

Identifying driving factors influencing summer air temperature and developing a high-resolution temperature map using regression Kriging in Seoul

Younju Lee & Changyeon Lee

To cite this article: Younju Lee & Changyeon Lee (05 May 2026): Identifying driving factors influencing summer air temperature and developing a high-resolution temperature map using regression Kriging in Seoul, Geo-spatial Information Science, DOI: [10.1080/10095020.2026.2664303](https://doi.org/10.1080/10095020.2026.2664303)

To link to this article: <https://doi.org/10.1080/10095020.2026.2664303>



© 2026 Wuhan University. Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 05 May 2026.



[Submit your article to this journal](#)



Article views: 292



[View related articles](#)



[View Crossmark data](#)

Identifying driving factors influencing summer air temperature and developing a high-resolution temperature map using regression Kriging in Seoul

Younju Lee  and Changyeon Lee 

Department of Urban Design and Studies, Chung-Ang University, Seoul, Republic of Korea

ABSTRACT

Urban temperature increases present significant challenges, necessitating the identification of key influencing factors and developing high-resolution temperature maps for effective climate analysis. This study constructs a 100 m-resolution near-surface air temperature map for Seoul, South Korea, by integrating “a distance decay regression selection strategy (ADDRESS)” with “regression Kriging (RK)” to enhance spatial temperature modeling. The analysis incorporates land use and land cover, building morphology, and geographical and topographical factors to determine the spatial influence range of each variable. The study uses smart Seoul urban data sensor (S-DoT) network data to ensure dense, spatially continuous observations, improving model accuracy. Results indicate that temperature influences vary across spatial extents, with green spaces, water bodies, and agricultural land exerting strong cooling effects, while built-environment factors such as gross floor area and building coverage ratio contribute to urban heat. While RK effectively reduces spatial bias in residuals, it does not significantly enhance numerical accuracy compared to the regression model. The final model achieved an RMSE of 0.4704°C and an MAE of 0.3517°C based on 20-fold cross-validation. These findings underscore the importance of optimal variable selection and spatial influence assessment in urban temperature modeling. The study provides valuable insights for urban planners and policymakers in mitigating urban heat and identifying heat-vulnerable regions. However, limitations such as incomplete airflow representation and data gaps in suburban areas remain. Future research should incorporate additional environmental factors, such as shading effects and wind dynamics, and expand sensor coverage to improve predictive accuracy.

ARTICLE HISTORY

Received 3 August 2025
Accepted 20 April 2026

KEYWORDS

Urban temperature modeling; regression Kriging; urban morphology; urban heat mitigation; built environment; geospatial analysis

1. Introduction

High summer temperatures in urban areas are a critical environmental concern. Elevated temperatures are associated with increased heat stress, heightened risks of heat-related morbidity and mortality, and adverse health effects due to the formation of tropospheric ozone (Ballester et al. 2023; Gabriel and Endlicher 2011; Jang et al. 2020). For instance, in Europe, each 1°C increase in summer temperature has been linked to 35.3 additional summer heat-related deaths per million (Ballester et al. 2023). A study conducted in South Korea found that a 0.1°C rise in urban temperatures, compared to surrounding areas, can lead to a 2%–3% increase in heat-related mortality risk (Jang et al. 2020). High summer temperatures not only lead to increased health risk but also affect energy consumption. Increased reliance on air conditioning raises electricity demand, exacerbating air pollution and greenhouse gas emissions (Kinney 2008). Additionally, temperature variations within cities, influenced by urban morphology, can

contribute to environmental injustice, disproportionately affecting vulnerable communities (Gonzalez-Trevizo et al. 2021; Hart and Sailor 2009; Yin et al. 2018). While extreme temperatures are also important, our objective is to model the persistent spatial thermal patterns determined by urban morphology. Multi-year average summer temperature provides a stable indicator of structural heat conditions, whereas maximum daily temperatures are strongly influenced by short-term weather fluctuations unrelated to built-environment characteristics.

Addressing urban heat effectively requires spatially detailed temperature data. Current temperature data sources include remote sensing and ground-based monitoring stations. Satellite-derived land surface temperature (LST) data provides broad spatial coverage but has limitations such as cloud obstruction, infrequent updates, and discrepancies with near-surface air temperatures (Chen et al. 2022). In contrast, ground-based air temperature measurements offer greater accuracy but are sparsely distributed,

limiting their ability to capture fine-scale urban temperature variations. These challenges underscore the need for high-resolution, spatially continuous air temperature models.

Various spatial prediction models have been developed to estimate urban air temperature across continuous surfaces, typically categorized into numerical climate models, statistical methods, and machine learning approaches. Numerical climate models, such as the Weather Research and Forecasting (WRF) model and microclimatic models like ENVI-Met and UrbClim, simulate urban temperature patterns based on physical parameters (Brousse et al. 2016; Yin et al. 2018). However, their reliance on fixed input parameters limits their adaptability to regional temperature variations. Moreover, WRF's resolution (approximately 0.9 km) restricts its ability to capture intra-urban temperature differences (Zhang et al. 2022). While computational fluid dynamics (CFD), urban canopy models (UCM), and ENVI-Met provide high-detail simulations at the neighborhood scale, their significant computational demands make them impractical for city-wide applications (Kadaverugu et al. 2021; Liu et al. 2021; Setyantho, Yuan, and Heo 2024). UrbClim, though more efficient than traditional mesoscale models, depends primarily on land cover datasets, limiting its ability to incorporate additional influencing variables (De Ridder, Lauwaet, and Maiheu 2015).

Machine learning techniques, particularly decision tree-based models such as random forest, have been widely applied to spatial temperature prediction (Cho et al. 2020; Han et al. 2022b; Venter et al. 2020; Xin et al. 2025; Zumwald et al. 2021). A recent study proposed TranSAT, which combines U-Net, deep neural network, and transfer learning (Wang, Brönnimann, et al. 2025). However, their reliance on training datasets reduces their reliability when extrapolating beyond observed conditions (Hengl et al. 2018; Wang et al. 2023). Therefore, machine learning methods are less suitable for developing sensor-based urban air temperature models, as sensor data often exhibit high uncertainty at unobserved locations, and values outside the training set range are likely to occur in practice.

A promising alternative involves statistical models that consider both influential factors and a spatial relationship, rather than standalone regression or interpolation methods. Both geographically weighted regression (GWR) and regression Kriging (RK) are regression-based spatial interpolation methods that aim to improve prediction accuracy by explicitly incorporating the spatial variation in the relationships between the target variable and explanatory covariates. GWR estimates spatially varying relationships by calibrating location-specific coefficients using a bandwidth, which defines the geographical extent

over which neighboring data points influence each local regression model (Gao et al. 2020; Jaber and Sengupta 2024; Li et al. 2010; Zhang, Middel, and Turner 2019). However, its reliance on bandwidth optimization (Gao et al. 2020) and the assumption that all variable effects vary across space (Li et al. 2010) can be problematic in urban areas, where clear criteria for delineating homogeneous regions are lacking and environmental conditions tend to be relatively uniform. Regression Kriging, in contrast, combines multiple linear regression (LR) with Kriging interpolation to enhance predictive accuracy. RK first identifies key influencing factors through regression analysis and then applies Kriging to account for spatial correlations among observations. However, many studies (Li, Griffith, and Shu 2020; Sekulić et al. 2020; Wu and Li 2013) have relied on limited predictor variables, such as latitude, longitude, and elevation, without fully considering the diverse urban factors that influence temperature variation.

Understanding the key drivers of urban summer temperature is essential for mitigating heat impacts. Recent studies have explored the relationship between microclimate and factors such as urban morphology, topography, and population density, emphasizing their roles in heat emission, absorption, and air circulation (Alavipanah et al. 2018; Hart and Sailor 2009; Peng et al. 2018; Venter et al. 2020). The influence of urban morphology on temperature can be categorized into two-dimensional (2D) and three-dimensional (3D) components. While 2D morphological factors include land use and land cover (LULC) (e.g. vegetated areas, impervious surfaces, water bodies) and urban morphology indicators (e.g. building coverage ratio) (Chun and Guldmann 2014; Estoque, Murayama, and Myint 2017; Yin et al. 2018), recent research has increasingly focused on the 3D built environment, including building height, volume, and shape (Alavipanah et al. 2018; Berger et al. 2017; Huang and Wang 2019; Park and Kim 2023).

Despite these advancements, many studies have not sufficiently addressed the spatial extent and influence of temperature-driving factors. Processes such as air convection between neighborhoods and the varying spatial influence of different variables require further investigation (Brousse et al. 2016). Some studies assess urban spatial factors using fixed units, such as grids or census blocks (Nikolaou et al. 2023; Yin et al. 2018), while others apply uniform distances across all variables (Park, Ha, and Lee 2017; Park and Kim 2023). However, only a few studies (e.g. Shi et al. 2019) have accounted for variations in the spatial influence of different variables when analyzing the urban thermal environment. Since neighboring features significantly affect temperature distribution, accurately determining these spatial extents is crucial for developing high-resolution spatial temperature maps.

Additionally, spatial interpolation requires a sufficient number of observed points to reliably capture spatial patterns and variability. A higher sample density enhances model precision, particularly for Kriging, which requires at least 100 samples for reliable estimations (Webster and Oliver 1992). However, beyond a certain threshold, additional samples yield diminishing returns (Li and Heap 2008). Many interpolation methods rely solely on spatial coverage and are constrained by sparse monitoring stations, leading to inaccuracies in small-area temperature variations (Vienneau, de Hoogh, and Briggs 2009). Optimizing sample density is essential to improving interpolation accuracy while avoiding unnecessary data collection.

To address these gaps, this study aims to:

- (1) Develop a high-resolution (100 m) spatial temperature model for Seoul using regression Kriging (RK).
- (2) Identify key urban determinants of summer temperature, including LULC, building morphology, and topographical factors.
- (3) Optimize spatial influence ranges for each variable using a distance decay regression selection strategy (ADDRESS).

Using data from the smart Seoul urban data sensor (S-DoT) network, this study ensures sufficient sample density for accurate temperature modeling. It also incorporates spatially independent variables representing specific urban conditions that influence summer temperatures. ADDRESS is used to determine the optimal spatial influence range for each variable, ensuring a precise representation of urban heat

distribution. Finally, a high-resolution summer temperature map is developed using RK, integrating linear regression with Kriging.

This study makes three key contributions to the literature on urban air-temperature modeling. First, we apply regression Kriging to combine deterministic temperature – environment relationships with spatial dependence in residuals, addressing a major limitation of both machine-learning and numerical climate models. Second, we integrate high-density S-DoT sensor data with detailed 2D and 3D urban morphological variables, enabling a city-scale 100 m air-temperature model with improved spatial bias reduction. Third, we determine variable-specific spatial influence ranges using ADDRESS, providing a physically informed and reproducible framework for interpreting how urban features affect temperature across different spatial scales.

The following section describes the study area, data sources, and methods. Section 3 presents the results of statistical analyses, and Section 4 discusses the key findings. Finally, Section 5 summarizes the study's conclusions, highlights its limitations, and suggests directions for future research.

2. Methodology

2.1. Study area

The study focuses on Seoul, the capital of South Korea (Figure 1), located in the temperate climate zone at a longitude of 126.97°E and latitude of 37.56°N. Seoul covers an area of 605.2 km² and has a population of 9,579,177 people as of 2025 (Seoul Metropolitan

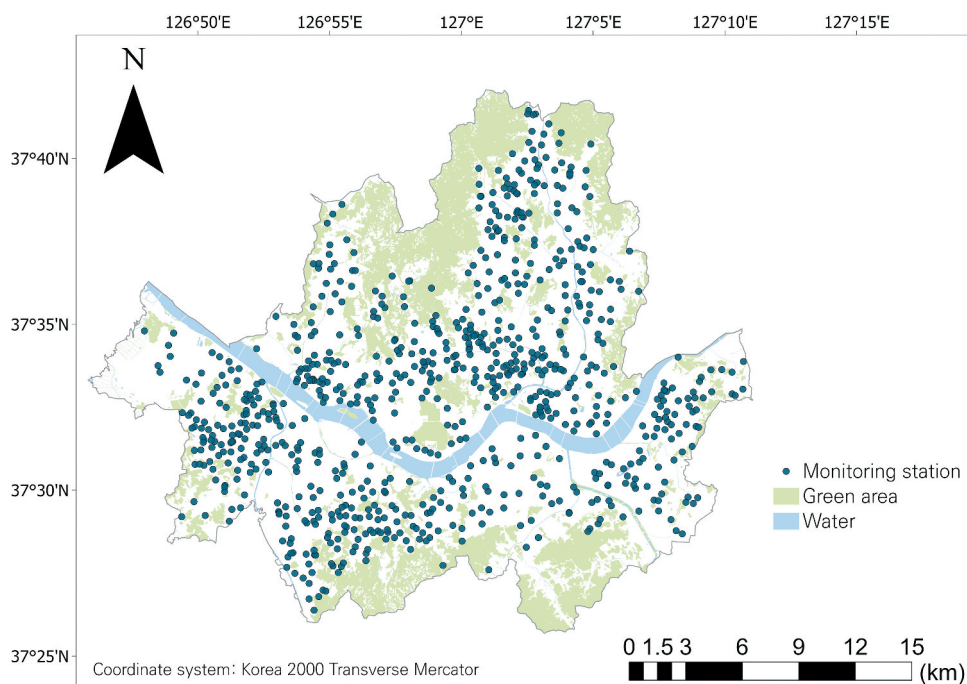


Figure 1. The study area and spatial distribution of monitoring stations (832 sensors).

Government 2026). The study examines summer temperatures from June to August during the years 2021 to 2023. By using data from three years, we account for year-to-year variations and minimize the impact of outliers. The analysis is conducted at a spatial resolution of 100 m × 100 m, which is the finest grid size available for the dataset.

