

Prioritizing environmental determinants of urban heat islands: A machine learning study for major cities in China

Haoran Hou^{a,b,c}, Qianqiu Longyang^a, Hongbo Su^d, Ruijie Zeng^a, Tianfang Xu^a, Zhi-Hua Wang^{a*}

^a *School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85287, USA*

^b *Key Laboratory of Water Cycle & Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources, Chinese Academy of Sciences, Beijing 100101, China*

^c *University of Chinese Academy of Sciences, Beijing 100049, China*

^d *Department of Civil, Environmental and Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, US*

* Corresponding author. Email: zhwang@asu.edu. Tel: +1-480-727-2933

Abstract

The exacerbated thermal environment in cities, the urban heat island (UHI) effect as a prominent example, has been the source of many adverse urban environmental issues, including the increase of health risks, degradation of air quality and ecosystem services, and reduced resiliency of engineering infrastructure. Last decades have witnessed tremendous efforts and resources being invested to find sustainable solutions for urban heat mitigation, whereas the relative contributions of different UHI attributes and their patterns of spatio-temporal variability remain obscure. In this study, we employed the random forest (RF) method to quantify the relative importance of four categories of urban surface characteristics that regulate the surface UHI, namely the urban greenery fraction, land surface albedo, urban morphology, and level of human activities. We selected seventeen major cities from six megaregions in China as our study areas, with the RF training and test sets obtained from multi-sourced remote sensing and observational data products. It is found that the urban greenery coverage manifests as the most important environmental determinants of UHI, followed by surface albedo. The results are informative for urban planners, policymakers, and engineering practitioners to design and implement sustainable strategies for urban heat mitigation.

Keywords: *Albedo; Land surface temperature; Normalized difference vegetation index (NDVI); Random forest; Urban heat island; Urban morphology*

1. Introduction

The urban heat island (UHI) effect, the phenomenon that the built environment is much warmer than its rural surroundings, has attracted rapidly increasing research efforts in past decades (Oke, 1967, 1982; Wang, 2022). In the face of climate changes, the UHI effect interacts positively with global warming as well as increasing heat extremes (McCarthy et al., 2010; Jiang et al., 2019); both tend to exacerbate the thermal environment in cities. Nevertheless, from the perspective of urban planners and engineering practitioners, it is the local and surface UHI determinants that are practically manageable for designing and implementing sustainable countermeasures and are therefore of pivotal importance for improving the thermal environment in urban areas.

The local environmental factors underlying the UHI effect contribute to different pathways of modifying the surface energy balance by human activities. These mechanistic pathways can be broadly classified as: (1) landuse landcover changes, especially the conversion of natural (vegetated) to artificial (paved and impervious) surfaces, and the concomitant alternation of land surface hydrothermal properties (Santero & Horvath, 2009; Yang et al., 2016; Wang et al., 2021a) that reduces the latent heat of evapotranspiration and converts it to sensible heat for heating the built environment, (2) the presence of urban morphology, especially building geometries, in modifying the radiative heat exchange in street canyons (Harman et al., 2004a; Wang, 2014a), (3) the presence of roughness elements (building and transportation infrastructure, shade trees, etc.) that modulates the turbulent transport of heat, moisture, and scalars (Grimmond and Oke, 2002; Harman, 2012; Giometto et al., 2016; Li, Yang. & Yang, 2021), and (4) the emission of anthropogenic heat and greenhouse gases (CO₂ in particular),

primarily through urban metabolism and fossil fuel burning (Allen et al., 2011; Menberg et al., 2013; Song et al., 2017) that directly or indirectly contributes to urban warming.

Past decades have also witnessed the tremendous research efforts devoted to unveiling and quantifying the environmental determinants of UHI that can lead to sustainable, especially nature-based, solutions for heat mitigation. The two most extensively studied UHI determinants are the surface reflectance of solar radiation, i.e. albedo, and the coverage of urban greenery, leading to widely advocated design and implementation of *white* (white roofs and pavements) and *green* (green roofs, green walls, urban gardens/forestry, lawns, shade trees, etc.) infrastructure for heat mitigation (Wang, 2021; Kirsch et al., 2022). The use of so-called cool or super-cool engineering materials can effectively reduce the surface (skin) temperature of urban facet by reflecting substantially larger amount of solar radiation directly back to the atmosphere, but not without unintended consequences (Yang et al., 2015; Wang et al., 2021a). In contrast, urban green space is a significant regulator to the thermal environment mainly through evapotranspirative cooling. Urban greening is among the most popular countermeasures of UHI not just because of its cooling capacity, but also for its environmental co-benefits including improvement of air quality, building energy efficiency, ecosystem services, and aesthetic value.

Other environmental determinants include the roughness and morphology of the built environment. Urban roughness elements mainly consist of the building and transportation infrastructure and, within a proper range, can be conducive to turbulent heat transport and effectively modulate the UHI effect (Wong et al., 2016; Venter et al., 2021), comparable to the effect of albedo or urban greenery. Meanwhile, urban morphology, often represented by various indices such as the building density, the urban canyon aspect (building-height-to-road-width) ratio, or the sky view factor, contributes to the UHI intensity to a not insignificant degree (Li et

al., 2021b; Hou et al., 2023), and often with a nonlinear effect in extremely densely or too sparsely built environment (Theeuwes et al., 2014; Song and Wang, 2015, 2016). Nevertheless, though the impact of roughness length or urban morphology is significant on UHI intensity, these parameters are less manageable than albedo or urban greenery for counteracting the UHI by urban planners or engineers.

The capacity of conventional physically- or statistically-based approaches for determining the impact of different environmental contributors to the UHI effect is largely limited in the sense that they often handle individual factors separately or multiple variables by linear regression or aggregation. It remains an outstanding challenge to quantify multiple determinants of UHI by considering their realistic (and often nonlinear) synergistic interplay in regulating the resultant UHI effect, especially when the dataset gets big with large spatial domains and fine temporal resolutions. This challenge, however, can be readily addressed by recent advances in data science, in particular, the use of machine learning (ML) techniques. The application of ML methods in the UHI study, albeit at its infancy, has already led to some promising results. For example, Venter et al. (2021) probed into different mechanisms of UHI effect using a machine learning algorithm and maintained aerodynamic roughness as the controlling factor of the UHI intensity. ML algorithms have also been used for multi-objective optimization (heat and carbon mitigation) of urban environment system design (Li et al., 2022), or predicting UHI and heatwaves (Oh et al., 2020; Li et al., 2023). A recent study also adopted ML-based analysis for projecting the future UHI effect in Chinese cities based on future climate scenarios and socioeconomic policies (Lan et al., 2023).

In this study, we aim to characterize and quantify the relative significance of different environmental determinants of heat islands, in particular, the surface urban heat island (SUHI)

using a particular ML, viz. the random forest (RF) method. The proposed method is applied to major Chinese cities in six (Northeastern, Northwestern, Jingjinji, Yangtze River Delta, Chengyu, and Pearl River Delta) megaregions. The UHI determinants are quantified using local urban factors including the vegetation coverage in urban areas, surface albedo, the canyon aspect ratio representing urban morphology, and measures of anthropogenic activities. These indices, together with the SUHI intensity, are calculated based on remotely sensed dataset and field observations. By quantifying their relative contribution to the SUHI intensity, we will be able to prioritize the urban planning strategies for more effective amelioration of the urban thermal environment.

2. Study Areas and Data Retrieval

2.1 The study areas

To adequately represent various geographic and climate conditions, in this study, we selected seventeen major cities from six megaregions in China (**Figure 1**). These cities are distributed across a range of climates (Domrös & Peng, 2012), from mid-temperate humid in the north to south-subtropical humid in the south, and from mid-temperate arid in the west to north-subtropical humid in the east. The annual average temperature varies significantly, from 5°C in the Northeastern megaregion to 21°C in the Pearl River Delta megaregion, while the annual average precipitation ranges from 300 mm in the Northwestern megaregion to 1800 mm in the Pearl River Delta megaregion. In addition, the seasonal differences in temperature and precipitation are highly diverse across the six megaregions. For instance, the amplitude of annual temperature variation is 40°C in the northeastern megaregion, but only 18°C in the Pearl River

Delta megaregion; the annual precipitation amplitude is 400 mm in the Jingjinji megaregion, 100 mm Northwestern.

