doi.org/10.1016/j.energy.2016.06.075>
- WMO, G. (1996). *Guide to meteorological instruments and methods of observation*.
- Wu, W.-B., Ma, J., Banzhaf, E., Meadows, M. E., Yu, Z.-W., Guo, F.-X., Sengupta, D., Cai, X.-X., & Zhao, B. (2023). A first Chinese building height estimate at 10 m resolution (CNBH-10 m) using multi-source earth observations and machine learning. *Remote Sensing of Environment*, 291, Article 113578. <https://doi.org/10.1016/j.rse.2023.113578>
- Yadav, N., Rajendra, K., Awasthi, A., Singh, C., & Bhushan, B. (2023). Systematic exploration of heat wave impact on mortality and urban heat island: A review from 2000 to 2022. *Urban Climate*, 51, Article 101622. <https://doi.org/10.1016/j.ulclim.2023.101622>
- Yao, R., Wang, L., Huang, X., Cao, Q., Wei, J., He, P., Wang, S., & Wang, L. (2023). Global seamless and high-resolution temperature dataset (GSHTD), 2001–2020. *Remote Sensing of Environment*, 286, Article 113422. <https://doi.org/10.1016/j.rse.2022.113422>
- Yu, Z., Chen, S., & Wong, N. H. (2020). Temporal variation in the impact of urban morphology on outdoor air temperature in the tropics: A campus case study. *Building and Environment*, 181, Article 107132. <https://doi.org/10.1016/j.buldev.2020.107132>
- Zekar, A., Milojevic-Dupont, N., Zumwald, M., Wagner, F., & Creutzig, F. (2023). Urban form features determine spatio-temporal variation of ambient temperature: A comparative study of three European cities. *Urban Climate*, 49, Article 101467. <https://doi.org/10.1016/j.ulclim.2023.101467>
- Zhang, T., Zhou, Y., Zhao, K., Zhu, Z., Chen, G., Hu, J., & Wang, L. (2022a). A global dataset of daily maximum and minimum near-surface air temperature at 1 km resolution over land (2003–2020). *Earth System Science Data*, 14(12), 5637–5649. <https://doi.org/10.5194/essd-14-5637-2022>
- Zhang, X., Liu, L., Zhao, T., Gao, Y., Chen, X., & Mi, J. (2022b). GISD30: Global 30 m impervious-surface dynamic dataset from 1985 to 2020 using time-series Landsat imagery on the Google Earth Engine platform. *Earth System Science Data*, 14(4), 1831–1856. <https://doi.org/10.5194/essd-14-1831-2022>
- Zhang, X., & Yan, X. (2014). Temporal change of climate zones in China in the context of climate warming. *Theoretical and Applied Climatology*, 115(1–2), 167–175. <https://doi.org/10.1007/s00704-013-0887-z>
- Zhu, S., Yan, Y., Zhao, B., & Wang, H. (2025). Assessing the impact of adjacent urban morphology on street temperature: A multisource analysis using random forest and SHAP. *Building and Environment*, 267, Article 112326. <https://doi.org/10.1016/j.buldev.2024.112326>
- Zhu, W., Lü, A., & Jia, S. (2013a). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment*, 130, 62–73. <https://doi.org/10.1016/j.rse.2012.10.034>
- Zhu, W., Lü, A., & Jia, S. (2013b). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment*, 130, 62–73. <https://doi.org/10.1016/j.rse.2012.10.034>
- Zhu, X., Guo, H., Huang, J. J., Tian, S., Xu, W., & Mai, Y. (2022). An ensemble machine learning model for water quality estimation in coastal area based on remote sensing imagery. *Journal of Environmental Management*, 323, Article 116187. <https://doi.org/10.1016/j.jenvman.2022.116187>