2.2. Measuring summer temperature using S-DoT

We used the three-year average summer temperature to capture stable and persistent spatial temperature patterns. While daily maximum temperatures fluctuate significantly with synoptic weather conditions, multi-year averages better reflect the long-term spatial thermal environment shaped by urban morphology. To overcome the limitations of the sparse meteorological observation network, we employed crowd-sensing data from S-DoT, an IoT-based system designed to collect urban microclimate data. The S-DoT network comprises 1100 sensors distributed across Seoul at heights ranging from 2 to 4 m above the ground (Seoul Metropolitan Government 2024). These sensors are positioned close enough to the pedestrian level while reflecting the surrounding built/natural environment, but their records are not interfered with by passersby due to the installation height (Kim, Jang, and Kim 2023). Therefore, this extensive network facilitates the collection of a large volume of spatially representative data, enabling an accurate depiction of temperature variations throughout the city. To analyze urban temperature patterns, we calculated the average daily summer temperature over three years. To ensure data consistency across different temporal scales (daily, monthly, yearly, and three-year summer periods), outliers were identified and removed based on pre-defined criteria (Figure 2). After filtering, data from 832 sensors were retained for further analysis (Figure 1), providing a robust and spatially consistent dataset for assessing urban microclimate variations.

2.3. Research process and statistical method

The study, as shown in Figure 3, follows a two-stage process: (1) identifying urban physical factors that influence summer temperatures, and (2) developing a spatial prediction model for summer temperature distribution. The first stage employs LR to identify key urban factors contributing to temperature variations, which can inform strategies for adapting urban environments to extreme heat. The second stage uses RK to create a high-resolution summer temperature map for Seoul by combining statistical modeling with spatial interpolation. This approach provides both

analytical insights into temperature influences and a practical tool for urban climate management.

RK integrates linear regression and Kriging methods. The temperature at a specific location is influenced not only by its physical characteristics but also by surrounding temperatures. LR identifies the key urban structural factors affecting temperature, while Kriging models the spatial correlations between neighboring locations. The RK equation is:

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0) = \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0) + \sum_{i=1}^n \lambda_i \cdot (s_i) \quad (1)$$

where $\hat{z}(s_0)$ is the interpolated value at the location s_0 ; (s_0) is the fitted drift; $\hat{e}(s_0)$ is the interpolated residual; $\hat{\beta}_k$ is the k -th estimated regression coefficient; q_k is the k -th external auxiliary variable at the location s_0 ; p is the number of auxiliary variables; λ_i is the weight of the i -th residual; and $e(s_i)$ is the regression residual.

The LR model is developed using best subset selection (BSS) via the “leaps” package in R. This method evaluates all possible subsets of predictors to identify the best combination (James et al. 2021). Model selection criteria include the adjusted multiple correlation coefficient (R_a^2), Mallows’ C_p , and the Bayesian information criterion (BIC). The optimal model is chosen based on the highest R_a^2 and the lowest C_p and BIC values. The corresponding equations are:

$$R_a^2 = 1 - \frac{SSE_p / (n - p - 1)}{SST / (n - 1)} \quad (2)$$

$$C_p = \frac{SSE_p}{\hat{\sigma}^2} + 2(p + 1) - n \quad (3)$$

$$BIC = n \ln(SSE_p / n) + (P + 1) \ln n \quad (4)$$

where n is the number of data points, p is the number of variables used, $\hat{\sigma}^2$ is the estimated variance of the error, SSE_p is the error sum of squares for p variables, and SST is the total sum of squares.

In ArcGIS Pro 3.2.0, spherical ordinary Kriging (OK) is applied to model the residuals from the BSS-based LR results. Kriging is a geostatistical interpolation method that provides unbiased estimates by assessing spatial correlations through semivariograms (Krige 1966). A semivariogram describes spatial variation using three key components:

- **Nugget:** Represents small-scale variation or measurement error.
- **Range:** The distance at which spatial correlation ceases.
- **Sill:** The maximum semivariance value, beyond which points are uncorrelated.

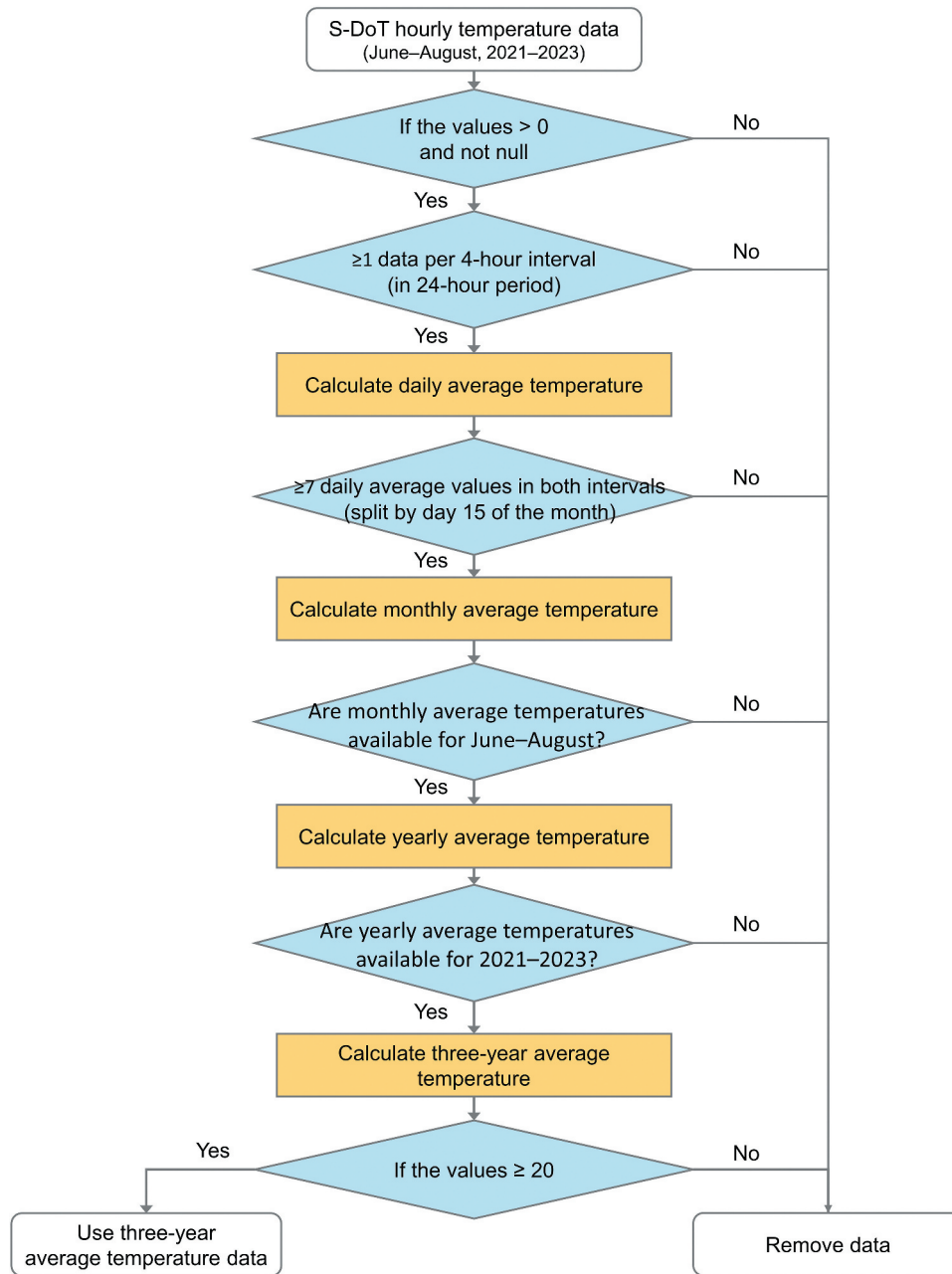


Figure 2. S-DoT temperature data processing flowchart.

The spherical model is used when short-distance variations are linear, and the nugget, range, and sill effects are prominent (Lee 2022). Ordinary Kriging assumes a constant mean and normal distribution of data, capturing spatial autocorrelation (ESRI (2021); SuperMap 2025). The spherical semivariogram model is commonly employed in ordinary Kriging to describe how spatial dependence increases with distance until it reaches a threshold, after which it remains constant (Pyrzcz and Deutsch 2006).

2.4. Variables

This study classifies variables into four categories: population, LULC, buildings, and geographical/topographical factors (Table 1). All variables are processed

as spatial data in ArcGIS Pro 3.2.0, using a 100 m × 100 m grid resolution (0.01 km²).

Population data serve as an indicator of heat sources, reflecting human activity that contributes to rising temperatures (Han et al. 2022a; Ma et al. 2024; Park and Kim 2023; Peng et al. 2018). The average population data from 2021 to 2023, obtained from the Statistical Geography Information Service (SGIS) (Ministry of Data and Statistics 2024), was used in the analysis.

LULC data, as shown in Figure 4(a), is sourced from the Environmental Geographic Information Service (EGIS) (Ministry of Environment 2024). Previous studies have broadly categorized LULC into impervious surfaces, vegetated areas, and water bodies (Berger et al. 2017; Estoque,

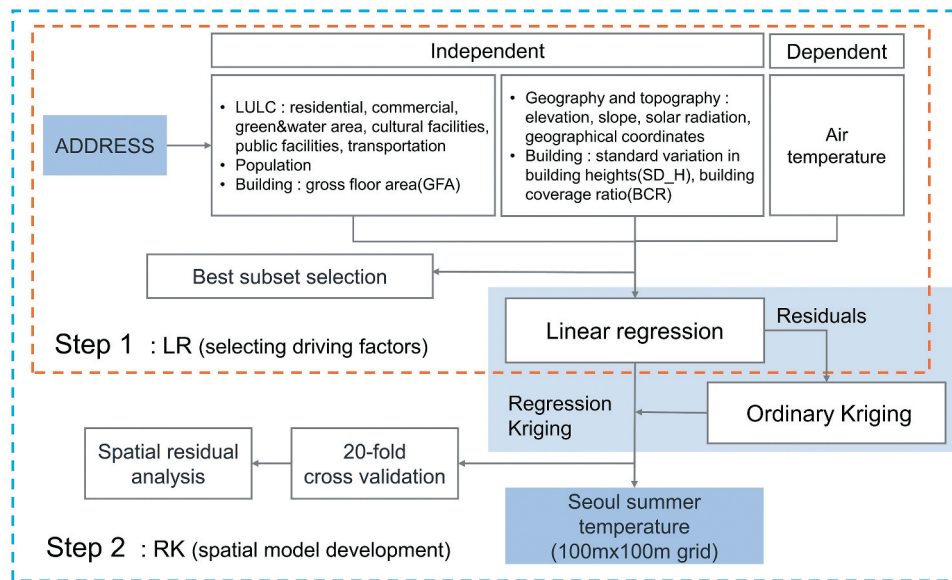


Figure 3. Two-step linear regression and regression Kriging framework for identifying driving factors and modeling summer air temperature in Seoul.

Table 1. List of variables used in the study, including their expected direction of effect, units, whether they were identified using address, and data source access.

Category	Variable	Expected direction of effect	Unit	Using ADDRESS	Access
Average summer temperature		–	°C	No	Seoul Open Data Plaza
Population		Positive	No.	Yes	SGIS
LULC	Green/Water	Negative	m ²	Yes	EGIS
	Agricultural	Negative	m ²	Yes	
	Transportation	Positive	m ²	Yes	
	Residential	Positive	m ²	Yes	
	Commercial	Positive	m ²	Yes	
	Industrial	Positive	m ²	Yes	
	Cultural	Positive	m ²	Yes	
	Public	Positive	m ²	Yes	
Building	GFA	Positive	m ²	Yes	V-WORLD
	BCR	Positive	%	No	
	SD_H	Positive, Negative	m	No	
Geography/Topography	Altitude	Negative	m	No	NGII
	Slope	Negative	(°)	No	Generated variable
	Longitude	Negative	(°)	No	
	Latitude	Negative	(°)	No	
	Solar radiation	Positive	kWh/m ²	No	

Murayama, and Myint 2017; Ma et al. 2024; Peng et al. 2018). Some studies further divide impervious surfaces into residential, commercial, and industrial sectors (Assaf and Assaad 2024; Huang and Wang 2019; Park and Kim 2023; Shi et al. 2019). This study assumes that more detailed distinctions within LULC categories may influence air temperature. Therefore, we classify built areas into transportation, residential, commercial, industrial, cultural, and public spaces, while non-built areas include green spaces, water bodies, and agricultural land – areas known for their cooling effects (Kong et al. 2014; Song et al. 2014). Green and water areas, which contribute to cooling through evaporation and transpiration (Liou, Tran, and Nguyen 2024), are grouped as a single variable, consistent with the concept of blue-green infrastructure (Douglas 2018).

Building-related variables are key to temperature regulation due to their three-dimensional structure. Using V-World data (Ministry of Land, Infrastructure and Transport 2024), we derive three key metrics: gross floor area (GFA), building coverage ratio (BCR), and the standard deviation of height (SD_H). GFA is expected to correlate positively with temperature due to high heat storage, radiation, and artificial heat emissions. SD_H measures height variation within the urban canopy (Figure 4(b)) and has two possible effects: it may trap heat by obstructing airflow or reduce temperatures by increasing shading (Cheng et al. 2006; Ng 2009; Park, Ha, and Lee 2017). BCR, which represents horizontal density, influences airflow and heat dissipation (Figure 4(c)) (Chen et al. 2022; Park and Kim 2023).