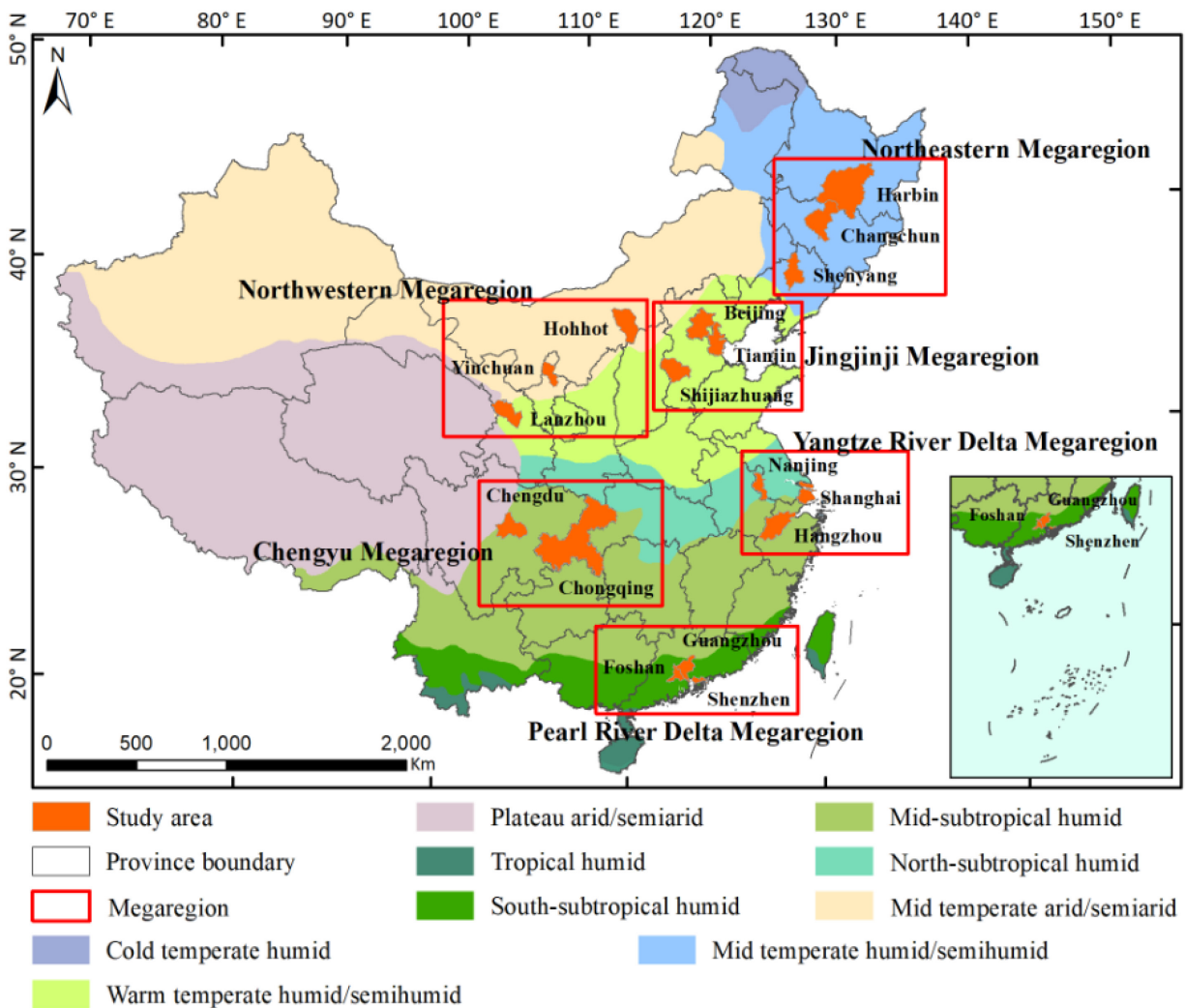


Figure 1. Map of study areas of six Chinese megaregions including Jingjinji, Yangtze River delta, Pearl River delta, Chengyu, Northwestern, and Northeastern megaregion, demarked in red rectangles, containing seventeen major cities. The background color indicates the classification of climate in different regions.

The seventeen cities selected for our study are also regional socioeconomic centers, characterized by concentrated populations and industrial production. The cities in the three east coastal megaregions, in particular, have experienced rapid urbanization and are at the forefront of China's development since the country's opening and reform (Year 1978). **Table 1** summarizes the socioeconomic statistics for these seventeen cities in 2017, obtained from the State Statistics Bureau (<https://data.stats.gov.cn/>). As the build-up areas of Guangzhou and Foshan have merged with each other in their decades-long expansion, they will be studied as one greater Guang-Fo metropolitan in the subsequent analysis.

Table 1. The socioeconomic statistics of the seventeen study cities in 2017

Megaregion	City	Residents (million)	GDP (billion CNY)
Northwestern	Hohhot	3.12	274.37
	Lanzhou	3.74	252.35
	Yinchuan	2.30	180.33
Jingjinji	Beijing	21.95	2988.30
	Shijiazhuang	10.88	646.09
	Tianjin	15.57	1245.06
Northeastern	Changchun	7.49	653.00
	Harbin	10.93	635.50
	Shenyang	8.29	586.50
Yangtze River Delta	Hangzhou	9.47	1260.34
	Nanjing	8.34	1171.51
	Shanghai	24.18	3292.50
Pearl River Delta	Foshan	7.66	954.96
	Guangzhou	14.50	2150.32
	Shenzhen	12.53	2249.01
Chengyu	Chengdu	16.05	1388.94
	Chongqing	31.44	2006.63

2.2 Remote sensing data

In this study, we employed multi-sourced data, including raster datasets for land surface temperature (LST), normalized difference vegetation index (NDVI), land surface albedo, and nighttime light (NTL), and vector datasets of urban buildings and streets to quantify urban

morphology. The raster datasets were retrieved from remotely sensed imagery of the Moderate Resolution Imaging Spectroradiometers (MODIS), onboard the Terra and Aqua satellites managed by U.S. National Aeronautics and Space Administration (NASA). Both Terra- and Aqua-MODIS provides global coverage of moderate resolution images at the frequency of every 1 to 2 days, since their launch around 2000. Raw data obtained by MODIS has been used to generate scientific data products, including the surface reflectance (MOD/MYD09) (Vermote et al., 2002) and vegetation indices (MOD/MYD13) (Huete et al., 1999; Zhang et al., 2017), among others.

MODIS data products have been widely used in various geographical and ecological studies due to their wide spatial coverage and high temporal frequency. In this study, we utilized MODIS data products from 2013 to 2017, including LST and emissivity (MOD/MYD11), vegetation indices (MOD/MYD13), and land cover type (MCD12). The MOD/MYD11 product provides LST images on a daily basis with 1000 m resolution, retrieved from two thermal infrared bands, band 31 (10.780 - 11.280 μm) and band 32 (11.770 - 12.270 μm), using the split-window algorithm (Duan et al., 2019). MOD/MYD13 product contains average NDVI and Enhanced Vegetation Index (EVI) for a 16-day period with a resolution up to 250 m (Huete et al., 2002), which are calculated from the red, near-infrared and blue bands (band 1, 2 and 3 respectively). Unlike the previous two products which processed data from Terra and Aqua separately, the MCD12 product integrates data obtained by both Terra and Aqua satellites and classifies annual land cover types using the decision tree algorithm based on five different standards, with a spatial resolution up to 500 m (Friedl et al., 2002).

The product of land surface albedo is also readily available from remotely sensed dataset (Wan & Li, 1997; Sobrino et al., 2005). In this study, we retrieved the shortwave white sky

albedo in 2013-2017 from the Global Land Surface Satellite (GLASS) dataset. This dataset provides long-term, no-gap global coverage for many critical eco-environmental features, such as the leaf area index, broadband emissivity and gross primary production (Liang et al., 2013, 2021).

In addition, we used nighttime light (NTL) images to represent the intensity of human activities in urban areas, such as population distribution and economic activities (Fan et al., 2014; Levin et al., 2020). The NTL imagery was obtained from the Visible Infrared Imaging Radiometer Suites (VIIRS) dataset (Cao et al., 2013; Wolfe et al., 2013). The VIIRS data products were produced by two satellites, namely Suomi NPP and NOAA-20 launched in 2011 and 2018, respectively, which provides visible and near-infrared data with resolutions of 375 m and 750 m, respectively. NASA has developed a data product called "Black Marble" to improve the precision of NTL imagery acquired by VIIRS, by removing cloud and moonlight pollution through atmospheric and terrain corrections (Román et al., 2018). In this study, we adopted the NTL data from the VNP46 dataset of Black Marble product. Furthermore, all remote sensing datasets used in this study were aggregated to 1000 m in spatial resolution and monthly means for consistency.

2.3 Field observation

In addition to the remote sensing data described above, we also utilized vector-based datasets to extract information on urban morphology. Specifically, we obtained a building distribution dataset from the Resource and Environment Science Data Center of the Chinese Academy of Sciences (<https://www.resdc.cn/data.aspx?DATAID=270>). This dataset includes the location, outline, and number of floors of buildings in build-up areas across several major

Chinese cities. Moreover, we used a road distribution dataset obtained from the OpenStreetMap (OSM) project, an online database of roads and amenities maintained by volunteers and publicly accessible at <https://www.OpenStreetMap.org>. The urban roads in this dataset have been classified into various, e.g. primary, secondary, or residential, categories, based on their applications and construction. Each road category has been assigned to a corresponding road width in accordance with relevant national standards and studies (Sun & Li, 2018). To ensure compatibility with the aforementioned raster data, we converted these vector data to raster format with a spatial resolution of 1000 m. A sample of urban morphology map for part of the city of Shenzhen, using the aforementioned two datasets, is shown in **Figure 2**, with different building height groups and road categories.

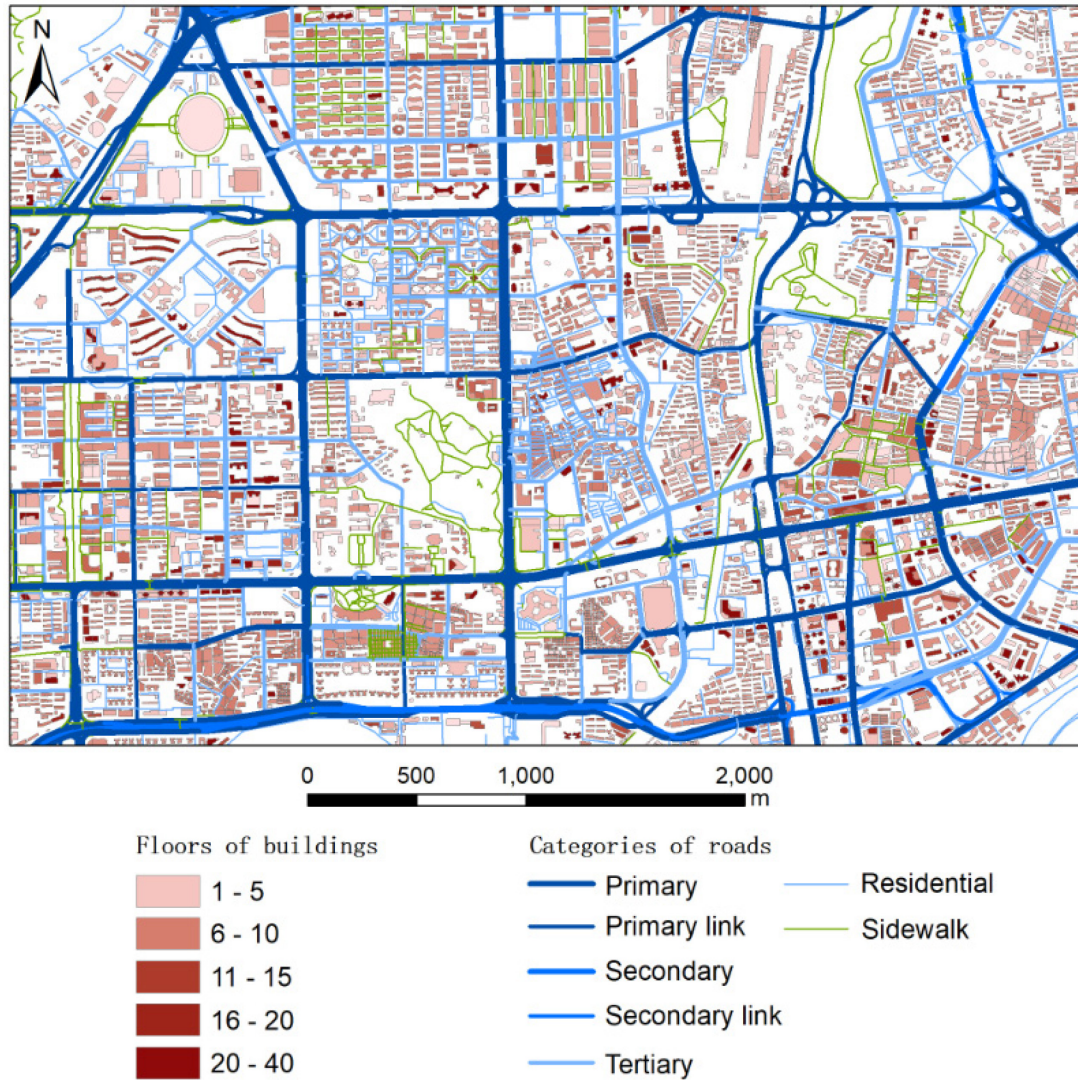


Figure 2. Map of urban morphology for a part of Shenzhen in Pearl River Delta megaregion by combining two vector datasets: the building and road information obtained from Resource and Environment Science Data Center and OpenStreetMap project, respectively.

3. Methods

3.1 Quantification of the SUHI intensity

Thermal environment in cities, in contrast to their rural surroundings, is conventionally quantified as canopy- or boundary-layer UHI (Oke, 1976), measured by air temperatures in the

urban canopy layer or atmospheric boundary layer. These two measures of UHI intensity have different characteristic footprints (source areas) of atmospheric measurements (Wang et al., 2018; Wang, 2022). On the other hand, with the advance of satellite-based thermal remote sensing techniques (Voogt & Oke, 2003), measurements of LST at large spatial coverage become more readily available. The difference between urban and rural LST has been widely used as an indicator of UHI at the surface level, viz. the SUHI (Peng et al., 2012), which enables the study of UHI effect at multiscales ranging from neighborhood to global scales with high economy and accessibility (Mirzaei & Haghighat, 2010). In this study, we adopted the remotely sensed LST, and quantified the SUHI intensity using,

$$\Delta T_s = T_{s,urban} - T_{s,rural} , \quad (1)$$

where T_s is the LST (skin temperature), retrieved from the MODIS MOD/MYD11 data product. We calculated the gridded SUHI intensity in each urban pixel within the boundaries of all study areas, defined by the spatial extent of the vectorial building distribution data. In defining the *urban* pixels, the following filtering criteria is applied, viz. the pixels with incomplete records were excluded from the selected cities, such as water bodies (with LST information from the MOD/MYD11 dataset but without NDVI information from the MOD/MYD13 dataset) or bare land (no building information for calculating urban morphology indicators). In addition, we applied outer buffer region of 10-km width based on each city's administrative boundary (Paschalis et al., 2020), in which the gridcells with natural land cover types (IGBP class 1 to 12) of MODIS MCD12 are classified as the rural surroundings to obtain $T_{s,rural}$ averaged over all *rural* pixels (Chakraborty & Lee; 2019).

3.2 Urban morphology

Urban morphology contributes to the energetic pathway of UHI through the presence of roughness elements and complex building geometry. Roughness elements are represented by key aerodynamic parameters such as roughness length and zero-plane displacement that strongly regulate the turbulent transport of sensible and latent heat (Raupach, 1994; Grimmond & Oke, 1999, 2002; Harman et al., 2004b; Giometto et al., 2016). The geometry of the built environment, on the other hand, alters the pathway of radiative heat exchange via two major competing processes: (1) the shading effect that cools the street canyon and (2) the warming effect by multiple reflections that retains more thermal radiation in street canyons. The synergistic interactions of two processes result in complex and nonlinear effects of urban geometry on the UHI intensity (Li et al., 2020; Hou et al., 2023).

The complex urban morphology can be represented by a variety of geometrical parameters, ranging from the one-dimensional (1D) roughness length and zero-plane displacement, to 2D building frontal and plane areas, to 3D building density and sky view factors (Grimmond & Oke, 1999; Li et al., 2021; Lu et al., 2022; Dewa et al., 2023). These various morphological parameters, however, are not independent, but are all strongly correlated to each other instead (Li et al., 2021; Hou et al., 2023). This is reasonable in the sense that densely built areas in urban cores tend to have high roughness length, large areas of built facets (roofs and walls), high building density, and small sky view factors. In urban climate modeling community, urban areas are customarily represented using “big canyons” (Nunez & Oke, 1977) with two rows of buildings separated by a road, both of infinite longitudinal dimension (canyon depth). In this setting, the urban morphology is quantified using the vertical aspect ratio (VAR) of the street canyon, or the building-height-to-road-width ratio, defined as,

$$\text{VAR} = H / W, \quad (2)$$

where H is the average height of buildings in an urban gridcell, and W the average road width. In this study, we adopt the VAR as the representative indicator of urban morphology, owing to the fact that it maintains an intricate balance of complexity and accuracy for urban climate modeling, UHI predictions included (Grimmond et al., 2010).

3.3 Random Forest algorithm

We utilized the RF algorithm to quantify the influence of various environmental determinants on SUHI as well as their relative significance. The RF model is a highly flexible nonparametric machine learning algorithm, capable of estimating or classifying a target variable (SUHI in this study) using explanatory variables (“features”) and estimating the importance of each feature. It leverages the bootstrap resampling method to draw multiple samples from the original dataset, constructs a decision tree for each bootstrap sample, and combines the predictions of multiple decision trees to reach the final decision prediction through voting for classification and averaging for regression (Breiman, 2001). Each node is split using the best feature selected among a random subset of all features, as opposed to selecting one from all features in the conventional decision tree algorithm (Breiman, 2001). Because of the bootstrapping and random subset strategy, random forest is more robust than decision trees against overfitting and has achieved competitive performance in numerous applications (Liaw & Wiener, 2002; Xu and Liang, 2021).

In this study, we adopted four local environmental variables as features to estimate the SUHI intensity using the RF algorithm. These features include NDVI, surface albedo, VAR, and NTL, representing the contribution from urban green space, paved surfaces, urban morphology,

and intensity of human activities, respectively. These variables, along with the target (SUHI intensity), were first normalized into the interval $[0, 1]$, by

$$X_N = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \quad (3)$$

for subsequent processing by RF.

To compare the impact of environmental variables on SUHI at different time scales and spatial regions, we conducted two sets of experiments. In the first experiment set, we calculated the annual average of all data for each city within a megaregion to investigate the effects of various variables on SUHI intensity across the megaregion on an annual basis. In the second set, we established a seasonal model for all selected cities to compare the impact of environmental variables on SUHI across seasons. In both experiments, the data was split into a training set and a test set, with the training set containing approximately 70% of the data and the test set containing the remaining. To fine-tune the hyperparameters of the RF model, we further divided the training set into five folds to perform 5-fold cross-validation. This approach involved training the model with a given set of hyperparameters on four of the folds and evaluating it on the remaining fold (i.e., the validation set). This process was repeated five times, with a different fold serving as the validation set each time. We evaluated R^2 and mean square error (MSE) on each validation set and average across five nodes to obtain an estimate of generalization error for the given set of hyperparameters. Next, the grid search technique was used to identify the best set of parameters that yielded the lowest 5-fold cross validation error. Using the optimal hyperparameters, we retrain the model on the entire training set and evaluate its performance, in terms of R^2 and MSE on the test set.

Subsequently, we calculated importance scores of the four environmental variables using a permutation-based technique. The technique calculates importance score of a feature by adding a

perturbation to the feature and then determining the difference between the model performance on the original dataset and on the permuted dataset. The score is normalized by dividing this difference with the standard deviation of model performance on the permuted datasets. A higher score signifies greater importance of a feature for predicting the target variable.

The schematic of the analysis framework in this study is shown in Figure 3 using the city of Shanghai as an example. The input to the RF consists of dataset extracted from remotely sensed imagery and ground-based measurements including the vegetation coverage (NDVI), land surface reflectivity (albedo), urban morphology (VAR), and nighttime light images (NTL).