The final category includes geographical and topographical factors. Latitude and longitude are derived

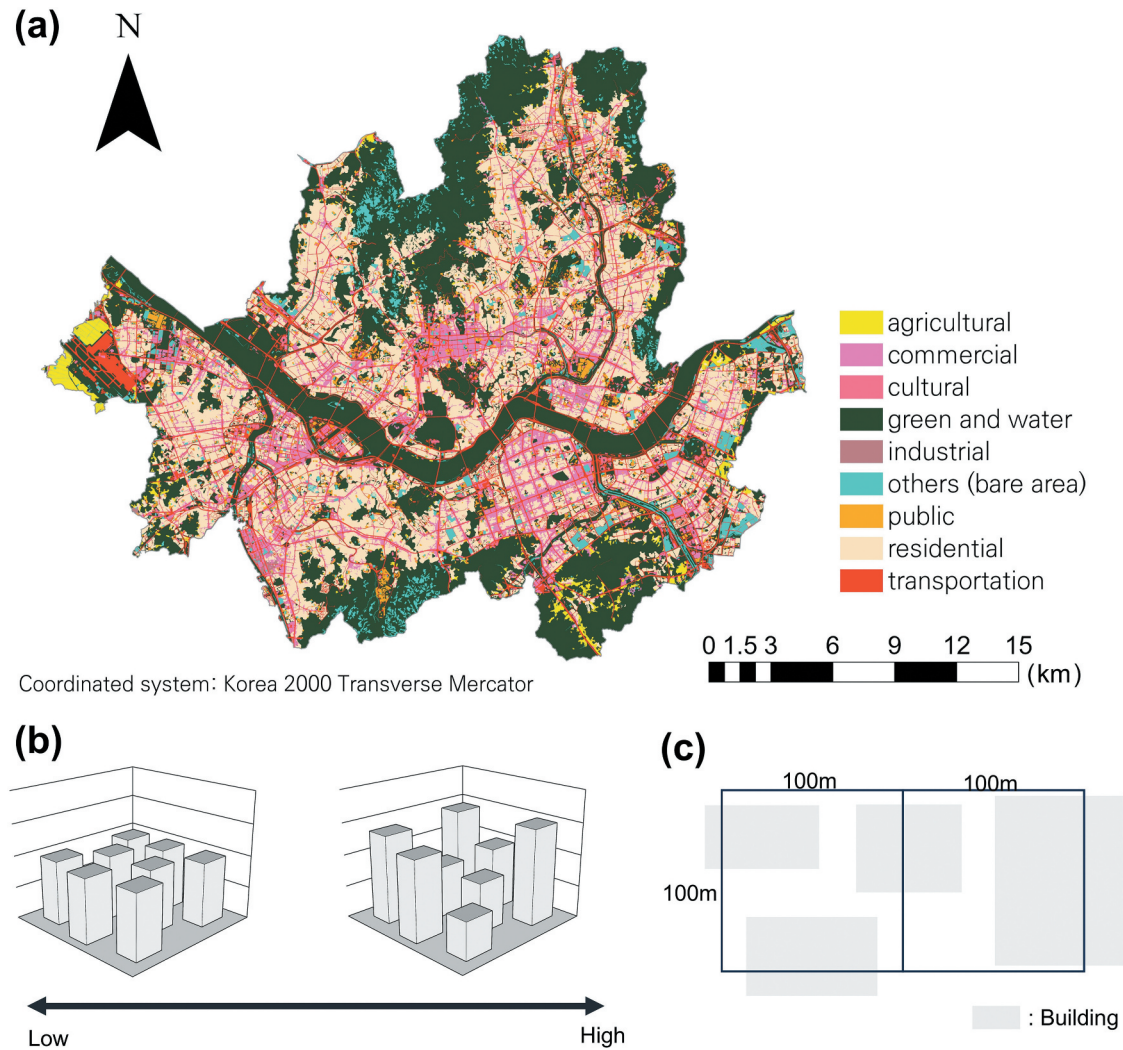


Figure 4. Visual representation of variables: (a) LULC classification used in the study, (b) the standard deviation of height, and (c) building coverage ratio.

using the ArcGIS Pro coordinate system. Topographical variables include altitude, slope, and solar radiation (Figure 5). Higher altitudes generally result in lower temperatures due to reduced radiative heat and air density, as described by the lapse rate effect (Kattel et al. 2022). Solar radiation varies with slope orientation (Rich, Hetrick, and Saving 1995), so slope is included to assess terrain effects. Altitude data comes from a 90-m resolution digital elevation model (DEM) provided by the National Geographic Information Institute (NGII) (Ministry of Land, Infrastructure and Transport 2023) (Figure 5(a)). Using ArcGIS Pro's spatial analyst tools, we generate slope and solar radiation maps based on the DEM (Figures 5(b,c)).

2.5. A distance decay regression selection strategy

In this study, we apply the ADDRESS method to optimize buffer distances. While temperatures at a given location are influenced by surrounding physical

conditions, the spatial extent of this influence varies depending on atmospheric mixing, radiative exchange, surface roughness, and the dispersal characteristics of the surrounding land cover and built environment. Because these processes differ across variables, for example, vegetation exerts cooling effects over shorter distances, while building density or topography may influence airflow and heat accumulation over broader areas, using fixed or arbitrary buffer ranges can result in suboptimal modeling performance. ADDRESS overcomes this limitation by systematically identifying the buffer distance at which each variable explains the largest share of temperature variance, thereby improving the physical relevance and precision of the model (Su, Jerrett, and Beckerman 2009). Although originally developed for air pollution analysis, ADDRESS has been increasingly applied in studies of the urban thermal environment to reflect variable-specific influence ranges. For example, Shi et al. (2019) applied ADDRESS to model thermal conditions, and M. Kim, Won, and Kim (2025) used it to identify key environmental determinants of urban air temperature.

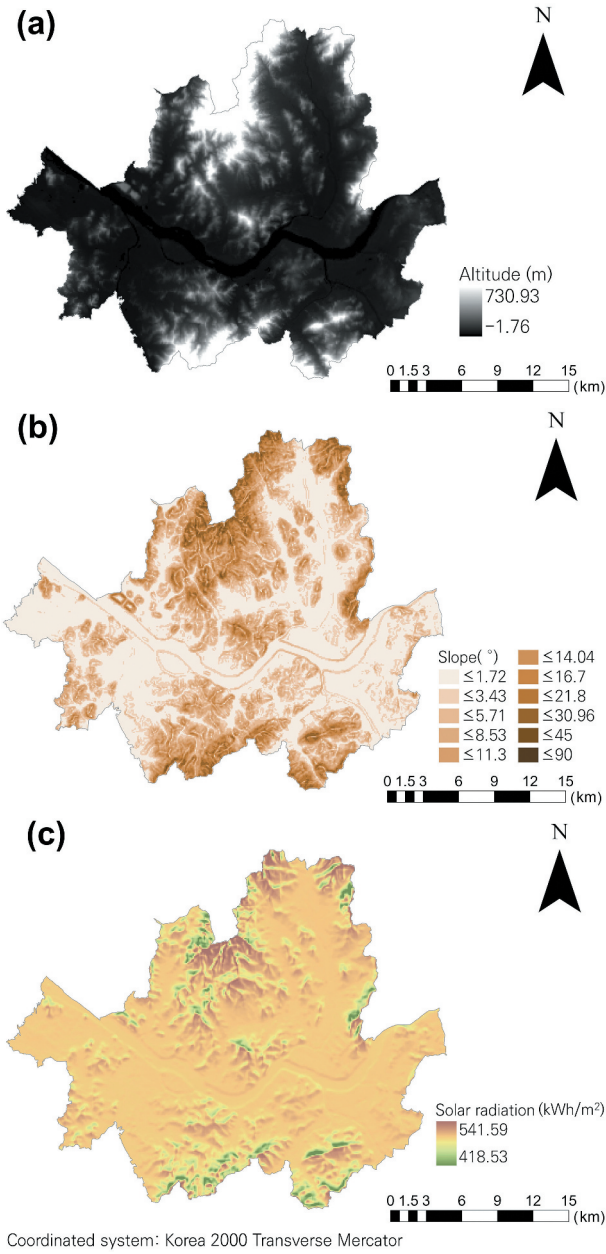


Figure 5. Visual representation of topographical variables: (a) altitude, (b) slope, and (c) solar radiation.

To implement this, we use the Spatial Analyst tool's "focal statistics" in ArcGIS Pro 3.2.0 to calculate variables at buffer distances ranging from 100 m to 5 km, in 100-m increments. Pearson correlation coefficients between each variable and temperature are then calculated for each buffer distance using R, and we select the buffer distance that yields the highest correlation. However, instead of simply selecting the buffer distance with the highest correlation, we examined the full distance – decay curves to assess whether the candidate critical buffers were appropriate. When the curves exhibited multiple inflection points, we chose the first point to avoid unnecessarily large buffers that would increase multicollinearity among predictors. This approach allows us to optimize variables through ADDRESS, providing a more accurate representation of the spatial influences on summer temperatures.

2.6. Cross-validation for model performance

To evaluate the model's performance beyond the observed data, 20-fold cross-validation (CV) was applied. K -fold CV is a method used to assess a model's ability to predict values at unmeasured locations (Fushiki 2011). The dataset is randomly divided into K equal-sized folds, where one fold is used to compute the prediction error, and the remaining $K - 1$ folds serve as the training set (Hastie, Tibshirani, and Friedman 2009; James et al. 2021). This process is repeated K times, and the average prediction error across all folds, denoted as $CV_{(k)}$, is used to evaluate model accuracy. Model performance is measured using two metrics: root mean squared error (RMSE) and mean absolute error (MAE), which are calculated as follows:

$$RMSE_k = \sqrt{\sum_{i \in C_k} (y_i - \hat{y}_i)^2 / n_k} \quad (5)$$

$$MAE_k = \sum_{i \in C_k} |y_i - \hat{y}_i| / n_k \quad (6)$$

$$CV_{(k)} = \sum_{k=1}^K \frac{n_k}{n} \times (\text{Prediction error of each fold}) \quad (7)$$

where n is the total number of data points; n_k is the number of data points in the k -th fold; C_k is the set of data included in the k -th fold; y_i and $\hat{y}_i$ are the actual and estimated values of the i -th data point, respectively; $RMSE_k$ and MAE_k represent the RMSE and MAE for the k -th fold, respectively.

To examine spatial variations in model performance, both global and local Moran's I coefficients are computed using the "spider" package in R. The global Moran's I detects spatial autocorrelation, while the local indicator of spatial association (LISA) reveals local variations (Xiang et al. 2024; Zhang, Gove, and Heath 2005). After assessing residual autocorrelation using global Moran's I, we evaluate whether local spatial patterns deviate from the global trend. The global Moran coefficient is given by:

$$MC = \frac{n \cdot \sum_{i=1}^n \sum_{j=1}^n w_{ij} (e_i - \bar{e})(e_j - \bar{e})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \cdot \sum_i (e_i - \bar{e})^2} \quad (8)$$

The local Moran coefficient is:

$$MC_i = (e_i - \bar{e}) \sum_{j=1}^n w_{ij} (e_j - \bar{e}), \quad (9)$$

where e_i and e_j represent the model errors at locations i and j , respectively; $\bar{e}$ is the mean error across all n locations; and w_{ij} is the spatial weight measure, calculated based on the k -th nearest neighbors (with $k=4$ in this study). If location j is a neighbor of location i , then $w_{ij} = 1$; otherwise $w_{ij} = 0$.

The global Moran coefficient ranges from -1 to 1 , where positive values indicate a positive spatial autocorrelation, meaning that neighboring areas tend to have similar values, while negative values suggest a negative spatial autocorrelation, where neighboring areas are more dissimilar. Values near zero indicate no spatial autocorrelation, whereas values approaching -1 or 1 indicate strong spatial autocorrelation. The local Moran coefficient allows for the identification of whether a given point is similar to or different from its neighbors. A positive value indicates that the point is surrounded by similar values, while a negative value indicates that it is surrounded by opposing values.

3. Results and discussion

3.1. Descriptive statistics and determination of spatial influence range

As shown in Table 2, the average summer temperature over the three years was 26.93°C , with a temperature range spanning 5.32°C between the minimum and maximum values. A temperature variation of 5.32°C within a city during the summer reflects the combined effects of urban heat island (UHI) phenomena, microclimates, and varying land use practices. For instance, a 4°C – 5°C difference compared to rural areas indicates a strong UHI intensity, which can significantly affect public health, energy consumption, and even local ecosystems (Assaf and Assaad 2024; Assaf, Hu, and Assaad 2023; Stillman 2019).

Additionally, we determined the spatial influence ranges of explanatory variables using the ADDRESS method. As shown in Figure 6, the distance – correlation curves identify the buffer ranges most strongly associated with temperature. For variables such as population, residential areas, and cultural facilities (Figures 6(a,e, h)), the curves contain multiple inflection points. To ensure reproducibility and avoid arbitrary selection, we selected the first stable local maximum that did not introduce multicollinearity

when incorporated into the regression model. Broader distances near later inflection points produced substantial multicollinearity; therefore, 500 m, 100 m, and 2500 m were selected for these variables. For variables with a single peak or monotonic behavior (Figures 6(b, c, d, f, g, i, j)), we selected the distance corresponding to the maximum correlation, consistent with the ADDRESS principle that variables exert their strongest climatological influence at this point. In total, 17 predictors and their optimal buffer radii were selected (Table 2).

3.2. Regression Kriging

3.2.1. Best subset selection regression analysis

As shown in Table 3, we conducted a BSS regression analysis and identified the optimal models based on adjusted R^2 (Figure 7(a)), C_p (Figure 7(b)), and BIC (Figure 7(c)). The best performing models were selected with 13, 11, and 9 variables, respectively. The model derived from the variables selected based on adjusted R^2 is designated as Model 1 (with 13 variables), the one based on C_p is Model 2 (with 11 variables), and the one based on BIC is Model 3 (with 9 variables).

To determine the best model, we evaluated the reliability and multicollinearity of the variables. All three models exhibited similar adjusted R^2 values, and multicollinearity was negligible, as indicated by variance inflation factors (VIFs) below 3. However, Model 1 included statistically insignificant variables, and its intercept was also non-significant. In contrast, Models 2 and 3 contained only statistically significant variables at the 0.1 confidence level. Although a 5% significance level is commonly used, Model 2 included a more diverse set of variables that better capture the 3D urban environment. Specifically, it included SD_H and solar radiation, which are linked to 3D urban characteristics. Despite the comparable performance of the models and the statistical significance of each variable, the final model was selected

Table 2. Descriptive statistics of the variables used in the study.