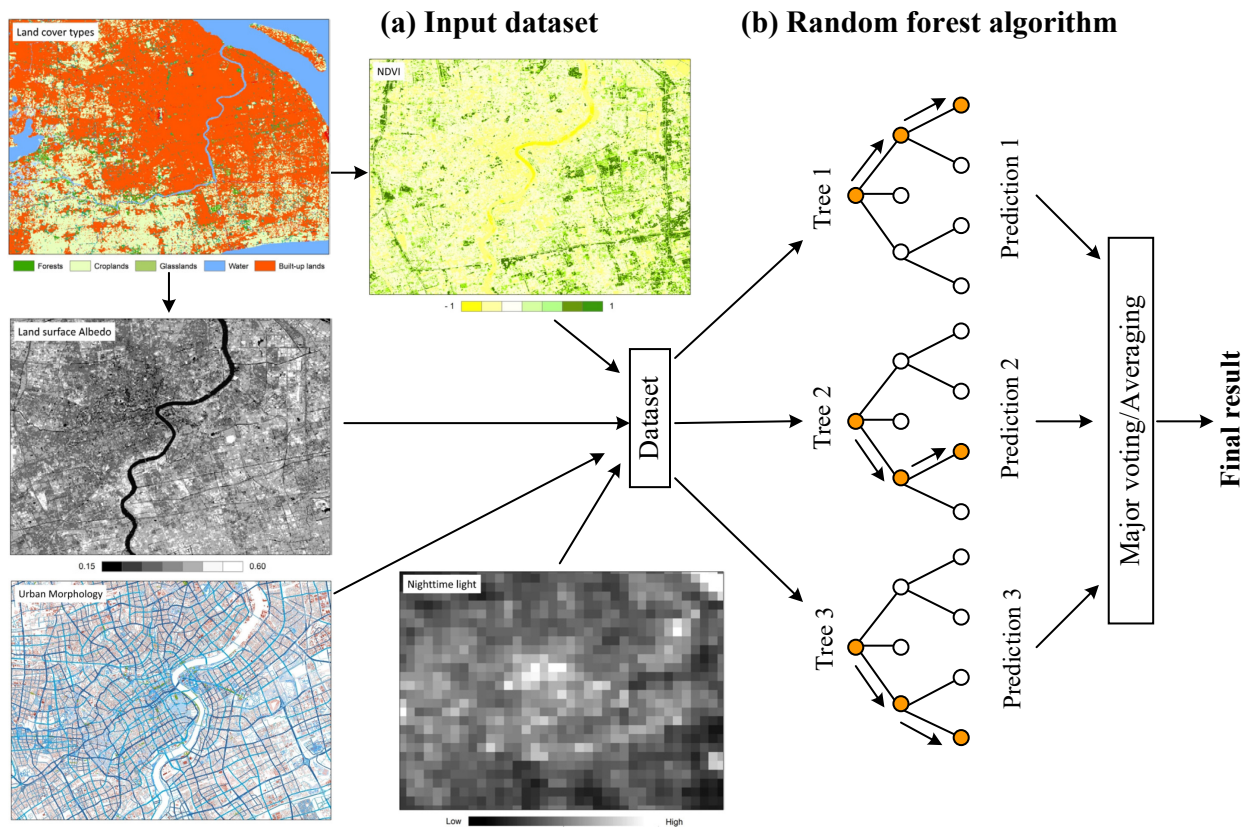


Figure 3. The proposed framework of the current study: (a) Input dataset consists of vegetation coverage, surface albedo, urban morphology, and nighttime light image in the city of Shanghai, and (b) the schematic of RF algorithm.

4. Results and Discussion

In this study, we devised two sets of ML experiments, trained and tested by RF, to evaluate how various environmental variables attributes to heat islands, including: (1) the spatial variability of annual mean SUHI attribution in different megaregions, and (2) the seasonal variation in individual cities. In each set of experiments, we evaluated the performance of RF model using the statistics of R^2 and MSE, and discussed the physical interpretation of the results. In addition, the RF hyperparameters used for training of the two experiments are listed in Tables S1 and S2 in the Supplementary Information, respectively.

4.1 Spatial variability of annual mean SUHI attribution

In addition to seasonal variability, the intensity of heat islands is also subject to climatic conditions and atmospheric forcing. To illustrate, the first set of RF experiments probed into the variation of *annual mean* SUHI attributes in different megaregions in China. The results are demonstrated in **Figure 4**. The first thing noticeable is the difference of dominating factors between northern and southern China: NDVI is the single predominant contributor to heat islands in southern Chinese megaregions, while other factors, such as albedo, become more important in northern megaregions. The predominance of NDVI to SUHI intensities in southern Chinese cities is possibly due to the presence of large area and high density of urban greenery, especially urban trees, as a result of abundant rainfall in sub-tropical climates (Li et al., 2015). It is noteworthy that impact of NDVI, albedo, and surface roughness can be strongly correlated in the presence of urban forest. The low albedo (~ 0.15 - 0.20) of vegetated surfaces can largely offset the effect of white paving materials (e.g. concrete with albedo of ~ 0.4), and the complex

geometry of urban trees synergizes with building clusters in regulating the aerodynamic transport of turbulent heat fluxes. The complex interactions of urban trees and urban forestry with surface albedo and urban morphology also help to explain the relatively low attributions of SUHI to albedo and VAR in cities in southern China. For instance, **Fig. S1** shows the dependence plot (PDP) of albedo and NDVI in the Pearl River Delta megapolitan, where SUHI exhibits varied responses contingent on different NDVI and albedo combinations. More specifically, SUHI intensity minimizes at relatively high NDVI (above 0.45) and low albedo (below 0.25) and stabilizes at a maximum within regions of low NDVI (below 0.3) and high albedo (above 0.35). When NDVI and albedo lie within mid-range, the PDP surface manifests high complexity with multiple peaks and ridges, suggesting interactions between these factors in their influence on SUHI. Similarly, complex interplays exist in other pairs of SUHI attributing factors: an example is shown in **Fig. S2** of the PDP of NDVI and VAR, where SUHI tends to rise sharply when either NDVI or VAR is low (below 0.15), whereas the contribution of VAR to SUHI reduces drastically when NDVI values exceed 0.45.

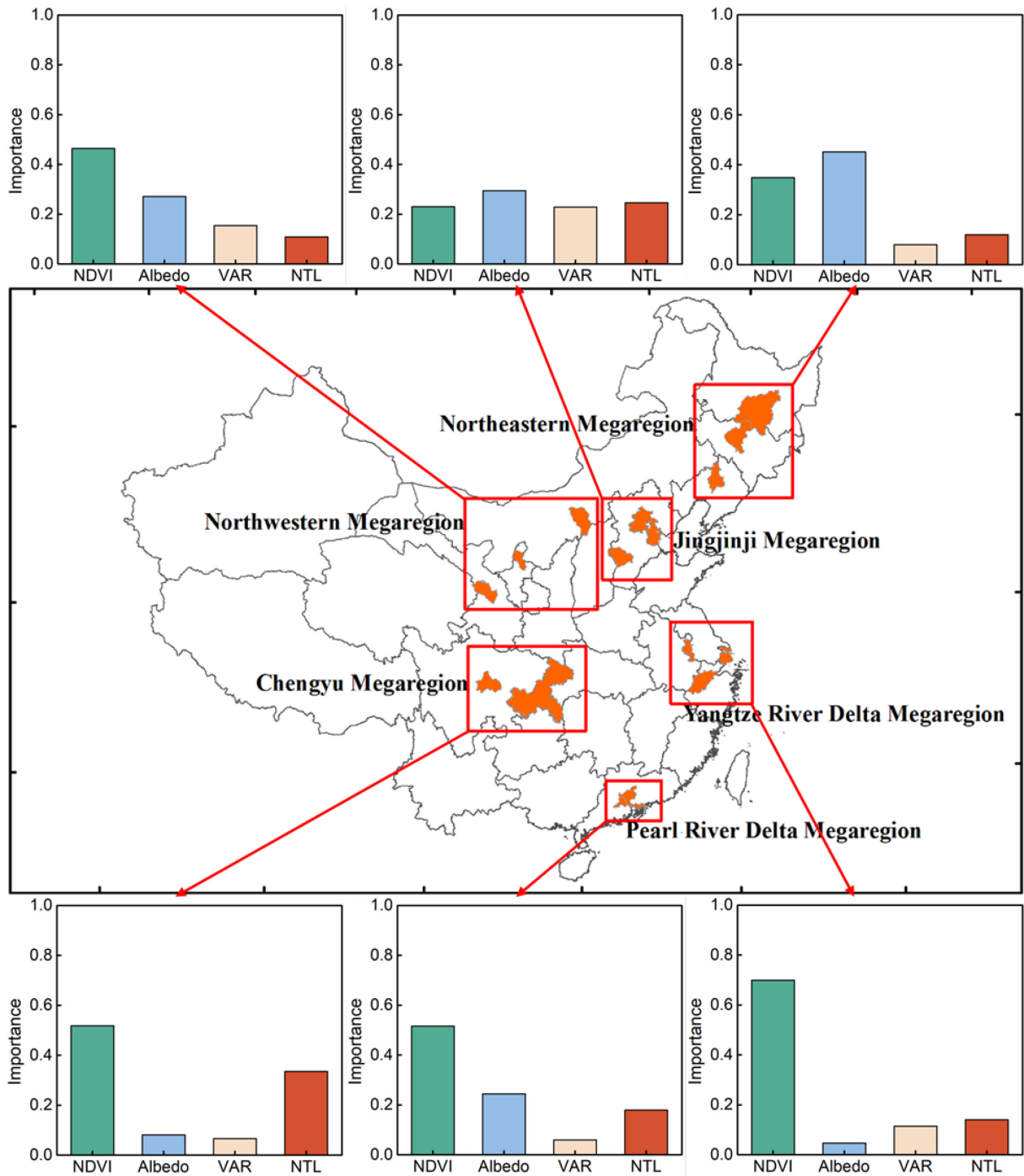


Figure 4. RF-quantified attribution to annual SUHI intensities by four environmental factors, viz. NDVI, albedo, VAR, and NTL, in six megaregions of China.

In addition to the relative scarcity of green spaces in the Jingjinji and Northeastern megaregions, urban vegetation dynamics is also subject to stronger seasonal variability. As a consequence, the NDVI values change significantly from growing to defoliation seasons due to the presence of prevailing deciduous trees in northern Chinese cities. Thus when the seasonal variability is averaged, the annual mean SUHI intensity in northern megaregions of China exhibit much milder response to NDVI. By the same token, we expect that the synergy between NDVI and surface albedo or urban morphology is weakened when temporally averaged as annual means. This also explains that the impact of albedo and VAR becomes comparable to or even more significant than that of NDVI for Jingjinji and Northeastern megaregions, where their effects are disentangled.

Moreover, it is noted that the Northwestern megaregion stand out as quite exceptional in **Figure 4** which could be partially due to the fact that cities in this region are relatively small, and the albedo in Northwestern cities is less distinguishable from that of the prevailing background of arid plateau areas (Zhang et al., 2013). For these relatively small cities in Northwestern China, urban morphology (VAR) becomes an outstanding factor in regulating the urban thermal environment, as it distinguishes the built environment from its rural surroundings.

Furthermore, we evaluated the RF performance using the statistics of R^2 and MSE. The results are summarized in **Table 2**. It is clear that most megaregions investigated in this study exhibit favorable model performance, with R^2 values ranging between 0.5 and 0.9 and MSE reasonably small for training sets. In addition, the performance levels observed for test sets are similar to that of the training sets, indicating that the model has achieved a balance between accuracy and generalizability.

Table 2. Statistics of R^2 and MSE of RF performance on training and test data, respectively, for predicting annual mean SUHI intensities in six Chinese megaregions

Megaregions	Train		Test	
	R^2	MSE	R^2	MSE
Chengyu	0.773	0.395	0.755	0.415
Jingjinji	0.588	1.342	0.593	1.352
Northeastern	0.901	0.567	0.887	0.556
Northwestern	0.661	0.549	0.623	0.460
Pearl River Delta	0.674	1.547	0.694	1.381
Yangtze River Delta	0.635	0.427	0.637	0.414

4.2 SUHI determinants for individual cities with seasonal variability

We then looked into the spatio-temporal variability of SUHI attribution in individual cities with seasonal variability. The results are shown in **Figures 5-10** for the megaregions of Northwestern, Jingjinji, Northwestern, Chengyu, Pearl River Delta, and Yangtze River Delta, respectively. The results look rather diverse and have strong locality as well as seasonal dependence. Nevertheless, we can observe some general patterns. First, NDVI appears to be, overall, the most important environmental determinant of the heat island effect. This is particularly true when we consider the seasonal variation of SUHI attributes, in comparison with the annual mean SUHI intensities in **Figure 4**, the distinction between cities in northern and southern China largely disappears when seasonality is taken into consideration.

In general, the importance of NDVI in regulating the SUHI intensity is more significant in warm seasons (summer and autumn) than in cold seasons, but with occasional exceptions. We speculate that there might be potential *hysteresis* effect in the seasonal variation of NDVI in modulating as well as responding to urban thermal environment. This is plausible as urban vegetation, in comparison to other non-biogenic factors, alters the surface energy balance in less direct pathway via evapotranspirative cooling by changing the relative partitioning of sensible and latent heat (Li & Wang, 2020). Furthermore, the growth of urban vegetation can be enhanced

in positive response to the warming in urban environment (Zhao et al., 2016). The bi-directional urban vegetation-heat interactions contain intrinsic hysteresis effect in both directions as it takes time for biogenic functions of urban vegetation to act on and react to the change of thermal environment. Such hysteresis effect has been observed in urban environment in terms of phase lags among different heat fluxes (Sun et al., 2013; Wang, 2014b) and temperature-CO₂ interactions (Zhang et al., 2015).

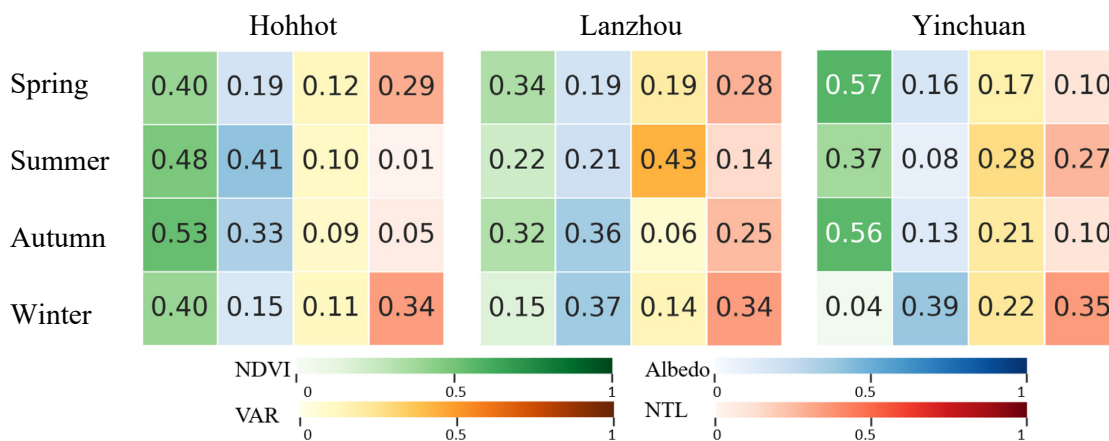


Figure 5. RF-quantified attribution of NDVI, albedo, VAR, and NTL to seasonal SUHI intensities in the cities of Hohhot, Lanzhou and Yinchuan of the Northwestern megaregion. The values shown in each cell represents the feature importance determined by the RF model, with number in each row of each city sum up to unity.

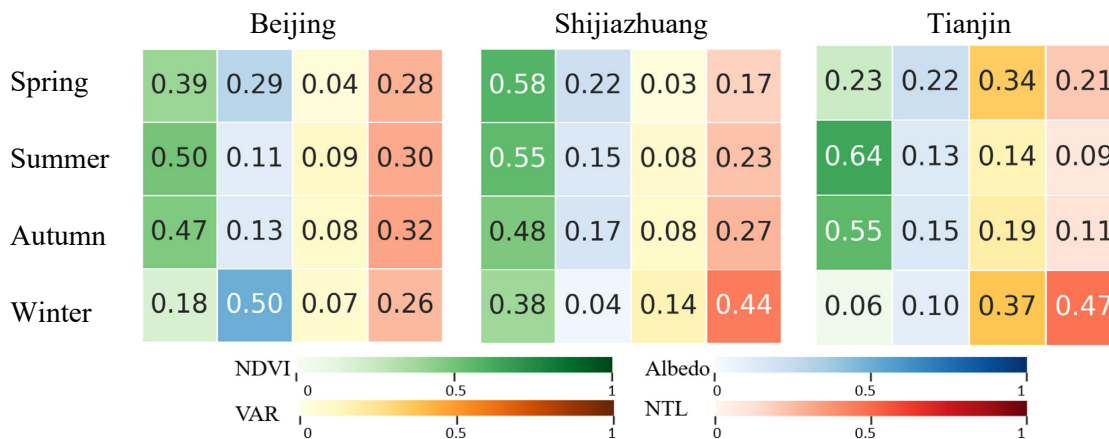
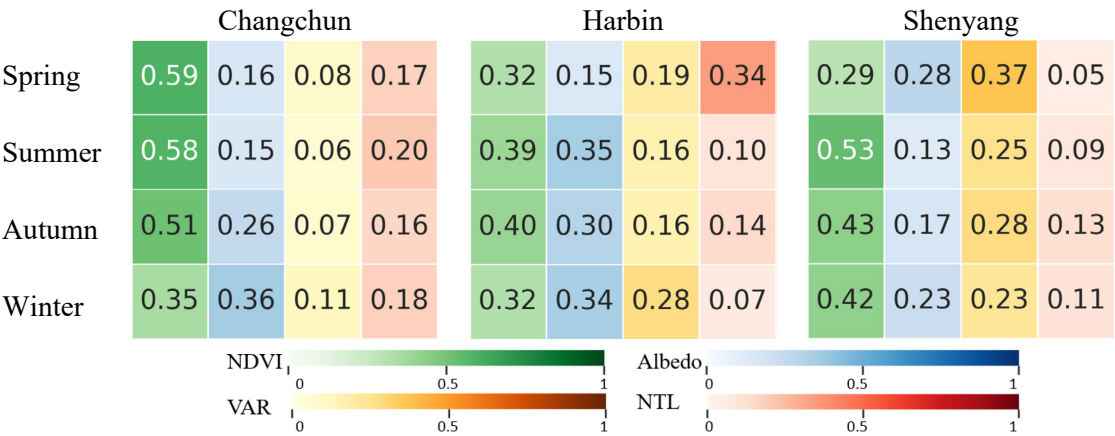


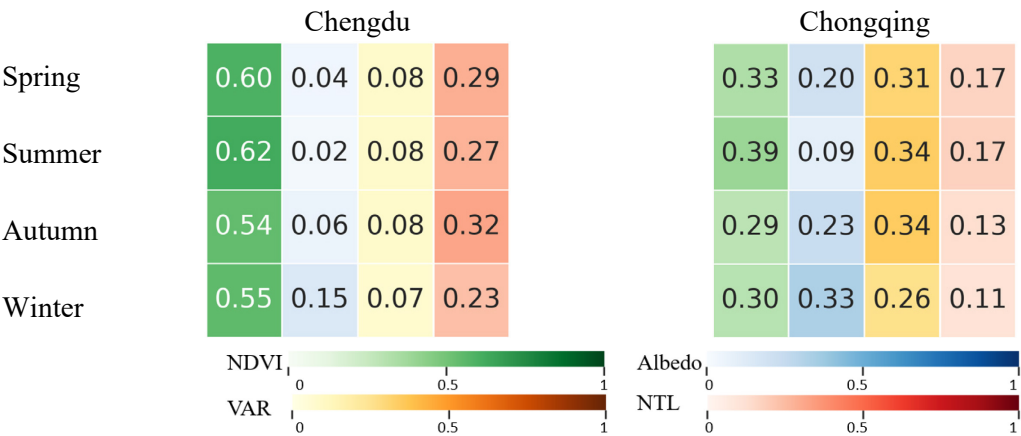
Figure 6. Same as Figure 5, but for the Jingjinji megaregion.