Category	Variable (Radii)	Min	Max	Mean	SD
Average summer temperature ($^{\circ}\text{C}$)		23.51	28.83	26.93	0.70
Population (500 m) (No.)		0	39,106	12,062.31	9776.61
LULC	Green/Water (400 m) (m^2)	0	490,000	171,861.58	156,745.74
	Agricultural (500 m) (m^2)	0	765,475	10,417.06	48,969.45
	Transportation (5 km) (m^2)	620,475	7,483,900	4,597,068.58	1,547,125.33
	Residential (100 m) (m^2)	0	50,000	11,530.41	14,051.64
	Commercial (400 m) (m^2)	0	356,975	44,632.97	57,870.89
	Industrial (1.1 km) (m^2)	0	324,100	4,793.98	22,148.62
	Cultural (2.5 km) (m^2)	0	308,325	82,930.82	55,766.27
	Public (4.7 km) (m^2)	119,675	2,497,975	1,204,796.71	431,238.66
Building	GFA (5 km) (m^2)	10,117,113	117,491,760	70,030,547.19	24,365,332.82
	BCR (%)	0	100	16.87	17.49
	SD_H (m)	0	239.71	4.22	7.68
	Altitude (m)	-1.76	730.93	59.72	82.26
Geography/Topography	Slope ($^{\circ}$)	0	41.25	4.99	5.91
	Longitude ($^{\circ}$)	126.77	127.18	126.99	0.09
	Latitude ($^{\circ}$)	37.43	37.70	37.55	0.06
Solar radiation (kWh/m^2)		418.53	541.59	498.27	8.00

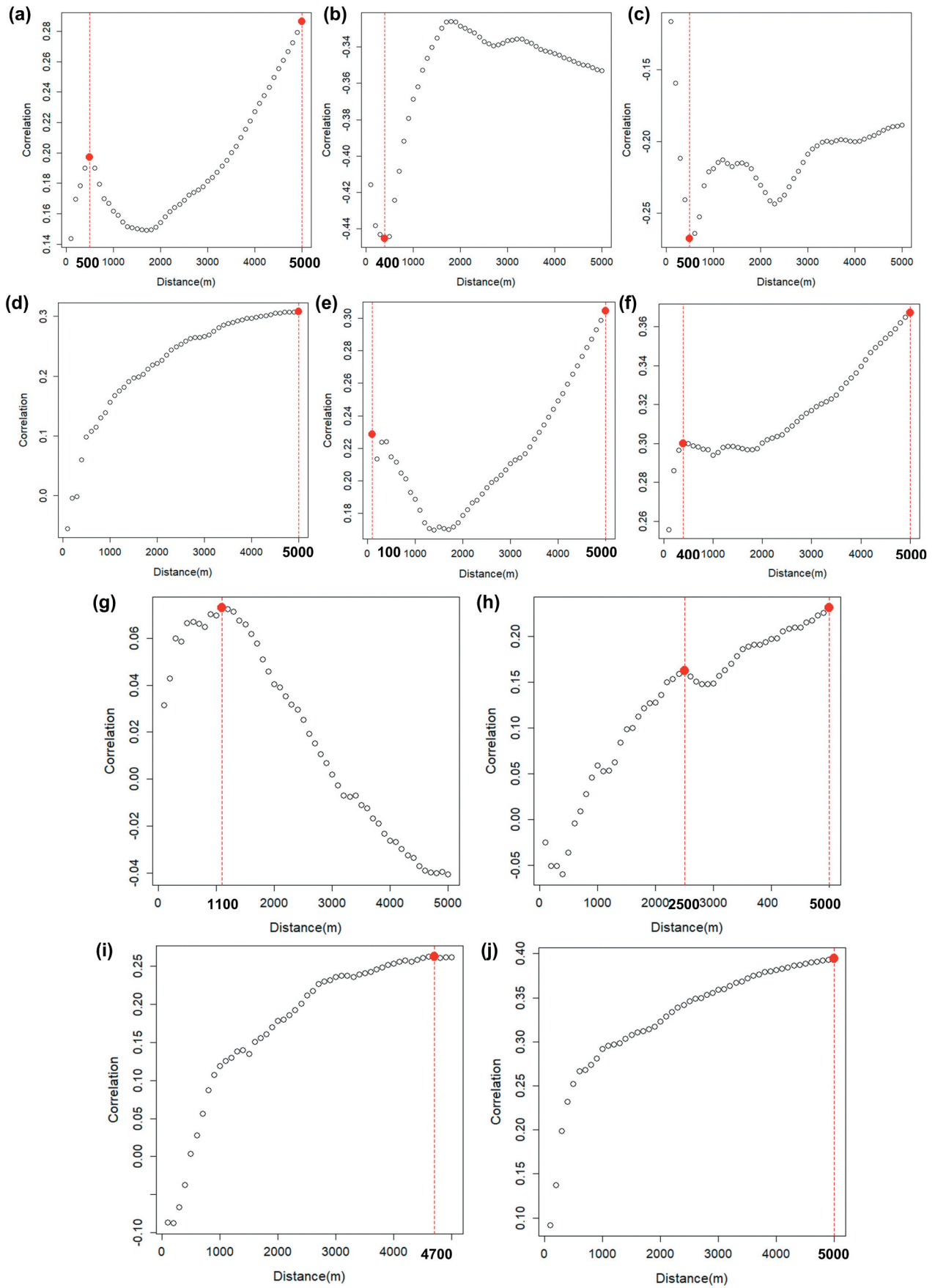


Figure 6. Results of address analysis shown through distance-decay curves of Pearson correlation coefficients. Red points indicate selected distances. (a) Population, (b) green/water, (c) agricultural, (d) transportation, (e) residential, (f) commercial, (g) industrial, (h) cultural, (i) public area, and (j) GFA.

Table 3. Regression model results using best subset selection.

Variables	Unit	Model 1 ($p = 13$)		Model 2 ($p = 11$)		Model 3 ($p = 9$)	
		Coefficient	VIF	Coefficient	VIF	Coefficient	VIF
Intercept	°C	37.260	—	71.370***	—	70.780 ***	—
Green/Water (400 m)	m ²	-1.797×10^{-6} ***	1.781	-1.829×10^{-6} ***	1.774	-1.802×10^{-6} ***	1.692
Agricultural (500 m)	m ²	-1.014×10^{-5} ***	1.174	-9.739×10^{-6} ***	1.158	-9.609×10^{-6} ***	1.152
Residential (100 m)	m ²	5.745×10^{-6} ***	1.301	5.508×10^{-6} ***	1.279	5.571×10^{-6} ***	1.279
Industrial (1.1 km)	m ²	8.254×10^{-7}	1.084	—	—	—	—
Cultural (2.5 km)	m ²	7.854×10^{-7} *	1.471	8.711×10^{-7} *	1.307	9.118×10^{-7} **	1.304
Public (4.7 km)	m ²	2.322×10^{-7} ***	1.903	2.418×10^{-7} ***	1.871	2.437×10^{-7} ***	1.860
GFA (5 km)	m ²	4.749×10^{-9} ***	2.723	4.342×10^{-9} **	2.564	4.242×10^{-9} **	2.524
BCR	%	1.057×10^{-2} ***	1.475	1.043×10^{-2} ***	1.460	1.079×10^{-2} ***	1.413
SD_H	m	-4.323×10^{-3} †	1.086	-4.116×10^{-3} †	1.082	—	—
Altitude	m	-8.170×10^{-3} ***	1.551	-8.272×10^{-3} ***	1.544	-8.320×10^{-3} ***	1.538
Longitude	(°)	2.574×10^{-1}	1.295	—	—	—	—
Latitude	(°)	-1.273 **	1.862	-1.312 **	1.833	-1.189 **	1.776
Solar radiation	kWh/m ²	8.023×10^{-3} †	1.074	8.190×10^{-3} †	1.073	—	—
Multiple R^2 (Adjusted R^2)		0.531 (0.523)		0.529 (0.523)		0.525 (0.520)	
C_p		11.426		10.659		13.595	
BIC		-535.494		-545.650		-552.078	

Note: Model 1 includes 13 variables, Model 2 includes 11 variables, and Model 3 includes 9 variables. "****" represents the significance level at 0.001, "***" represents the significance level at 0.01, "**" represents the significance level at 0.05, "†" represents the significance level at 0.1, and "—" represents the level at 1.

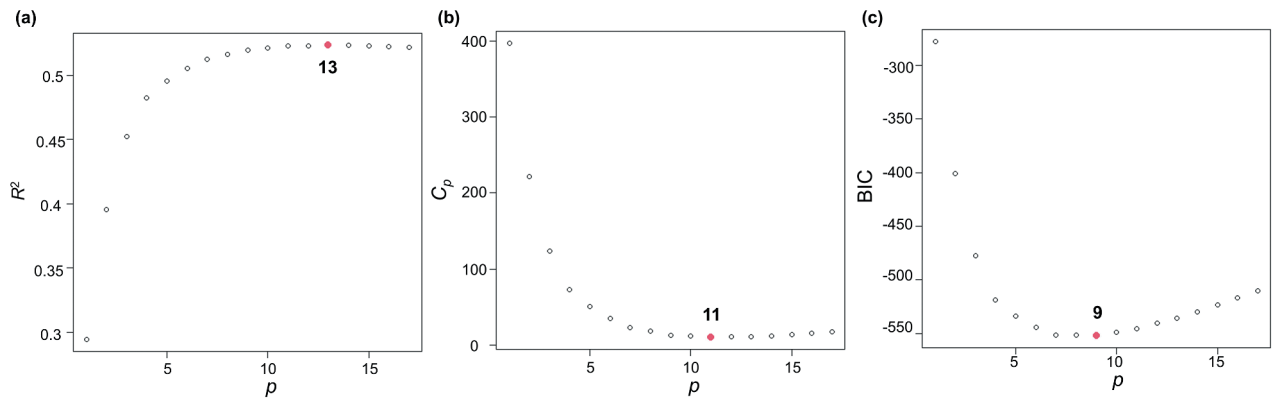


Figure 7. Adjusted R^2 (a), C_p (b), and BIC (c) are shown for the best models with each number of predictors, where p represents the number of predictors. The red point indicates the best model according to each variable selection criterion.

based on its performance with Kriging applied in both Models 2 and 3.

Regarding individual variables, there were no differences in the direction of coefficients across the three models. At the 5% significance level, variables positively correlated with temperature included GFA, BCR, public spaces, residential areas, and cultural facilities. In contrast, green/water areas, altitude, agricultural zones, and latitude were negatively correlated with temperature. At the 10% significance level, SD_H (p -value: 0.065) and solar radiation (p -value: 0.0719) were also included. Built-up areas had varying temperature effects: for instance, a 1000 m² increase in residential area within a 100 m radius raised the temperature by 0.005508°C and 0.005571°C in Models 2 and 3, respectively. Similarly, for cultural areas (at a 2.5 km radius), the temperature increased by 0.0008711°C and 0.0009118°C in Models 2 and 3, respectively, while public areas (4.7 km radius) showed an increase of 0.0002418°C and 0.0002437°C in Models 2 and 3, respectively. Due to the differences in variable radii, coefficient values are not directly

comparable. Cooling effects also varied across land use types. Agricultural areas (500 m radius) reduced temperature by 0.009739°C and 0.009609°C in Models 2 and 3, respectively, whereas green/water areas (400 m radius) decreased temperature by 0.001829°C and 0.001802°C, respectively. The cooling effect of agricultural areas was approximately five times greater than that of green/water areas.

In addition to LULC, an increase of 1 km² in total floor area within a 5 km radius raised the temperature by 0.004342°C and 0.004242°C in Models 2 and 3, respectively. When the building area increased by 10% (1000 m²), the temperature rose by 0.1043°C and 0.1079°C in Models 2 and 3, respectively. Each 100 m increase in altitude resulted in a temperature decrease of 0.08272°C and 0.0832°C in Models 2 and 3, respectively, while a 0.1° change in latitude led to a decrease of 0.1312°C and 0.1189°C in Models 2 and 3, respectively. Industrial areas, which were excluded from Models 2 and 3, showed a positive but insignificant coefficient in Model 1, suggesting a minimal impact. Latitude consistently demonstrated a significant negative

relationship with temperature across all models. While SD_H and solar radiation were significant at the 10% level in Models 1 and 2, they were excluded from Model 3, indicating reduced explanatory power when other variables were included. These findings underscore the complex interactions between spatial, environmental, and built environment factors influencing temperature variations.

3.2.2. Ordinary Kriging

OK was applied to the residuals from the BSS regression analysis. Using a spherical model, the semivariograms of both models revealed strong spatial correlations between the residuals (Figure 8). Compared to the semivariogram parameters for temperature and residuals, the residuals from Model 2 (Figure 8(a)) and Model 3 (Figure 8(b)) exhibited shorter ranges and lower nugget, partial sill, and sill values (Table 4). These patterns align with findings from previous studies (Hengl, Heuvelink, and Rossiter 2007; Hengl, Heuvelink, and Stein 2004). Since regression explains much of the temperature variation due to environmental factors, the shorter range and smaller sill values suggest that RK effectively accounts

Table 4. Comparison of the range, nugget, and sill of the OK semivariograms.

Metric	Air temperature	Model 2	Model 3
Range (m)	16,421	13,970	14,080
Nugget ($^{\circ}\text{C}^2$)	0.2839	0.1848	0.1865
Partial sill ($^{\circ}\text{C}^2$)	0.1817	0.0247	0.0249
Sill ($^{\circ}\text{C}^2$)	0.4656	0.2095	0.2114

Note: Range is the distance at which spatial correlation ceases. Nugget represents small-scale variation or measurement error. Partial sill is the difference between the sill and the nugget, and sill indicates the maximum semivariance value, beyond which points are uncorrelated.

for environmental influences, reducing overall variability. This results in a more rapid decline in spatial dependence and lower semivariance values than temperature Kriging. Additionally, the smaller nugget values indicate that localized environmental differences are well captured by the regression model.

Table 4 presents a comparison of range, nugget, and sill values from the semivariograms used in OK. The results show that air temperature exhibits the largest spatial dependence, with a range of 16,421 m, while Models 2 and 3 show shorter ranges of 13,970 and 14,080 m, respectively. Nugget values for Models 2 and 3 are significantly lower (0.1848 and 0.1865) compared to temperature (0.2839), indicating that regression reduces unexplained spatial variability. Both partial sill and total sill values also decrease, further supporting the effectiveness of regression in explaining environmental influences.

3.2.3. Model performance

To assess model performance, we conducted 20-fold cross-validation, evaluating two key error metrics: MAE and RMSE. MAE reflects the average magnitude of errors, while RMSE penalizes larger deviations more heavily, providing complementary perspectives on model performance. As shown in Table 5, we compared four models: two derived from BSS regression (Models 2 and 3) and two combined models integrating BSS and OK (Models 4 and 5). RMSE results indicate that temperature errors are less than 0.5°C on average, with differences between models under 0.02°C . However, BSS-based models performed better, achieving lower RMSE and MAE, which suggests that they are better at minimizing large prediction errors and providing more consistent performance, respectively.