397



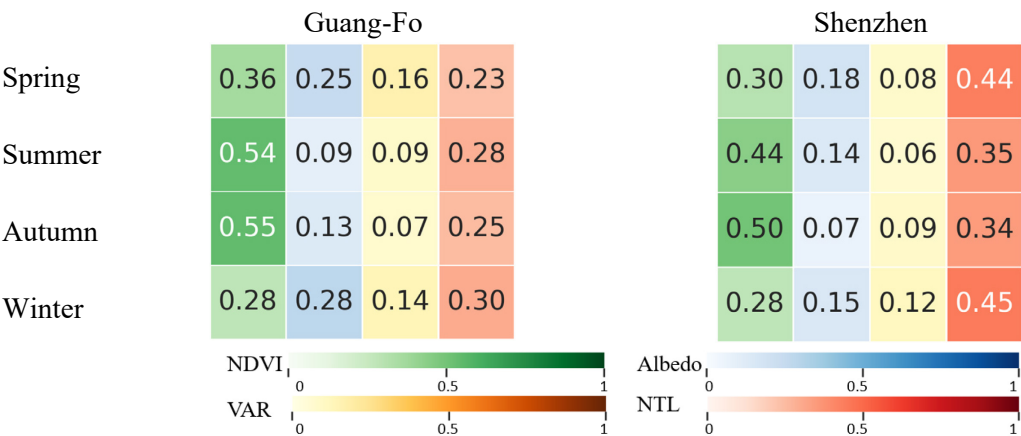
398 **Figure 7.** Same as Figure 5, but for the Northeastern megaregion.

399



400 **Figure 8.** Same as Figure 5, but for the Chengyu megaregion.

401



402 **Figure 9.** Same as Figure 5, but for the Pearl River Delta megaregion.

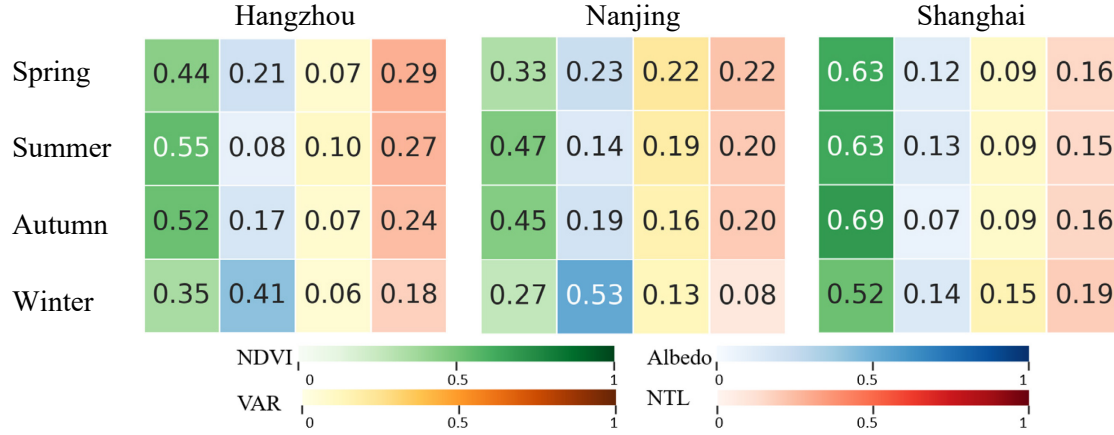


Figure 10. Same as Figure 5, but for the Yangtze River Delta megaregion.

Besides NDVI, albedo is the second most important attribute to SUHI in northern Chinese cities, but its significance is weaker and often out-weighted by VAR or NTL in southern megaregions. In addition, the importance of albedo is usually more manifest in cold seasons (winter and spring), possible due to the reduced impact of NDVI in these seasons, especially if deciduous broadleaf urban trees prevail in the built environment. Considering the remaining factors of VAR and NTL, their significance shows up in some cities in certain seasons (e.g. in Shenzhen or Chongqing), but there is no general pattern observed. Moreover, these parameters are less manageable as far as UHI mitigation strategies are concerned.

The model performance for this set of RF experiments is shown in **Table 3** for individual cities. In most cases, the RF performance is of good accuracy and robust. A number of cities, including Beijing, Chengdu, Guang-Fo, Hangzhou, Hohhot, Shenzhen, and Shijiazhuang, exhibited consistently better-the-average performance throughout all seasons, with R^2 scores ranging from 0.7 to 0.9 for both training and test sets. Nevertheless, in some cities, the RF model exhibits good performance in specific seasons but encountered difficulties in others. For instance, in Tianjin, the RF model effectively captured the relationships between explanatory variables and SUHI during the summer season but failed to replicate these patterns in the other

three seasons. This inconsistency in model performance might be attributed to the fluctuating seasonal patterns of SUHI's dependency on the selected environmental variables. In a particular season, when at least one variable exhibits a pronounced relationship with SUHI, model performance tends to be satisfactory. As an example, in Beijing, also located in the Jingjinji city group, all four factors—NDVI, albedo, NTL, and VAR—display distinct correlations with SUHI. These varying dependency patterns can substantially impact the model's ability to accurately characterize the intricate relationships between surface characteristics and SUHI intensity throughout the year.

Table 3. The statistics of RF performance (R^2) for predicting SUHI attributes in selected Chinese cities with seasonal variability

	Spring		Summer		Autumn		Winter	
	Train	Test	Train	Test	Train	Test	Train	Test
Beijing	0.862	0.850	0.941	0.952	0.879	0.885	0.798	0.755
Changchun	0.650	0.607	0.899	0.905	0.780	0.779	0.722	0.749
Chengdu	0.790	0.787	0.909	0.907	0.766	0.766	0.790	0.744
Chongqing	0.720	0.713	0.849	0.843	0.828	0.814	0.428	0.411
Guang-Fo	0.751	0.754	0.869	0.870	0.892	0.883	0.819	0.804
Hangzhou	0.851	0.759	0.907	0.894	0.883	0.825	0.840	0.836
Harbin	0.492	0.467	0.857	0.827	0.842	0.831	0.838	0.787
Hohhot	0.804	0.839	0.944	0.917	0.877	0.857	0.751	0.748
Lanzhou	0.554	0.593	0.639	0.631	0.856	0.832	0.861	0.859
Nanjing	0.764	0.750	0.877	0.856	0.805	0.792	0.420	0.431
Shanghai	0.745	0.764	0.849	0.842	0.844	0.864	0.541	0.530
Shenyang	0.466	0.438	0.931	0.930	0.712	0.787	0.578	0.531
Shenzhen	0.832	0.815	0.872	0.860	0.876	0.853	0.778	0.777
Shijiazhuang	0.900	0.854	0.952	0.939	0.772	0.705	0.712	0.702
Tianjin	0.390	0.318	0.790	0.802	0.494	0.493	0.191	0.185
Yinchuan	0.553	0.548	0.656	0.656	0.592	0.578	0.276	0.261

A critical constraint in this study is the size of available data points, which is closely associated to the spatial and temporal resolutions and can impact the model accuracy. The amount of training data available for seasonal models for each city depends on the city size at the same spatial resolution, which may result in an inadequate representation of real-world conditions due to insufficient observations in smaller cities (e.g. those in the Northwestern megaregion). Furthermore, the spatial and temporal resolutions of the dataset may not be sufficiently fine-grained to capture the intricate dynamics of SUHI with adequate sub-grid heterogeneity. Employing finer resolution data, when available, could unveil more complex patterns and dependencies, leading to enhanced model performance and a deeper understanding of the underlying processes. Consequently, future studies should consider incorporating more detailed data with higher resolutions to improve the accuracy and precision of SUHI modeling.

4.3 Implications to UHI mitigation strategies

Previous results showed that the presence of urban greenery, measured by NDVI, is overall the most significant determinant to SUHI intensity in all the selected Chinese cities, especially in southern megaregions and in warm (growing) seasons. This is consistent with the prevailing practice of sustainable urban planning in advocating the use of nature-based solutions for heat mitigation. In this regard, urban greening, together with irrigation that supports the biogenic functions of vegetation, remains the most important UHI countermeasure, and is more sustainable, ecosystem friendly, and aesthetically appealing than the use of white (highly reflective) materials in urban infrastructure.

In contrast, the use of white roofs and white pavements continues to attract advocates in many cities, most likely due to the simplicity of its working mechanism as well as low

maintenance cost. Despite its simplicity in working mechanisms, there remains knowledge gaps that needs to be addressed in the use of white pavements for heat mitigation, especially the lack of transparency and trustworthy information regarding its implementation and efficacy (Wang et al., 2021a, 2021b). In particular, the unintended consequence of high albedo paving materials has hitherto been largely overlooked, such as its potential degradation to urban ecosystems, and health risks to sensitive population groups such as children and the elderly (Yang et al., 2015). Moreover, white roofs and white pavements are not aesthetically appealing or help to meet recreational needs to many urban residents in diverse cultural groups (Wang, 2021).

On the other hand, the parameters NTL signaling human activities and VAR representing urban morphology, in general, are of less importance than NDVI and albedo in regulating the SUHI intensity and its seasonal variation. This could be partly due to the fact that these two indices are relative constant and lack seasonal variability. In terms of heat mitigation, it is also less manageable for urban planners and engineers to significantly alter the patterns of urban morphology or anthropogenic activities in existing cities without causing substantial disturbance or even chaos in urban life. Nevertheless, accurate quantification of the attribution of urban morphology and anthropogenic activities to the UHI effect is imperatively necessary. This will be particularly informative to urban planners and policy makers in building sustainable future cities over still unoccupied and natural terrains.