To determine whether the differences in model performance were statistically significant, we conducted paired t -tests using the absolute value of residuals. A p -value of 0.4837 between Models 2 and 5, and 0.0946 between Models 3 and 5, indicated no statistically significant difference. This suggests that the average error magnitudes of the BSS and combined BSS-OK models are practically equivalent. However, the difference between Models 2 and 4 was

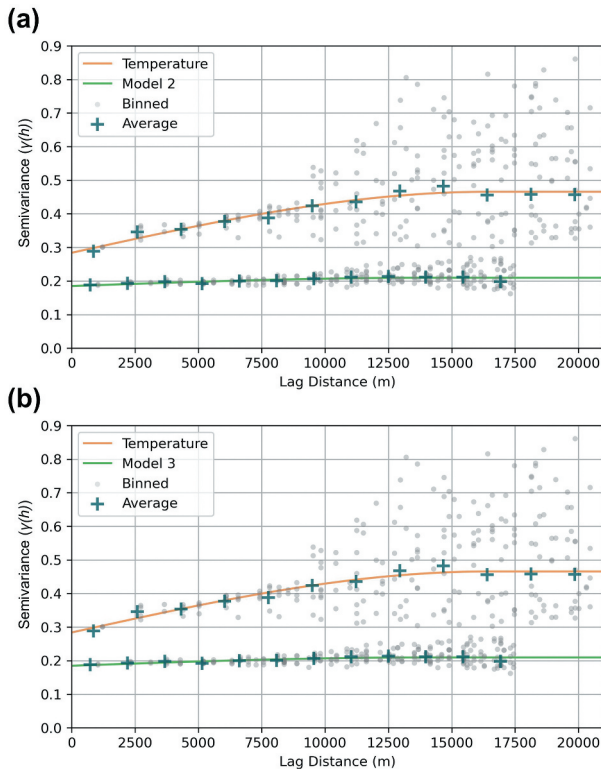


Figure 8. Semivariogram used in OK. Orange lines indicate the semivariogram of temperature, and green lines indicate the semivariogram of each LR model's residuals. Grey dots indicate binned values which were generated by grouping semivariogram points using square cells that are one lag wide. Blue crosses show average points generated by grouping points that fall within angular sectors. (a) Results of temperature and model 2 using 11 variables, and (b) results of temperature and model 3 using 9 variables.

Table 5. Comparison of model performance and the standard deviation of residuals.

Number of variables	Model	RMSE	MAE	Paired t-test			
				Model 4		Model 5	
				Mean difference	<i>p</i> -value	Mean difference	<i>p</i> -value
11	Model 2	0.4590	0.3408	0.0179	0.0049*	0.0120	0.4837
	Model 4	0.4756	0.3586	–	–	0.0176	0.2578
9	Model 3	0.4589	0.3410	0.0176	0.2578	0.0107	0.0946
	Model 5	0.4704	0.3517	0.0069	0.6606	–	–

Note: *Significant at 5% significance level. RMSE and MAE were computed for each model using the full dataset and 20-fold cross-validation. Paired t-tests were conducted to compare RK models. Model 4 combines Model 2 with OK residuals to form the RK model, and Model 5 combines Model 3 with OK residuals.

statistically significant (p -value = 0.0049), indicating that Model 2 outperforms Model 4. These results highlight that the overall performance of BSS-based models is comparable to that of Model 5.

Scatter plots between measured and predicted temperatures were created to assess quantitative bias in error terms (Figure 9), and spatial distributions of error terms were analyzed to identify spatial bias (Figure 10). As shown in Figure 9, all models exhibited a consistent trend: predicted temperatures below 26°C were overestimated, while those above 28°C were underestimated. This pattern may reflect systematic model behavior or limitations in the training data. In terms of spatial

bias, Figure 10(a, c) reveal clustering of errors, where values within a circle are predominantly overestimated or underestimated. In contrast, RK-based models (Figures 10(b, d)) showed spatially unbiased error distributions, suggesting improved spatial generalization.

To assess the statistical significance of these spatial biases, we calculated global and local Moran coefficients (Table 6). The BSS-based models displayed statistically significant global Moran coefficients (0.0539, 0.0569) and Z-values (2.3965, 2.5302) at a p -value of 0.05, indicating that similar error values tend to cluster spatially. On the other hand, the RK-based models had statistically

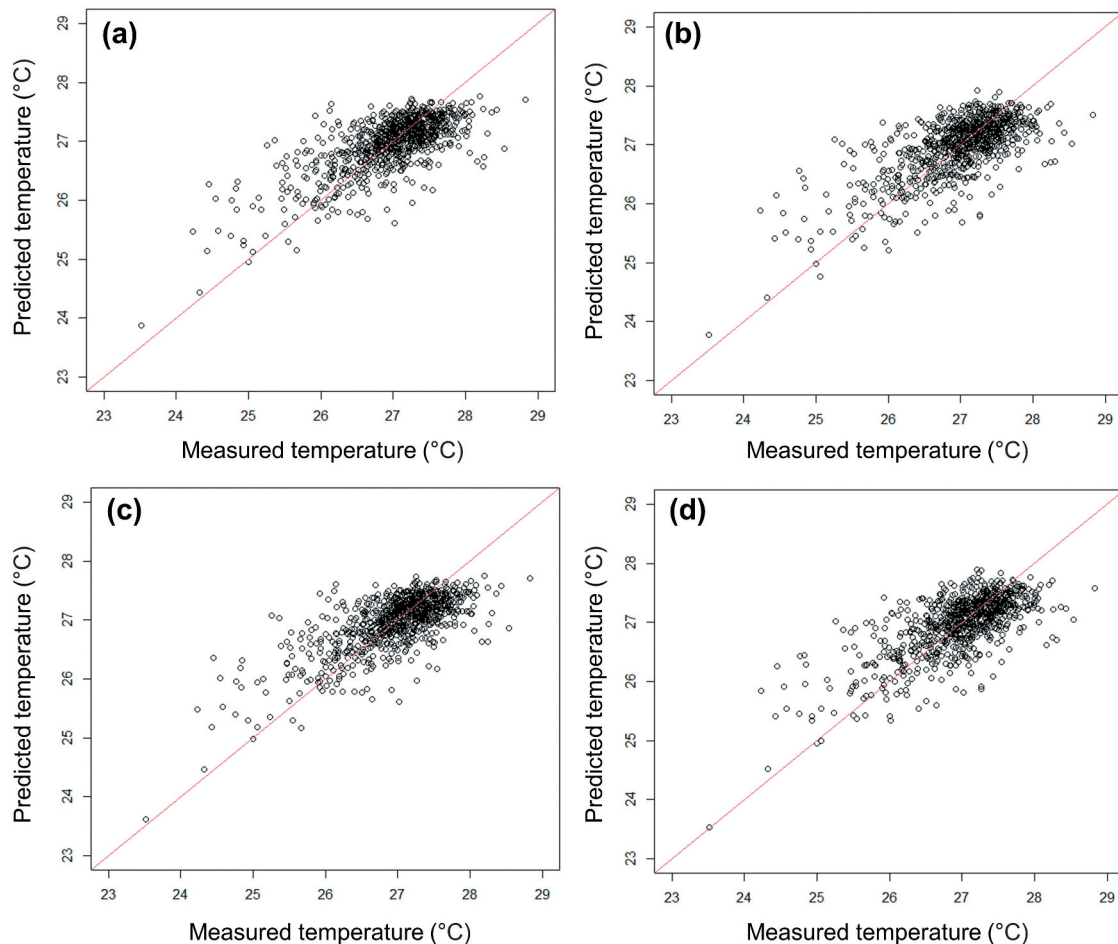


Figure 9. Comparison of predicted and measured temperatures using 20-fold cross-validation. The red line represents the 1:1 line where predicted values equal actual values. (a) and (b) show results using 11 variables for models 2 and 4, respectively, while (c) and (d) show results using 9 variables for models 3 and 5.

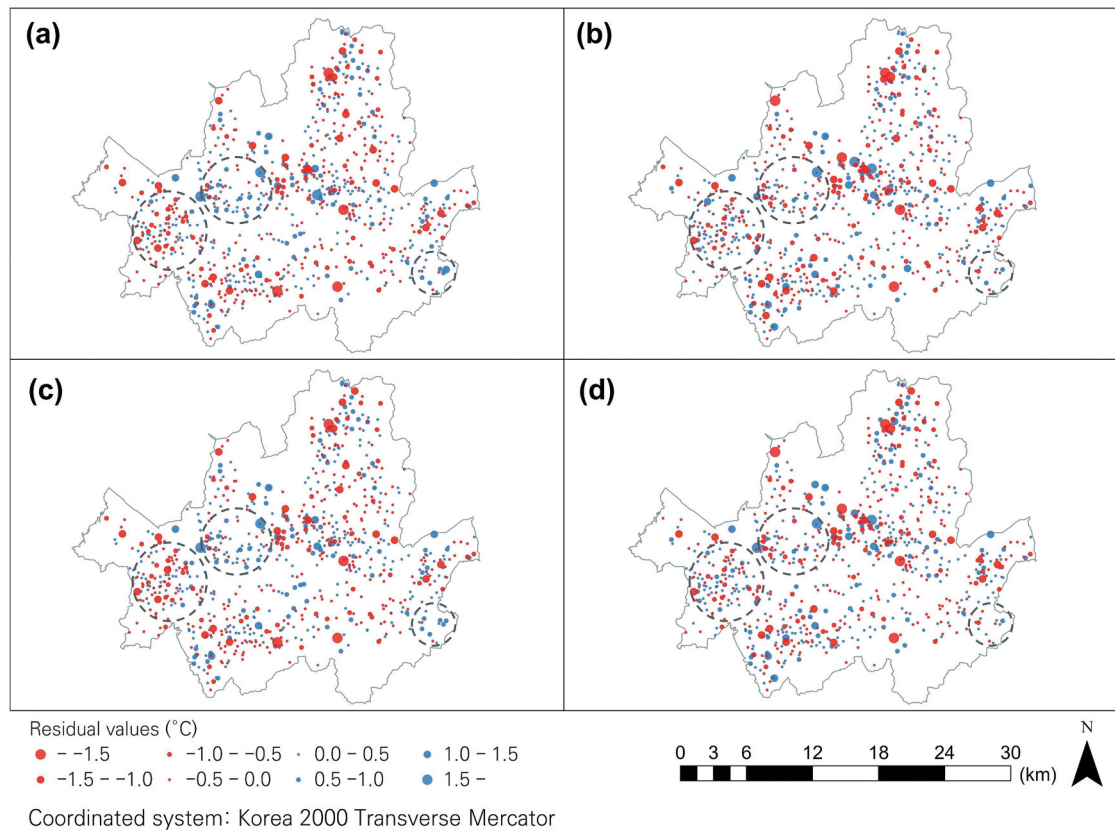


Figure 10. Spatial distribution of error values based on 20-fold cross-validation. (a) Model 2 using LR, (b) Model 4 using RK, (c) Model 3 using LR, and (d) Model 5 using RK. Circles indicate areas with biased estimates, such as overestimations or underestimations, when using LR.

Table 6. Global Moran coefficient and comparison of significant Z-values for the local Moran coefficient.

Model	Global			Local Number of positive spatial autocorrelation (Z-value ≥ 1.96)
	Moran coefficient	Z-value	p-value	
Model 2	0.0539	2.3965	0.00827*	32
Model 3	0.0569	2.5302	0.00569*	31
Model 4	-0.0871	-3.7351	0.999990	7
Model 5	-0.0823	-3.5283	0.999880	7

Note: *Significant at the 5% level. The spatial distribution of residuals was analyzed using the global Moran's I coefficient and the local Moran's I coefficient. The number of data points with significant positive spatial autocorrelation (Z-values ≥ 1.96) was compared across models.

insignificant Moran's I values, failing to reject the null hypothesis of no spatial autocorrelation. At the local level, we assessed the percentage of data points with significant spatial clusters (above 1.96), finding that Models 2 and 3 had more spatially biased points (32, 31) than Models 4 and 5 (7, 7).

These results suggest that while quantitative and spatial biases exist in most models, Model 5 demonstrates superior performance in minimizing spatial bias and providing a more balanced error distribution. Therefore, Model 5 was selected as the final model. Based on this model, we created a summer daily average temperature map for Seoul (Figure 11). The map reveals that inner-city areas tend to have higher temperatures compared to the outskirts. Within the inner city, areas near green and water spaces are predicted to be relatively cooler.

4. Discussion

4.1. Spatial model of temperature using RK

This study examined the performance of the RK model, which integrates regression and Kriging. When using LR alone, residuals exhibited statistically significant spatial autocorrelation, indicating that the model violated the independence assumption of regression residuals and was not fully appropriate for spatial prediction. This spatial variation may partly result from the uneven distribution of S-DoT sensors: densely instrumented areas are well represented in the fitting process, whereas morphological conditions typical of sparsely monitored regions are underrepresented. When strong residual clustering persists, certain areas become systematically over- or underestimated, leading to spatially biased interpretations. Although spatial bias is a critical issue for temperature modeling, previous RK-based studies diagnosed it only visually using error

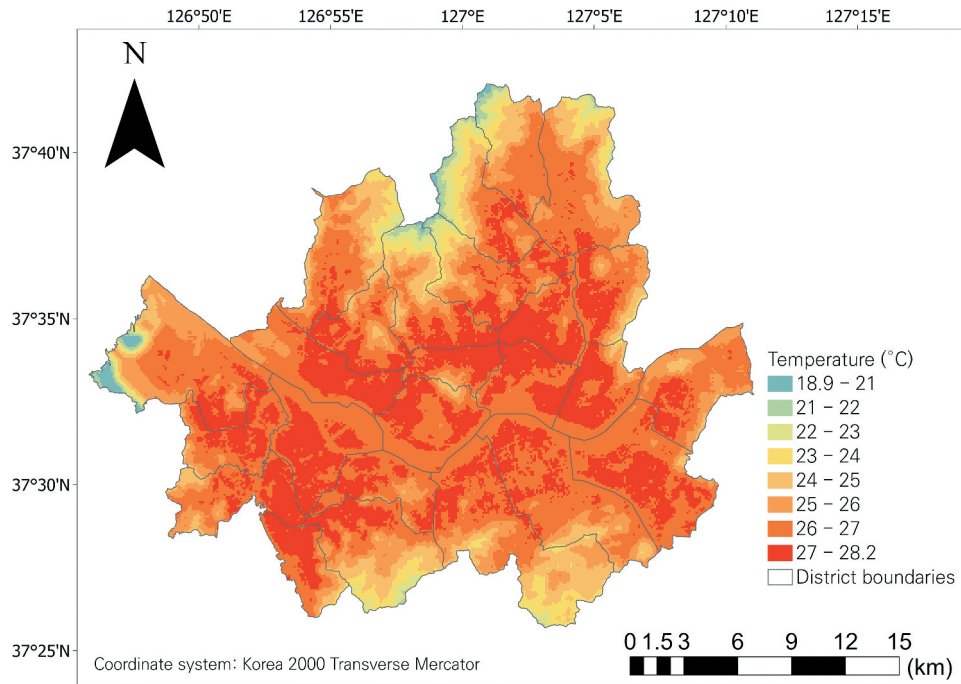


Figure 11. Final spatial model of average summer temperature (2021–2023) utilizing RK with nine variables.

maps (Sekulić et al. 2020; Wu and Li 2013). In contrast, by applying RK, this study partially captures the unexplained spatial dependence in residuals and verifies the reduction in spatial autocorrelation using both global and local Moran's I.

However, despite its effectiveness in mitigating spatial bias, RK did not produce substantial improvements in averaged numerical accuracy relative to the regression model. This contrasts with earlier findings where RK improved both numerical performance and spatial interpolation quality (Li, Griffith, and Shu 2020; Szymanowski, Kryza, and Spallek 2013). Two factors may explain this discrepancy. First, previous studies generally employed a limited number of simple two-dimensional variables such as latitude, longitude, and elevation. By contrast, our model includes detailed 2D and 3D urban physical variables, allowing LR to account for a much larger share of the heat exchange processes across space. When regression already captures most spatial variation, the marginal numerical contribution of RK naturally declines. Second, earlier studies typically relied on sparse and widely spaced meteorological stations covering large geographic areas. Sparse networks produce higher residual variance and larger RMSE values, leaving more spatial structure for RK to explain. For example, Szymanowski, Kryza, and Spallek (2013) reported RMSE values around 1.109°C for multiple linear regression, while ordinary Kriging yielded an RMSE of 2.674°C in Shen, Xu, and Hu (2023). By comparison, the relatively low RMSE values of our BSS model (0.4590°C and 0.4589°C) indicate that a substantial portion of spatial variability was already captured by the explanatory variables. Under these conditions, RK

contributes more to reducing spatial clustering than to improving numerical accuracy. These methodological improvements thus explain the limited incremental accuracy gain, while highlighting RK's continued utility in mitigating spatial bias in residuals. However, because cross-validation reflects the underlying spatial sampling distribution of the S-DoT network, the resulting accuracy metrics may underrepresent prediction uncertainty in sparsely monitored peripheral regions such as river corridors, green spaces, and mountainous areas. Additionally, given the uneven spatial distribution of S-DoT sensors, model reliability is likely higher in densely monitored central districts than in peripheral regions, where sparse observations may lead to greater prediction uncertainty. This spatial sampling imbalance should be considered when interpreting local temperature patterns, particularly near the Han River and mountainous areas.

Regarding temperature patterns, the model reveals that while temperatures in green spaces and along the Han River in the city center are lower than those in surrounding areas, they remain higher than those in suburban regions. The model enhances understanding by capturing the relatively cooler conditions of urban green and water areas, compared to other temperature maps using S-DoT (Cho et al. 2020; Park, Park, and Lee 2023). The LR model alone also reflected this trend, suggesting that an optimal selection of variables enables the model to account for surrounding contextual factors, even in areas lacking S-DoT installations. However, the higher temperatures observed in the city center compared to suburban regions can be attributed to the absence of S-DoT in those areas, making validation of these results more challenging.

4.2. Urban spatial factors influencing summer temperature

The BSS regression analysis identifies the combination of variables that maximizes the model's performance at each step, ensuring that the most frequently selected variables significantly contribute to explaining temperature patterns. The variables chosen at each step suggest that green/water areas were the most consistently selected, highlighting their importance as a key factor influencing temperature, followed by GFA, BCR, altitude, agricultural areas, public spaces, residential zones, and others (Table 7). Most of the selected variables align with our hypotheses.

Green/water areas emerged as one of the most influential variables, showing a strong negative relationship with temperature. This finding aligns with previous studies emphasizing the cooling role of green spaces and water bodies (Chen et al. 2022; Chun and Guldmann 2014; Lee 2024; Yu et al. 2020). However, our model indicates that increasing agricultural land within a 500 m radius produces a stronger cooling effect than expanding green/water areas within a 400 m radius. This suggests that agricultural land – most of which is located on the urban periphery – exerts a more pronounced cooling influence.

Several microclimatic mechanisms help explain this result. Agricultural land generally exhibits lower temperatures due to irrigation-driven evaporative cooling (Kumar et al. 2017) and heightened evapotranspiration during the growing season, particularly in summer (Qiu et al. 2013; Zhou et al. 2016). Croplands also tend to have higher surface albedo relative to forests, reducing heat absorption (Zhou et al. 2016). In contrast, the cooling performance of urban green spaces can be weakened by heat and drought stress, which reduces tree health and evapotranspiration (Gobatti et al. 2025; Wang, Wu, et al. 2025). Furthermore, cooling intensity varies by vegetation type:

broadleaved forests cool more effectively than coniferous stands (Hwang et al. 2023), implying that not all green spaces provide equal thermal benefits.

The spatial context of Seoul further reinforces these patterns. Agricultural land accounts for only 1.749% of the city's area and is concentrated near the outskirts (Figure 4(a)), where low development intensity, adjacency to extensive green zones, and higher ventilation create inherently cooler microclimates (Huang et al. 2025). In such peripheral settings, agricultural land can act as an effective ventilation corridor that enhances airflow and convection (Bowler et al. 2010; Osínska-Skotak and Zawulich 2016). These microclimatic and locational characteristics collectively explain why agricultural land demonstrates a stronger temperature-reducing effect than green/water areas in our model.

Despite the model's reliance on variables with presumed relationships to temperature, some variables were not included in the final model (see Table 7 for discarded variables). In addition, when p -values were calculated based on the full model including all variables, none of the variables except for SD_H were found to be statistically significant at the 5% level. The residential variable was retained because it likely captures the effects of population density, making other variables, such as population, redundant. Among the LULC variables, commercial, industrial, and transportation areas were excluded. Commercial areas, characterized by high density and diverse activities, were considered potential variables but likely excluded because the high floor area ratio standards for commercial zones are effectively represented by the GFA variable. Industrial areas, while initially expected to have a distinct impact on temperature, were excluded because of their sparse distribution within Seoul. Similarly, transportation zones, which expand with urban and population density, are likely accounted for by the GFA and residential variables,

Table 7. Comparison of selected and discarded variables based on BSS results.

Variables		Selected frequency	Initial selection step		p -value
Selected	Green/Water	17	1		<0.001
	GFA	16	2		0.007
Discarded	BCR	15	3		<0.001
	Altitude	14	4		<0.001
	Agricultural	13	5		<0.001
	Public	12	6		<0.001
	Residential	11	7		<0.001
	Latitude	10	8		0.002
	Cultural	9	9		0.026
	SD_H	8	10		0.048
	Solar radiation	7	11		0.117
	Industrial	6	12		0.150
	Longitude	5	13		0.268
	Slope	4	14		0.315
	Commercial	3	15		0.643
	Population	2	16		0.800
	Transportation	1	17		0.823

Note: Variables are ordered by the frequency with which they were selected across models. For comparability, p -values were calculated based on the full model that includes all 17 variables.

which capture the temperature increase linked to transportation infrastructure.

In the geography and topography category, longitude showed no significant relationship with temperature at the urban level. The exclusion of both solar radiation and slope variables suggests that latitude and altitude alone adequately capture variations in solar radiation related to topographic features. Within the building category, SD_H was excluded from the final model, indicating that its conflicting effects, those that both increase and decrease temperature, reduced its coefficient significance and diminished the explanatory power of the variable. Although the *p*-value of SD_H is close to the 5% significance level, it remains relatively higher compared to other selected variables and shows instability across models (Tables 3 and 7). Therefore, this study reflects airflow solely based on BCR, which only partially captures morphology-driven ventilation. Recent studies indicate that urban ventilation and local winds are effective levers for lowering urban temperatures, mitigating UHIs, and improving outdoor thermal comfort, and that local wind fields are modified by urban morphology (He, Ding, and Prasad 2020a, 2020b). Although directly incorporating airflow would likely yield a more accurate model, our study differs from prior studies that focused on precinct scale, short-term work in that we predict city scale, three-year mean summer air temperature. Consequently, spatially continuous airflow observations that reflect regional variation are not available. Therefore, we represent airflow indirectly through morphological attributes.

The results from the ADDRESS model reveal differences in the spatial influence range and magnitude depending on land use types and building density in urban areas. To account for neighborhood interactions, the optimal radius for each variable was determined based on hypothesized temperature relationships. As a result, selected variables were classified into two categories: those with a shorter range (within 1 km), such as green/water, agricultural, and residential areas, and those with longer ranges. When multiple inflection points arose, we selected the first peak and checked VIF. While the ADDRESS by Su, Jerrett, and Beckerman (2009) did not discuss this aspect, our empirical analysis showed that incorporating variables over a wider range around these points led to multicollinearity in the regression models. However, due to the lack of theoretical frameworks for determining the precise range of influence for each variable, this study provides influence levels solely based on statistical methods.

The present study provides practical implications for urban temperature management, particularly for high-density metropolitan contexts such as Seoul. Fine-scale spatial mapping of air temperature offers a foundation for assessing heat vulnerability by

integrating temperature patterns with the spatial distribution of elderly residents, low-income households, outdoor workers, and other at-risk groups. This allows planners to identify priority neighborhoods and to incorporate heat mitigation into ongoing neighborhood-scale planning processes. The ADDRESS results show that green and water areas exert their strongest cooling influence within approximately 400 m – a distance comparable to Seoul's walkable neighborhood unit commonly used in park accessibility and living-environment planning. This suggests that Seoul should expand distributed, small- and medium-scale green-blue spaces within everyday walking distances, rather than relying solely on large, centralized parks. Such neighborhood-level interventions align with global best practices and are particularly relevant for compact Asian megacities. Our findings also highlight the importance of multiscale density management. BCR, representing horizontal density at the block scale, and GFA, capturing vertical density over broader urban areas, were both positively associated with higher temperatures. This indicates that temperature mitigation should be incorporated into Seoul's existing planning instruments: block-scale regulations (e.g. site coverage controls, ventilation corridors) and district- or city-scale development strategies (e.g. coordinated high-rise clustering). Integrating these temperature-sensitive design considerations into density planning can help improve thermal comfort and resilience in high-density urban environments.

5. Conclusions

This study developed a high-resolution (100 m × 100 m) near-surface summer air temperature map for Seoul, utilizing ADDRESS and RK. It examined key factors influencing temperature and the methodological implications of RK in urban climate modeling. The findings offer valuable insights for offering practical guidance for urban planners to mitigate heat, assess heat vulnerability, and visualize urban temperature distributions to inform adaptive urban planning. The key conclusions are as follows:

- (1) Importance of built-environment variables: The careful selection of built-environment variables significantly enhances temperature prediction accuracy, reducing the need for complex interpolation methods like RK.
- (2) Effectiveness of RK: The results demonstrate that RK is applicable not only at macro-scale settings with sparse observations but also at the city level with high-density monitoring networks. Although RK did not substantially enhance numerical accuracy relative to

regression models, it consistently reduced spatial bias in residuals, confirming its value in improving spatial realism and mitigating clustered errors.

- (3) Role of S-DoT data: The dense S-DoT sensor network provides high-resolution, real-time air temperature measurements across Seoul, substantially reducing spatial uncertainty. This rich empirical dataset strengthens the regression component of RK, enabling accurate urban heat distribution mapping and explaining why the incremental numerical benefits of Kriging are smaller than in previous sparse-sensor studies.
- (4) Spatial influence ranges: ADDRESS confirmed that different variables impact temperature over distinct spatial extents, identifying both broader and narrower influence ranges.

Despite these contributions, the study has several limitations:

- (1) Model accuracy constraints: Although RK mitigated spatial bias, it did not significantly improve overall model accuracy. Incorporating additional variables, such as shaded area representation and factors that capture low- and high-temperature extremes, could enhance predictions by addressing unexplained variance and the conflicting effects of SD_H. While we attempted to simulate shading variables for buildings, computational limitations hindered large-scale implementation.
- (2) Incomplete airflow representation: This analysis relies on city-scale spatial predictions using three-year mean air temperature. Because temporal averaging smooths short-term atmospheric variability, wind effects are largely averaged out, and the spatial variability of airflow cannot be directly measured. In our model, airflow is only indirectly represented through the building coverage ratio (BCR), which partially reflects surface roughness and potential ventilation pathways. However, this proxy does not fully capture dynamic airflow, particularly near water bodies such as the Han River, where stronger and more variable winds frequently occur. Since wind and urban ventilation structures play an important role in heat dispersion and thermal regulation, incorporating detailed airflow data – if available at comparable spatial and temporal resolution – would likely improve the model's predictive accuracy.
- (3) Lack of theoretical frameworks for spatial influence ranges: Due to the absence of established theories for optimal buffer ranges, this study relied solely on statistical methods to define

spatial influence levels. Future research should focus on developing theoretical models to define these spatial influence distances.

- (4) Data coverage limitations: The study indicated that temperatures near the Han River were lower than in surrounding areas, but still relatively high compared to Seoul's outskirts. This discrepancy may be due to limited S-DoT measurements in green spaces, water bodies, and mountainous areas, limiting the accuracy of temperature validation. Incorporating LST or other temperature datasets could improve validation.
- (5) Uneven sensor distribution: S-DoT sensors are mainly concentrated in urban centers, leaving suburban and natural areas with sparse data. A more uniform and denser sensor network would provide a more comprehensive dataset, thus improving future urban climate models. Consequently, this uneven sampling structure limits the generalizability of the model in peripheral or natural regions, where prediction uncertainty is inherently higher due to sparse sensor coverage.

By addressing these limitations, future research can refine spatial temperature modeling, improve prediction accuracy, and support more effective urban heat mitigation strategies.