Lastly, we can learn from the high spatio-temporal variability of SUHI attributions that practical design and implementation of sustainable heat mitigation solutions should be based on local environmental (and socioeconomic) conditions. It is clear from the results of seasonal variation of SUHI attribution in individual cities (**Figures 5-10**) that the most effective means for combating SUHI in a given city may not be compatible with that suggested by the spatially-

and/or temporally-aggregated result (c.f. **Figure 4** for annual mean SUHI aggregated in megaregions). More importantly, to effectively counteract the UHI effect in different cities, it is more practical to have a portfolio of mixed and well-coordinated solutions, instead of a singular strategy (such as urban greening or the use of white pavements). For example, in Harbin or Nanjing, all four SUHI attributes are of comparable importance, albeit varying seasonally, and they jointly determine the heat island effect. Therefore *an* optimal countermeasure of UHI in these cities could be a combined approach using white and green infrastructure, together with careful planning of building density and regulation of building operation and transportation, in future urban expansion or re-development.

5. Concluding Remarks

In this study, we quantified different environmental attributes to urban heat islands measured by the land surface temperature in six megaregions of China, using the RF method, a popular machine learning algorithm suitable for nonlinear regression problems. Future research endeavor to employ ML algorithms for UHI studies is therefore encouraged to make intercomparisons of numerical performance of different ML algorithms, in the hope that more specific ML method tailored for urban climate dynamics should prevail. Out of the four representative environmental indices, viz. NDVI, albedo, VAR, and NTL, we found that overall the presence of urban greenery (NDVI) emerges as the most significant determinant of the surface UHI effect in the selected major Chinese cities, followed by the surface albedo. It is caveated that the list of selected environmental indices in this study is neither unique nor exhaustive. For all the selected major categories of SUHI attributes, e.g. vegetation fraction, built-up areas, urban morphology, and anthropogenic activities, each can be measured by

multiple indices. For example, urban morphology can be readily measured by the aerodynamic roughness length, building density, or sky view factor, in addition to VAR, to name a few. Moreover, the influences of the selected indices on heat islands are by no means independent. On the contrary, they often interact synergistically to regulate the urban thermal environment via complex pathways for redistributing surface energy budgets.

In addition, in terms of UHI mitigation to promote human thermal comfort, some critical factors may have been missed from the current study. For instance, for cities located in humid environment, such as those in the Pearl River Delta or Hong Kong, an effective means for ameliorating human thermal comfort of urban residents is to maintain breezeways of air flow (HKPSG, 2022). The impact of breezeway design, albeit related to urban morphology, is inadequately represented by the current study, as it requires much finer temporal (sub-hour) scale to capture the highly turbulent air flow that affect human thermal comfort combining with temperature and humidity variables. Nevertheless, the findings of RF simulations in this study provide useful urban planning guidelines to counteract the UHI effect. In particular, it corroborates and echoes what we have proposed in a previous study that there is unlikely to be a *one-solution-fits-all strategy* for urban heat mitigation (Yang et al., 2015). The sooner we abandon the myth of seeking a “silver bullet” for UHI, the better the chance we have for finding more sustainable solutions that are likely to be locality-dependent and consist of a portfolio of coordinated urban heat mitigation and adaptation strategies.

521 **Acknowledgement**

522 ZHW acknowledges the funding support from the U. S. National Science Foundation
523 (NSF) under grant No. CBET-2028868 and AGS-2300548, and the National Aeronautics and
524 Space Administration (NASA) under grant No. 80NSSC20K1263. The study is also supported
525 by the National Natural Science Foundation of China under grand No. 41971315.

References:

- Allen, L., Lindberg, F., & Grimmond, C. S. B. (2011), Global to city scale urban anthropogenic heat flux: model and variability, *International Journal of Climatology*, 31(13), 1990-2005.
- Breiman, L. (2001), Random forests. *Machine learning*, 45, 5-32.
- Cao, C., Xiong, J., Blonski, S., Liu, Q., Upreti, S., Shao, X., et al. (2013), Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring, *Journal of Geophysical Research: Atmospheres*, 118(20), 11664-11678.
- Chakraborty, T., & Lee, X. (2019). A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. *International Journal of Applied Earth Observation and Geoinformation*, 74, 269-280.
- Dewa, D.D., Buchori, I., Rudiarto, I., & Sejati, A.W. (2023), Modifying the contact perimeter approach for measuring urban compactness gradients in the Joglosemar urban region, Indonesia, *Journal of Geovisualization and Spatial Analysis*, 7(1), 4.
- Domrös, M., & Peng, G. (2012), The climate of China. Springer Science & Business Media.
- Duan, S.-B., Li, Z.-L., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., et al. (2019), Validation of Collection 6 MODIS land surface temperature product using in situ measurements, *Remote Sensing of Environment*, 225, 16-29.
- Fan, J., Ma, T., Zhou, C., Zhou, Y., & Xu, T. (2014), Comparative estimation of urban development in China's cities using socioeconomic and DMSP/OLS night light data, *Remote Sensing*, 6(8), 7840-7856.
- Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002), Global land cover mapping from MODIS: algorithms and early results, *Remote*

549 *Sensing of Environment*, 83(1), 287-302.

550 Giometto, M. G., Christen, A., Meneveau, C., Fang, J., Krafczyk, M., & Parlange, M. B. (2016),

551 Spatial characteristics of roughness sublayer mean flow and turbulence over a realistic

552 urban surface, *Boundary-Layer Meteorology*, 160(3), 425-452.

553 Grimmond, C. S. B., Blackett, M., Best, M. J., Barlow, J., Baik, J. J., Belcher, S. E., et al. (2010),

554 The international urban energy balance models comparison project: First results from Phase

555 1, *Journal of Applied Meteorology and Climatology*, 49(6), 1268-1292.

556 Grimmond, C. S. B., & Oke, T. R. (1999), Aerodynamic properties of urban areas derived from

557 analysis of surface form, *J. Appl. Meteorol.*, 38(9), 1262-1292.

558 Grimmond, C. S. B., & Oke, T. R. (2002), Turbulent heat fluxes in urban areas: Observations

559 and a local-scale urban meteorological parameterization scheme (LUMPS), *Journal of*

560 *Applied Meteorology*, 41(7), 792-810.

561 Harman, I. N. (2012), The role of roughness sublayer dynamics within surface exchange

562 schemes, *Boundary-Layer Meteorology*, 142(1), 1-20.

563 Harman, I. N., Best, M. J., & Belcher, S. E. (2004a), Radiative exchange in an urban street

564 canyon, *Boundary-Layer Meteorol.*, 110(2), 301-316.

565 Harman, I. N., Barlow, J. F., & Belcher, S. E. (2004b), Scalar fluxes from urban street canyons.

566 Part II: Model, *Boundary-Layer Meteorology*, 113(3), 387-409.

567 HKPSG, (the Hong Kong Planning Standards and Guidelines) (2022), *Chapter 11: Urban*

568 *Design Guidelines*, Planning Department of the Government of the Hong Kong Special

569 Administrative Region, 56 pp.

570 Hou, H., Su, H., Yao, C., & Wang, Z.-H. (2023), Spatiotemporal patterns of the impact of

571 surface roughness and morphology on urban heat island, *Sustainable Cities and Society*, 92,

572 104513.

573 Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002), Overview
574 of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote*
575 *Sensing of Environment*, 83(1), 195-213.

576 Huete, A., Justice, C., & van Leeuwen, W. (1999), MODIS vegetation index (MOD 13)
577 algorithm theoretical basis document version 3. University of arizona 1200.

578 Jiang, S. J., Lee, X., Wang, J. K., & Wang, K. C. (2019), Amplified urban heat islands during
579 heat wave periods, *Journal of Geophysical Research-Atmospheres*, 124(14), 7797-7812.

580 Kirsch, K.R., Newman, G.D., Zhu, R., McDonald, T.J., Xu, X., & Horney, J.A. (2022), Applying
581 and integrating urban contamination factors into community garden siting, *Journal of*
582 *Geovisualization and Spatial Analysis*, 6(2), 33.

583 Lan, T., Peng, J., Liu, Y., Zhao, Y., Dong, J., Jiang, S., et al. (2023), The future of China's urban
584 heat island effects: A machine learning based scenario analysis on climatic-socioeconomic
585 policies, *Urban Climate*, 49, 101463.

586 Levin, N., Kyba, C. C. M., Zhang, Q., Sánchez de Miguel, A., Román, M. O., Li, X., et al.
587 (2020), Remote sensing of night lights: A review and an outlook for the future, *Remote*
588 *Sensing of Environment*, 237, 111443.

589 Li, F., Wang, R., Liu, X., & Zhang, X. (2005), Urban forest in China: Development patterns,
590 influencing factors and research prospects, *International Journal of Sustainable*
591 *Development & World Ecology*, 12(2), 197-204.

592 Li, H., Li, Y., Wang, T., Wang, Z. H., Gao, M., & Shen, H. (2021), Quantifying 3D building
593 form effects on urban land surface temperature and modeling seasonal correlation patterns,
594 *Building and Environment*, 204, 108132.

595 Li, P. & Wang, Z. H. (2020), A nonequilibrium thermodynamic approach for surface energy
596 balance closure, *Geophysical Research Letters*, 47(3), e2019GL085835.

597 Li, P., Xu, T., Wei, S., & Wang, Z. H. (2022), Multi-objective optimization of urban
598 environmental system design using machine learning, *Computers, Environment and Urban
599 Systems*, 94, 101796.

600 Li, P., Yu, Y., Huang, D., Wang, Z. H., & Sharma, A. (2023), Regional heatwave prediction
601 using Graph Neural Network and weather station data, *Geophysical Research Letters*, 50,
602 e2023GL103405.

603 Li, Q., Yang, J., & Yang, L. (2021), Impact of urban roughness representation on regional
604 hydrometeorology: An idealized study, *Journal of Geophysical Research: Atmospheres*,
605 126(4), e2020JD033812.