Author contributions

CRedit: **Younju Lee:** Conceptualization, Data curation, Formal analysis, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Changyeon Lee:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by the Chung-Ang University Graduate Research Scholarship in 2024, and this work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) [Grant number RS-2025-16065102].

Notes on contributors

Younju Lee is an M.S. student in the Department of Urban Design and Studies at Chung-Ang University, where she

also earned her B.S. degree in 2024. Her research interests focus on urban thermal environments, urban heat mitigation, and the application of spatial analysis and modeling to address heat vulnerability.

Changyeon Lee is an Assistant Professor of Urban Design and Studies at Chung-Ang University. He received his Ph.D. from the College of Design, Construction and Planning at the University of Florida. He specializes in land use and environmental planning, focusing on how urban physical changes impact the environment, public health, and energy efficiency across various scales. To explore these dynamics and broader spatial phenomena, he leverages big data, advanced statistical methods, GIS spatial analysis, and AI deep learning.

ORCID

Younju Lee  <http://orcid.org/0009-0001-7288-5122>

Changyeon Lee  <http://orcid.org/0000-0001-9936-3415>

Data availability statement

The results and codes of this study are available at <https://github.com/lyj12h/Identifying-Driving-Factors-and-Developing-a-Summer-Air-Temperature-Map.git>

References

- Alavipanah, S., J. Schreyer, D. Haase, T. Lakes, and S. Qureshi. 2018. "The Effect of Multi-Dimensional Indicators on Urban Thermal Conditions." *Journal of Cleaner Production* 177:115–123. <https://doi.org/10.1016/j.jclepro.2017.12.187>.
- Assaf, G., and R. H. Assaad. 2024. "Modeling the Impact of Land Use/Land Cover (LULC) Factors on Diurnal and Nocturnal Urban Heat Island (UHI) Intensities Using Spatial Regression Models." *Urban Climate* 55:101971. <https://doi.org/10.1016/J.UCLIM.2024.101971>.
- Assaf, G., X. Hu, and R. H. Assaad. 2023. "Mining and Modeling the Direct and Indirect Causalities Among Factors Affecting the Urban Heat Island Severity Using Structural Machine Learned Bayesian Networks." *Urban Climate* 49:101570. <https://doi.org/10.1016/J.UCLIM.2023.101570>.
- Ballester, J., M. Quijal-Zamorano, R. F. M. Turrubiates, F. Pegenaute, F. R. Herrmann, J. M. Robine, X. Basagaña, C. Tonne, J. M. Antó, and H. Achebak. 2023. "Heat-Related Mortality in Europe During the Summer of 2022." *Nature Medicine* 29 (7): 1857–1866. <https://doi.org/10.1038/s41591-023-02419-z>.
- Berger, C., J. Rosentreter, M. Voltersen, C. Baumgart, C. Schmullius, and S. Hese. 2017. "Spatio-Temporal Analysis of the Relationship Between 2D/3D Urban Site Characteristics and Land Surface Temperature." *Remote Sensing of Environment* 193:225–243. <https://doi.org/10.1016/j.rse.2017.02.020>.
- Bowler, D. E., L. Buyung-Ali, T. M. Knight, and A. S. Pullin. 2010. "Urban Greening to Cool Towns and Cities: A Systematic Review of the Empirical Evidence." *Landscape and Urban Planning* 97 (3): 147–155. <https://doi.org/10.1016/J.LANDURBPLAN.2010.05.006>.
- Brousse, O., A. Martilli, M. Foley, G. Mills, and B. Bechtel. 2016. "WUDAPT, an Efficient Land Use Producing Data Tool for Mesoscale Models? Integration of Urban LCZ in WRF Over Madrid." *Urban Climate* 17:116–134. <https://doi.org/10.1016/j.uclim.2016.04.001>.
- Chen, J., W. Zhan, P. Du, L. Li, J. Li, Z. Liu, F. Huang, J. Lai, and J. Xia. 2022. "Seasonally Disparate Responses of Surface Thermal Environment to 2D/3D Urban Morphology." *Building & Environment* 214:108928. <https://doi.org/10.1016/j.buildenv.2022.108928>.
- Cheng, V., K. Steemers, M. Montavon, and R. Compagnon. 2006. "Urban Form, Density and Solar Potential." Paper presented at the 23rd Conference on Passive and Low Energy Architecture, Geneva, Switzerland. September 6–8.
- Cho, D., C. Yoo, J. Im, Y. Lee, and J. Lee. 2020. "Improvement of Spatial Interpolation Accuracy of Daily Maximum Air Temperature in Urban Areas Using a Stacking Ensemble Technique." *GIScience and Remote Sensing* 57 (5): 633–649. <https://doi.org/10.1080/15481603.2020.1766768>.
- Chun, B., and J.-M. Guldmann. 2014. "Spatial Statistical Analysis and Simulation of the Urban Heat Island in High-Density Central Cities." *Landscape and Urban Planning* 125:76–88. <https://doi.org/10.1016/j.landurbplan.2014.01.016>.
- De Ridder, K., D. Lauwaet, and B. Maiheu. 2015. "UrbClim – A Fast Urban Boundary Layer Climate Model." *Urban Climate* 12:21–48. <https://doi.org/10.1016/J.UCLIM.2015.01.001>.
- Douglas, I. 2018. "Ecosystems and Human Well-being." *Encyclopedia of the Anthropocene* 4:185–197. <https://doi.org/10.1016/B978-0-12-809665-9.09206-5>.
- ESRI. 2021. "Understanding Ordinary Kriging." <https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/understanding-ordinary-kriging.htm#:~:text=Ordinary%20kriging%20assumes%20the%20model,to%20create%20a%20prediction%20map>.
- Estoque, R. C., Y. Murayama, and S. W. Myint. 2017. "Effects of Landscape Composition and Pattern on Land Surface Temperature: An Urban Heat Island Study in the Megacities of Southeast Asia." *Science of the Total Environment* 577:349–359. <https://doi.org/10.1016/j.scitotenv.2016.10.195>.
- Fushiki, T. 2011. "Estimation of Prediction Error by Using K-Fold Cross-Validation." *Statistics and Computing* 21 (2): 137–146. <https://doi.org/10.1007/s11222-009-9153-8>.
- Gabriel, K. M. A., and W. R. Endlicher. 2011. "Urban and Rural Mortality Rates During Heat Waves in Berlin and Brandenburg, Germany." *Environmental Pollution* 159 (8–9): 2044–2050. <https://doi.org/10.1016/j.envpol.2011.01.016>.
- Gao, C., Y. Feng, X. Tong, Z. Lei, S. Chen, and S. Zhai. 2020. "Modeling Urban Growth Using Spatially Heterogeneous Cellular Automata Models: Comparison of Spatial Lag, Spatial Error and GWR." *Computers, Environment and Urban Systems* 81:101459. <https://doi.org/10.1016/j.compenvurbsys.2020.101459>.
- Gobatti, L., P. M. Bach, M. Maurer, and J. P. Leitão. 2025. "Impact of Soil Moisture Content on Urban Tree Evaporative Cooling and Human Thermal Comfort." *NPJ Urban Sustainability* 5 (1): 26. <https://doi.org/10.1038/s42949-025-00220-0>.
- Gonzalez-Trevizo, M. E., K. E. Martinez-Torres, J. F. Armendariz-Lopez, M. Santamouris, G. Bojorquez-Morales, and A. Luna-Leon. 2021. "Research Trends on Environmental, Energy and Vulnerability Impacts of Urban Heat Islands: An Overview." *Energy & Buildings*

- 246:111051. <https://doi.org/10.1016/j.enbuild.2021.111051>.
- Han, D., H. An, F. Wang, X. Xu, Z. Qiao, M. Wang, X. Sui, et al. 2022a. "Understanding Seasonal Contributions of Urban Morphology to Thermal Environment Based on Boosted Regression Tree Approach." *Building & Environment* 226:109770. <https://doi.org/10.1016/j.buildenv.2022.109770>.
- Han, L., J. Zhao, Y. Gao, and Z. Gu. 2022b. "Prediction and Evaluation of Spatial Distributions of Ozone and Urban Heat Island Using a Machine Learning Modified Land Use Regression Method." *Sustainable Cities and Society* 78:103643. <https://doi.org/10.1016/j.scs.2021.103643>.
- Hart, M. A., and D. J. Sailor. 2009. "Quantifying the Influence of Land-Use and Surface Characteristics on Spatial Variability in the Urban Heat Island." *Theoretical and Applied Climatology* 95 (3–4): 397–406. <https://doi.org/10.1007/s00704-008-0017-5>.
- Hastie, T., R. Tibshirani, and J. Friedman. 2009. "Model Assessment and Selection." In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, edited by T. Hastie, R. Tibshirani, and J. Friedman, 219–259. New York, NY: Springer. https://doi.org/10.1007/978-0-387-84858-7_7.
- He, B.-J., L. Ding, and D. Prasad. 2020a. "Wind-Sensitive Urban Planning and Design: Precinct Ventilation Performance and Its Potential for Local Warming Mitigation in an Open Midrise Gridiron Precinct." *Journal of Building Engineering* 29:101145. <https://doi.org/10.1016/j.jobe.2019.101145>.
- He, B.-J., L. Ding, and D. Prasad. 2020b. "Relationships Among Local-Scale Urban Morphology, Urban Ventilation, Urban Heat Island and Outdoor Thermal Comfort Under Sea Breeze Influence." *Sustainable Cities and Society* 60:102289. <https://doi.org/10.1016/j.scs.2020.102289>.
- Hengl, T., G. B. M. Heuvelink, and D. G. Rossiter. 2007. "About Regression-Kriging: From Equations to Case Studies." *Computers and Geosciences* 33 (10): 1301–1315. <https://doi.org/10.1016/j.cageo.2007.05.001>.
- Hengl, T., G. B. M. Heuvelink, and A. Stein. 2004. "A Generic Framework for Spatial Prediction of Soil Variables Based on Regression-Kriging." *Geoderma* 120 (1–2): 75–93. <https://doi.org/10.1016/j.geoderma.2003.08.018>.
- Hengl, T., M. Nussbaum, M. N. Wright, G. B. M. Heuvelink, and B. Gräler. 2018. "Random Forest as a Generic Framework for Predictive Modeling of Spatial and Spatio-Temporal Variables." *PeerJ* 6:e5518. <https://doi.org/10.7717/peerj.5518>.
- Huang, C., K. Liu, T. Ma, H. Xue, P. Wang, and L. Li. 2025. "Analysis of the Impact Mechanisms and Driving Factors of Urban Spatial Morphology on Urban Heat Islands." *Scientific Reports* 15 (1): 18589. <https://doi.org/10.1038/s41598-025-04025-0>.
- Huang, X., and Y. Wang. 2019. "Investigating the Effects of 3D Urban Morphology on the Surface Urban Heat Island Effect in Urban Functional Zones by Using High-Resolution Remote Sensing Data: A Case Study of Wuhan, Central China." *ISPRS Journal of Photogrammetry & Remote Sensing* 152:119–131. <https://doi.org/10.1016/j.isprsjprs.2019.04.010>.
- Hwang, B. M., H.-D. Sou, J.-H. Oh, and C.-R. Park. 2023. "Cooling Effect of Urban Forests on the Urban Heat Island in Seoul, South Korea." *PLOS ONE* 18 (7): e0288774. <https://doi.org/10.1371/journal.pone.0288774>.
- Jaber, S. M., and R. Sengupta. 2024. "Spatial and Temporal Variabilities in Land Surface Temperatures and Near-Surface Air Temperatures in an Arid to Semiarid Urban Region: Implications for Urban Heat Island Research." *Geo-Spatial Information Science* 27 (6): 2137–2161. <https://doi.org/10.1080/10095020.2024.2310589>.
- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2021. *An Introduction to Statistical Learning with Applications in R*. 2nd ed. New York, NY: Springer. <https://doi.org/10.1007/978-1-0716-1418-1>.
- Jang, J., W. Lee, M. Choi, C. Kang, and H. Kim. 2020. "Roles of Urban Heat Anomaly and Land-Use/Land-Cover on the Heat-Related Mortality in the National Capital Region of South Korea: A Multi-Districts Time-Series Study." *Environment International* 145:106127. <https://doi.org/10.1016/j.envint.2020.106127>.
- Kadaverugu, R., V. Purohit, C. Matli, and R. Biniwale. 2021. "Improving Accuracy in Simulation of Urban Wind Flows by Dynamic Downscaling WRF with OpenFOAM." *Urban Climate* 38:100912. <https://doi.org/10.1016/j.uclim.2021.100912>.
- Kattel, D. B., H. A. M. Salih, T. Yao, and K. Ullah. 2022. "Near-Surface Air Temperature Dependence on Elevation and Geographical Coordinates Over Tropical Desert Land Surfaces." *Frontiers in Earth Science* 9:777381. <https://doi.org/10.3389/feart.2021.777381>.
- Kim, M., J. Won, and H. Kim. 2025. "Assessing Land Use and Urban Form Effects on Summer Air Temperatures Using a City-Wide Environmental Sensor Network in Seoul, South Korea." *Land* 14 (8): 1628. <https://doi.org/10.3390/land14081628>.
- Kim, Y. J., S. Jang, and K. B. Kim. 2023. "Impact of Urban Microclimate on Walking Volume by Street Type and Heat-Vulnerable Age Groups: Seoul's IoT Sensor Big Data." *Urban Climate* 51:101658. <https://doi.org/10.1016/J.UCLIM.2023.101658>.
- Kinney, P. L. 2008. "Climate Change, Air Quality, and Human Health." *The American Journal of Preventive Medicine* 35 (5): 459–467. <https://doi.org/10.1016/J.AMEPRE.2008.08.025>.
- Kong, F., H. Yin, P. James, L. R. Huttyra, and H. S. He. 2014. "Effects of Spatial Pattern of Greenspace on Urban Cooling in a Large Metropolitan Area of Eastern China." *Landscape and Urban Planning* 128:35–47. <https://doi.org/10.1016/J.LANDURBPLAN.2014.04.018>.
- Krige, D. G. 1966. "Two-Dimensional Weighted Moving Average Trend Surfaces for Ore Evaluation." Paper presented at the Symposium on Mathematical Statistics and Computer Applications in Ore Valuation, Johannesburg, March.
- Kumar, R., V. Mishra, J. Buzan, R. Kumar, D. Shindell, and M. Huber. 2017. "Dominant Control of Agriculture and Irrigation on Urban Heat Island in India." *Scientific Reports* 7 (1): 14054. <https://doi.org/10.1038/s41598-017-14213-2>.
- Lee, C. 2022. "Long-Term Wind Speed Interpolation Using Anisotropic Regression Kriging with Regional Heterogeneous Terrain and Solar Insolation in the United States." *Energy Reports* 8:12–23. <https://doi.org/10.1016/j.egyr.2021.11.285>.
- Lee, C. 2024. "How Do Spatial Scale-Dependent Urban and Forest Areas Influence Regionally Seasonal Temperatures in Mid-Latitude South Korean Cities?" *Journal of Environmental Management* 360:121171. <https://doi.org/10.1016/j.jenvman.2024.121171>.

- Li, J., and A. D. Heap. 2008. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Record 2008/23. Canberra: Geoscience Australia.
- Li, S., D. A. Griffith, and H. Shu. 2020. "Temperature Prediction Based on a Space-Time Regression-Kriging Model." *Journal of Applied Statistics* 47 (7): 1168–1190. <https://doi.org/10.1080/02664763.2019.1671962>.
- Li, S., Z. Zhao, X. Miaomiao, and Y. Wang. 2010. "Investigating Spatial Non-Stationary and Scale-Dependent Relationships Between Urban Surface Temperature and Environmental Factors Using Geographically Weighted Regression." *Environmental Modelling & Software* 25 (12): 1789–1800. <https://doi.org/10.1016/J.ENVSOF.2010.06.011>.
- Liou, Y.-A., D.-P. Tran, and K.-A. Nguyen. 2024. "Spatio-Temporal Patterns and Driving Forces of Surface Urban Heat Island in Taiwan." *Urban Climate* 53:101806. <https://doi.org/10.1016/j.uclim.2024.101806>.
- Liu, Z., W. Cheng, C. Y. Jim, T. E. Morakinyo, Y. Shi, and E. Ng. 2021. "Heat Mitigation Benefits of Urban Green and Blue Infrastructures: A Systematic Review of Modeling Techniques, Validation and Scenario Simulation in ENVI-Met V4." *Building & Environment* 200:107939. <https://doi.org/10.1016/j.buildenv.2021.107939>.
- Ma, X., J. Yang, R. Zhang, W. Yu, J. Ren, X. Xiao, and J. Xia. 2024. "XGBoost-Based Analysis of the Relationship Between Urban 2D/3D Morphology and Seasonal Gradient Land Surface Temperature." *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing* 17:4109–4124. <https://doi.org/10.1109/JSTARS.2023.3348476>.
- Ministry of Data and Statistics. 2024. "Grid Statistics (Population)." *Statistical Geographic Information Service*. <https://sgis.kostat.go.kr/view/pss/openDataIntrcn>.
- Ministry of Environment. 2024. "Land Use and Land Cover." *Environmental Geographic Information Service*. <https://egis.me.go.kr>.
- Ministry of Land, Infrastructure and Transport. 2023. "Digital Elevation Model (DEM)." *National Geographic Information Institute*. <http://map.ngii.go.kr/ms/map/NlipMap.do?tabGb=total>.
- Ministry of Land, Infrastructure and Transport. 2024. "Building Use Information and Road Name Address Building Data." V-World. https://www.vworld.kr/dtmk/dtmk_ntads_s002.do?svcCde=NA&dsId=2.
- Ng, E. 2009. "Policies and Technical Guidelines for Urban Planning of High-Density Cities – Air Ventilation Assessment (AVA) of Hong Kong." *Building & Environment* 44 (7): 1478–1488. <https://doi.org/10.1016/J.BUILDENV.2008.06.013>.
- Nikolaou, N., M. Dallavalle, M. Stafoggia, L. M. Bouwer, A. Peters, K. Chen, K. Wolf, and A. Schneider. 2023. "High-Resolution Spatiotemporal Modeling of Daily Near-Surface Air Temperature in Germany Over the Period 2000–2020." *Environmental Research* 219:115062. <https://doi.org/10.1016/j.envres.2022.115062>.
- Osińska-Skotak, K., and J. Zawlich. 2016. "Analysis of Land Use Changes of Urban Ventilation Corridors in Warsaw in 1992–2015." *Geographia Polonica* 89 (3). <https://doi.org/10.7163/GPol.0057>.
- Park, C., J. Ha, and S. Lee. 2017. "Association Between Three-Dimensional Built Environment and Urban Air Temperature: Seasonal and Temporal Differences." *Sustainability* 9 (8): 1338. <https://doi.org/10.3390/su9081338>.
- Park, M., and H. Kim. 2023. "Interaction of Urban Configuration, Temperature, and De Facto Population in Seoul, Republic of Korea: Insights from Two-Stage Least-Squares Regression Using S-Dot Data." *Land* 12 (12): 2110. <https://doi.org/10.3390/land12122110>.
- Park, S., J. Park, and S. Lee. 2023. "Unpacking the Nonlinear Relationships and Interaction Effects Between Urban Environment Factors and the Urban Nighttime Heat Index." *Journal of Cleaner Production* 428:139407. <https://doi.org/10.1016/j.jclepro.2023.139407>.
- Peng, J., J. Jia, Y. Liu, H. Li, and J. Wu. 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>.
- Pyrz, M. J., and C. V. Deutsch. 2006. "Semivariogram Models Based on Geometric Offsets." *Mathematical Geology* 38 (4): 475–488. <https://doi.org/10.1007/s11004-005-9025-5>.
- Qiu, G.-Y., H.-Y. Li, Q.-T. Zhang, W. Chen, X.-J. Liang, and X.-Z. Li. 2013. "Effects of Evapotranspiration on Mitigation of Urban Temperature by Vegetation and Urban Agriculture." *Journal of Integrative Agriculture* 12 (8): 1307–1315. [https://doi.org/10.1016/S2095-3119\(13\)60543-2](https://doi.org/10.1016/S2095-3119(13)60543-2).
- Rich, P. M., W. A. Hetrick, and S. C. Saving. 1995. *Modeling Topographic Influences on Solar Radiation: A Manual for the SOLARFLUX Model*. Los Alamos: NM. <https://doi.org/10.2172/200698>.
- Sekulić, A., M. Kilbarda, D. Protić, M. Perčec Tadić, and B. Bajat. 2020. "Spatio-Temporal Regression Kriging Model of Mean Daily Temperature for Croatia." *Theoretical and Applied Climatology* 140 (1–2): 101–114. <https://doi.org/10.1007/s00704-019-03077-3>.
- Seoul Metropolitan Government. 2024. "S-DoT (Smart Seoul Data of Things) Environmental Information." Seoul Open Data Plaza. <https://data.seoul.go.kr/dataList/OA-15969/S/1/datasetView.do>.
- Seoul Metropolitan Government. 2026. "Registered Population of Seoul Statistics." Seoul Open Data Plaza. <https://data.seoul.go.kr/dataList/419/S/2/datasetView.do>.
- Setyantho, G. R., C. Yuan, and Y. Heo. 2024. "Evaluation of Multi-Layer Urban Canopy Model (MLUCM) for Urban Microclimate Predictions at Different Urban Contexts." *Urban Climate* 55:101882. <https://doi.org/10.1016/j.uclim.2024.101882>.
- Shen, F., C. Xu, and M. Hu. 2023. "Comparison of Approaches to Spatiotemporally Interpolate Land Surface Air Temperature for the Qinghai-Tibet Plateau." *Environmental Earth Sciences* 82 (19): 452. <https://doi.org/10.1007/s12665-023-11151-3>.
- Shi, Y., C. Ren, M. Cai, K. K.-L. Lau, T.-C. Lee, and W.-K. Wong. 2019. "Assessing Spatial Variability of Extreme Hot Weather Conditions in Hong Kong: A Land Use Regression Approach." *Environmental Research* 171:403–415. <https://doi.org/10.1016/j.envres.2019.01.041>.
- Song, J., S. Du, X. Feng, and L. Guo. 2014. "The Relationships Between Landscape Compositions and Land Surface Temperature: Quantifying Their Resolution Sensitivity with Spatial Regression Models." *Landscape and Urban Planning* 123:145–157. <https://doi.org/10.1016/J.LANDURBPLAN.2013.11.014>.
- Stillman, J. H. 2019. "Heat Waves, the New Normal: Summertime Temperature Extremes Will Impact Animals, Ecosystems, and Human Communities." *Physiology* 34 (2): 86–100. <https://doi.org/10.1152/physiol.00040.2018>.

- Su, J. G., M. Jerrett, and B. Beckerman. 2009. "A Distance-Decay Variable Selection Strategy for Land Use Regression Modeling of Ambient Air Pollution Exposures." *Science of the Total Environment* 407 (12): 3890–3898. <https://doi.org/10.1016/j.scitotenv.2009.01.061>.
- SuperMap. 2025. "Ordinary Kriging Interpolation." <https://help.supermap.com/iDesktopX/1101/en/tutorial/Analyst/Raster/interpolation/OrdinaryKriging#parameter-specification>.
- Szymanowski, M., M. Kryza, and W. Spallek. 2013. "Regression-Based Air Temperature Spatial Prediction Models: An Example from Poland." *Meteorologische Zeitschrift* 22 (5): 577–585. <https://doi.org/10.1127/0941-2948/2013/0440>.
- Venter, Z. S., O. Brousse, I. Esau, and F. Meier. 2020. "Hyperlocal Mapping of Urban Air Temperature Using Remote Sensing and Crowdsourced Weather Data." *Remote Sensing of Environment* 242:111791. <https://doi.org/10.1016/j.rse.2020.111791>.
- Vienneau, D., K. de Hoogh, and D. Briggs. 2009. "A GIS-Based Method for Modelling Air Pollution Exposures Across Europe." *Science of the Total Environment* 408 (2): 255–266. <https://doi.org/10.1016/j.scitotenv.2009.09.048>.
- Wang, H., J. Yang, G. Chen, C. Ren, and J. Zhang. 2023. "Machine Learning Applications on Air Temperature Prediction in the Urban Canopy Layer: A Critical Review of 2011–2022." *Urban Climate* 49:101499. <https://doi.org/10.1016/j.uclim.2023.101499>.
- Wang, W., S. Brönnimann, J. Zhou, S. Li, and Z. Wang. 2025. "Near-Surface Air Temperature Estimation for Areas with Sparse Observations Based on Transfer Learning." *ISPRS Journal of Photogrammetry & Remote Sensing* 220:712–727. <https://doi.org/10.1016/j.isprsjprs.2025.01.021>.
- Wang, W., M. Wu, H. Yan, Y. Sang, R. Wu, W. Yang, Z. Zhao, X. Nan, L. Zhu, and Z. Bao. 2025. "Do Healthier Trees Cool Better? Assessing the Impact of Street Tree Health on Urban Thermal Regulation and Pedestrian Comfort." *Building & Environment* 285:113665. <https://doi.org/10.1016/j.buildenv.2025.113665>.
- Webster, R., and M. A. Oliver. 1992. "Sample Adequately to Estimate Variograms of Soil Properties." *Journal of Soil Science* 43 (1): 177–192. <https://doi.org/10.1111/j.1365-2389.1992.tb00128.x>.
- Wu, T., and Y. Li. 2013. "Spatial Interpolation of Temperature in the United States Using Residual Kriging." *Applied Geography* 44:112–120. <https://doi.org/10.1016/j.apgeog.2013.07.012>.
- Xiang, Y., C. Yuan, Q. Cen, C. Huang, C. Wu, M. Teng, and Z. Zhou. 2024. "Heat Risk Assessment and Response to Green Infrastructure Based on Local Climate Zones." *Building & Environment* 248:111040. <https://doi.org/10.1016/j.buildenv.2023.111040>.
- Xin, Y., Y. Xu, X. Tong, Y. Mo, Y. Liu, and S. Zhu. 2025. "Mapping the Air Temperature in China from Time-Normalized MODIS Land Surface Temperature Data via Zone-Based Stacking Ensemble Models." *Geo-Spatial Information Science* (online published). <https://doi.org/10.1080/10095020.2025.2532579>.
- Yin, C., M. Yuan, Y. Lu, Y. Huang, and Y. Liu. 2018. "Effects of Urban Form on the Urban Heat Island Effect Based on Spatial Regression Model." *Science of the Total Environment* 634:696–704. <https://doi.org/10.1016/j.scitotenv.2018.03.350>.
- Yu, Z., G. Yang, S. Zuo, G. Jørgensen, M. Koga, and H. Vejre. 2020. "Critical Review on the Cooling Effect of Urban Blue-Green Space: A Threshold-Size Perspective." *Urban Forestry and Urban Greening* 49:126630. <https://doi.org/10.1016/j.ufug.2020.126630>.
- Zhang, L., J. H. Gove, and L. S. Heath. 2005. "Spatial Residual Analysis of Six Modeling Techniques." *Ecological Modelling* 186 (2): 154–177. <https://doi.org/10.1016/j.ecolmodel.2005.01.007>.
- Zhang, Y., S. Cao, L. Zhao, and J. Cao. 2022. "A Case Application of WRF-UCM Models to the Simulation of Urban Wind Speed Profiles in a Typhoon." *Journal of Wind Engineering & Industrial Aerodynamics* 220:104874. <https://doi.org/10.1016/j.jweia.2021.104874>.
- Zhang, Y., A. Middel, and B. L. Turner. 2019. "Evaluating the Effect of 3D Urban Form on Neighborhood Land Surface Temperature Using Google Street View and Geographically Weighted Regression." *Landscape Ecology* 34 (3): 681–697. <https://doi.org/10.1007/s10980-019-00794-y>.
- Zhou, D., D. Li, G. Sun, L. Zhang, Y. Liu, and L. Hao. 2016. "Contrasting Effects of Urbanization and Agriculture on Surface Temperature in Eastern China." *Journal of Geophysical Research Atmospheres* 121 (16): 9597–9606. <https://doi.org/10.1002/2016JD025359>.
- Zumwald, M., B. Knüsel, D. N. Bresch, and R. Knutti. 2021. "Mapping Urban Temperature Using Crowd-Sensing Data and Machine Learning." *Urban Climate* 35:100739. <https://doi.org/10.1016/j.uclim.2020.100739>.