606 Li, Y., Schubert, S., Kropp, J. P., & Rybski, D. (2020), On the influence of density and
607 morphology on the Urban Heat Island intensity, *Nature Communications*, 11(1), 2647.

608 Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., et al. (2021), The Global Land Surface
609 Satellite (GLASS) Product Suite, *Bulletin of the American Meteorological Society*, 102(2),
610 E323-E337.

611 Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., et al. (2013), A long-term Global
612 LAnd Surface Satellite (GLASS) data-set for environmental studies, *International Journal
613 of Digital Earth*, 6, 5-33.

614 Liaw, A., & Wiener, M (2002), Classification and regression by randomForest. *R News*, 2(3), 18-
615 22.

616 Lu, H., Li, F., Yang, G., & Sun, W. (2021), Multi-scale impacts of 2D/3D urban building pattern
617 in intra-annual thermal environment of Hangzhou, China, *International Journal of Applied*

618 *Earth Observation and Geoinformation*, 104, 102558.

619 McCarthy, M. P., Best, M. J., & Betts, R. A. (2010), Climate change in cities due to global
620 warming and urban effects, *Geophysical Research Letters*, 37, L09705.

621 Menberg, K., Blum, P., Schaffitel, A., & Bayer, P. (2013), Long-term evolution of anthropogenic
622 heat fluxes into a subsurface urban heat island, *Environmental Science & Technology*,
623 47(17), 9747-9755.

624 Mirzaei, P. A., & Haghighat, F. (2010), Approaches to study Urban Heat Island – Abilities and
625 limitations, *Building and Environment*, 45(10), 2192-2201.

626 Nunez, M., & Oke, T. R. (1977), The energy balance of an urban canyon, *Journal of Applied*
627 *Meteorology*, 16(1), 11-19.

628 Oh, J. W., Ngarambe, J., Duhirwe, P. N., Yun, G. Y., & Santamouris, M. (2020), Using deep-
629 learning to forecast the magnitude and characteristics of urban heat island in Seoul Korea,
630 *Scientific Reports*, 10(1), 3559.

631 Oke, T. R. (1967), City size and the urban heat island, *Atmospheric Environment*, 7, 769-779.

632 Oke, T. R. (1976), The distinction between canopy and boundary-layer urban heat islands,
633 *Atmosphere*, 14, 268-277.

634 Oke, T. R. (1982), The energetic basis of the urban heat island, *Quarterly Journal of the Royal*
635 *Meteorological Society*, 108(455), 1-24.

636 Paschalis, A., Chakraborty, T. C., Fatichi, S., Meili, N., & Manoli, G. (2021). Urban forests as
637 main regulator of the evaporative cooling effect in cities. *AGU Advances*, 2(2),
638 e2020AV000303.

639 Peng, S. S., Piao, S. L., Ciais, P., Friedlingstein, P., Ottle, C., Breon, F. M., et al. (2012), Surface
640 urban heat island across 419 global big cities, *Environmental Science & Technology*, 46(2),

641 696-703.

642 Raupach, M. R. (1994), Simplified expressions for vegetation roughness length and zero-plane
643 displacement as functions of canopy height and area index, *Boundary-Layer Meteorology*,
644 71(1), 211-216.

645 Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., et al. (2018), NASA's
646 Black Marble nighttime lights product suite, *Remote Sensing of Environment*, 210, 113-
647 143.

648 Santero, N. J., & Horvath, A. (2009), Global warming potential of pavements, *Environmental*
649 *Research Letters*, 4(3), 034011.

650 Sobrino, J. A., Jiménez-Muñoz, J. C., & Verhoef, W. (2005), Canopy directional emissivity:
651 Comparison between models, *Remote Sensing of Environment*, 99(3), 304-314.

652 Song, J., & Wang, Z. H. (2015), Interfacing urban land-atmosphere through coupled urban
653 canopy and atmospheric models, *Boundary-Layer Meteorology*, 154(3), 427-448.

654 Song, J., & Wang, Z. H. (2016), Evaluating the impact of built environment characteristics on
655 urban boundary layer dynamics using an advanced stochastic approach, *Atmospheric*
656 *Chemistry and Physics*, 16, 6285-6301.

657 Song, J., Wang, Z. H., & Wang, C. (2017), Biospheric and anthropogenic contributors to
658 atmospheric CO₂ variability in a residential neighborhood of Phoenix, Arizona, *Journal of*
659 *Geophysical Research: Atmospheres*, 122, 3317-3329.

660 Sun, J., & Li, T. (2018) Relationship of lane width to speed for urban expressway: A case study
661 in Shanghai, In: *2018 International Conference on Network, Communication, Computer*
662 *Engineering (NCCE 2018)*, Atlantis Press, pp. 1-6.

663 Sun, T., Wang, Z. H., & Ni, G. (2013), Revisiting the hysteresis effect in surface energy budgets,

664 *Geophysical Research Letters*, 40, 1741-1747.

665 Theeuwes, N. E., Steeneveld, G. J., Ronda, R. J., Heusinkveld, B. G., van Hove, L. W. A., &
666 Holtslag, A. A. M. (2014), Seasonal dependence of the urban heat island on the street
667 canyon aspect ratio, *Quarterly Journal of the Royal Meteorological Society*, 140(684),
668 2197-2210.

669 Venter, Z. S., Chakraborty, T., & Lee, X. (2021), Crowdsourced air temperatures contrast
670 satellite measures of the urban heat island and its mechanisms, *Science Advances*, 7(22),
671 eabb9569.

672 Vermote, E. F., El Saleous, N. Z., & Justice, C. O. (2002), Atmospheric correction of MODIS
673 data in the visible to middle infrared: first results, *Remote Sensing of Environment*, 83(1),
674 97-111.

675 Voogt, J. A., & Oke, T. R. (2003), Thermal remote sensing of urban climates, *Remote Sensing of*
676 *Environment*, 86(3), 370-384.

677 Wan, Z., & Li, Z.-L. (1997), A physics-based algorithm for retrieving land-surface emissivity
678 and temperature from EOS/MODIS data, *IEEE Transactions on Geoscience and Remote*
679 *Sensing*, 35(4), 980-996.

680 Wang, C., Wang, Z. H., Kaloush, K. E., & Shacat, J. (2021a), Cool pavements for urban heat
681 island mitigation: A synthetic review, *Renewable & Sustainable Energy Reviews*, 146,
682 111171.

683 Wang, C., Wang, Z. H., Kaloush, K. E., & Shacat, J. (2021b), Perceptions of urban heat island
684 mitigation and implementation strategies: survey and gap analysis, *Sustainable Cities and*
685 *Society*, 66, 102687.

686 Wang, C., Wang, Z. H., Yang, J., & Li, Q. (2018), A backward-Lagrangian-stochastic footprint

687 model for the urban environment, *Boundary-Layer Meteorology*, 168(1), 59-80.

688 Wang, Z. H. (2014a), Monte Carlo simulations of radiative heat exchange in a street canyon with
689 trees, *Solar Energy*, 110, 704-713.

690 Wang, Z. H. (2014b), A new perspective of urban-rural differences: The impact of soil water
691 advection, *Urban Climate*, 10, 19-34.

692 Wang, Z. H. (2021), Compound environmental impact of urban mitigation strategies: Co-
693 benefits, trade-offs, and unintended consequence, *Sustainable Cities and Society*, 75,
694 103284.

695 Wang, Z. H. (2022), Reconceptualizing urban heat island: Beyond the urban-rural dichotomy,
696 *Sustainable Cities and Society*, 77, 103581.

697 Wolfe, R. E., Lin, G., Nishihama, M., Tewari, K. P., Tilton, J. C., & Isaacman, A. R. (2013),
698 Suomi NPP VIIRS prelaunch and on-orbit geometric calibration and characterization,
699 *Journal of Geophysical Research: Atmospheres*, 118(20), 11508-11521.

700 Wong, P. P. Y., Lai, P. C., Low, C. T., Chen, S., & Hart, M. (2016), The impact of
701 environmental and human factors on urban heat and microclimate variability, *Building and*
702 *Environment*, 95, 199-208.

703 Xu, T., & Liang, F. (2021). Machine learning for hydrologic sciences: An introductory
704 overview. *Wiley Interdisciplinary Reviews: Water*, 8(5), e1533.

705 Yang, J., Wang, Z. H., Kaloush, K., & Dylla, H. (2016), Effect of pavement thermal properties
706 on mitigating urban heat islands: A multi-scale modeling case study in Phoenix, *Building*
707 *and Environment*, 108, 110-121.

708 Yang, J., Wang, Z. H., & Kaloush, K. E. (2015), Environmental impacts of reflective materials:
709 Is high albedo a 'silver bullet' for mitigating urban heat island?, *Renewable and Sustainable*

710 *Energy Reviews*, 47, 830-843.

711 Zhang, Q., Katul, G. G., Oren, R., Daly, E., Manzoni, S., & Yang, D. W. (2015), The hysteresis
712 response of soil CO₂ concentration and soil respiration to soil temperature, *Journal of*
713 *Geophysical Research-Biogeosciences*, 120(8), 1605-1618.

714 Zhang, Y. F., Wang, X. P., Pan, Y. X., & Hu, R. (2013), Diurnal and seasonal variations of
715 surface albedo in a spring wheat field of arid lands of Northwestern China, *International*
716 *Journal of Biometeorology*, 57(1), 67-73.

717 Zhang, Y., Song, C., Band, L. E., Sun, G., & Li, J. (2017), Reanalysis of global terrestrial
718 vegetation trends from MODIS products: Browning or greening?, *Remote Sensing of*
719 *Environment*, 191, 145-155.

720 Zhao, S. Q., Liu, S. G., & Zhou, D. C. (2016), Prevalent vegetation growth enhancement in
721 urban environment, *Proceedings of the National Academy of Sciences of the United States*
722 *of America*, 113(22), 6313-6318.