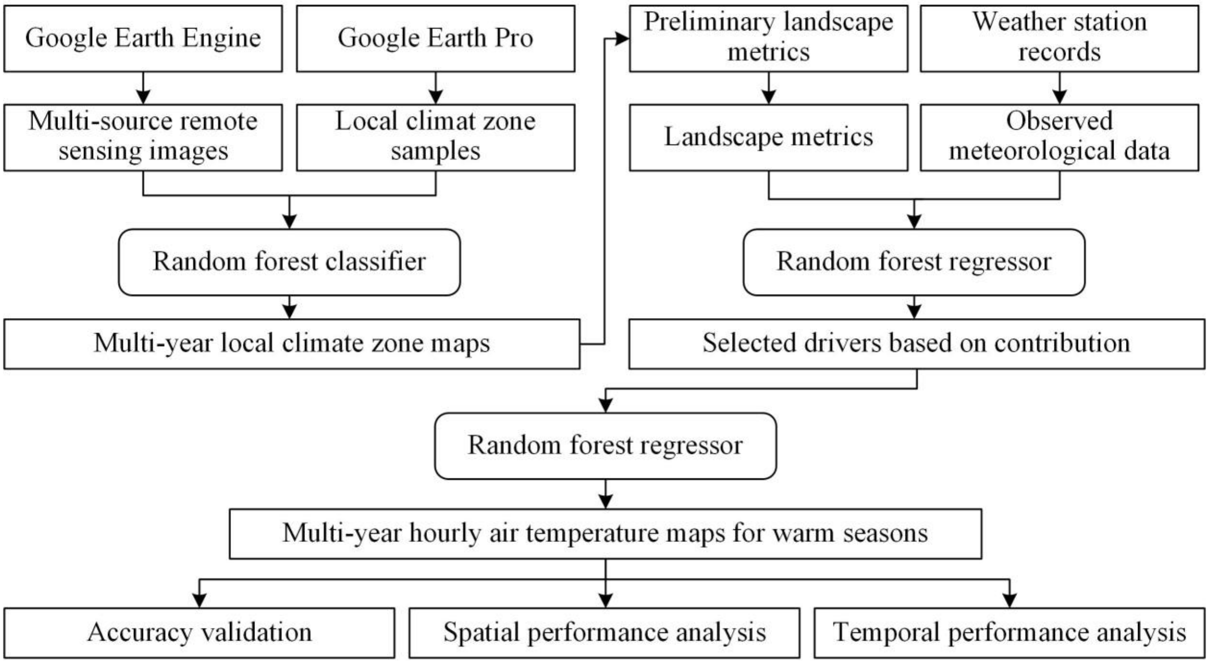
Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning

Guangzhao Chen[1,4], Yuan Shi[2], Ran Wang[3], Chao Ren[4,*], Edward Ng[5], Xiaoyi Fang[6], Zhihua Ren[7]

1. Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong, SAR, China
2. Department of Geography & Planning, University of Liverpool, Liverpool, UK
3. Nankai University
4. Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China
5. School of Architecture, The Chinese University of Hong Kong, Hong Kong, SAR, China
6. Chinese Academy of Meteorological Sciences, China
7. National Meteorological Information Center, China

*corresponding author: Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China (renchao@hku.hk)

Graphical Abstract



Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning

Highlight

- We performed hourly air temperature mapping with 1-km resolution across multi-year warm seasons using local-climate-zone-based landscape metrics and random forest algorithms.
- Nighttime results: Analysis revealed that the maps steadily maintained high accuracy at nighttime (20:00–7:00), which is important to investigate the nighttime urban climate conditions, especially the urban heat island effect.
- Spatial pattern of the air temperature estimations exhibited a pronounced landscape divide that air temperatures in contiguous mountainous areas with dense trees were significantly lower than those in the plains.
- Air temperatures tend to fall more slowly in the core of metropolitan areas than in the urban fringe.

Keywords

Machine learning; Hourly air temperature mapping; High spatial resolution; Local climate zone,

 Click here to view linked References

Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning

Guangzhao Chen[1,4], Yuan Shi[2], Ran Wang[3], Chao Ren[4]*, Edward Ng[5], Xiaoyi Fang[6], Zhihua Ren[7]

1. Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong, SAR, China
2. Department of Geography & Planning, University of Liverpool, Liverpool, UK
3. Nankai University
4. Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China
5. School of Architecture, The Chinese University of Hong Kong, Hong Kong, SAR, China
6. Chinese Academy of Meteorological Sciences, China
7. National Meteorological Information Center, China

*corresponding author: Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China (renchao@hku.hk)

Abstract

Air temperature is a crucial variable of urban meteorology and is essential to many urban environments, urban climate and climate-change-related studies. However, due to the limited observational records of air temperature and the complex urban morphology and environment, it might not be easy to map the hourly air temperature with a fine resolution at the surface level within and around cities via conventional methods. Thus, this study employed machine learning (ML) algorithms and meteorological and landscape data to develop hourly air temperature mapping techniques and methods at the 1-km resolution over a multi-year warm seasons period. Guangdong Province, China was selected for the case study. Random forest algorithm was employed for the hourly air temperature mapping. The validation results showed that the hourly air temperature maps exhibit good accuracy from 2008 to 2019, with mean R^2, root mean square error (RMSE) and mean absolute error (MAE) values of 0.8001, 1.4821℃ and 1.0872℃, respectively. The importance assessment of the driving factors showed that meteorological factors, especially relative humidity, contributed the most to the air temperature mapping. Simultaneously, landscape factors also played a non-negligible role. Further analysis revealed that the maps steadily maintained high accuracy at nighttime (20:00–7:00), which is essential for investigating nighttime urban climate conditions, especially the urban heat island effect. Moreover, a correlation existed between the nighttime air temperature changes and urban morphology represented by the local climate zones. Air temperatures tended to fall more slowly in the core of metropolitan areas than in the urban fringe. Using ML, this study reliably improves the spatial refinement of hourly air temperature mapping and reveals the spatially explicit air temperature patterns in and around cities at different times in a day during the warm seasons. Moreover, it provides a novel valuable and reliable dataset for air-temperature-related implementation and studies.

Highlight

- We performed hourly air temperature mapping with 1-km resolution across multi-year warm seasons using local-climate-zone-based landscape metrics and random forest algorithms.
- Nighttime results: Analysis revealed that the maps steadily maintained high accuracy at

nighttime (20:00–7:00), which is important to investigate the nighttime urban climate conditions, especially the urban heat island effect.

- Spatial pattern of the air temperature estimations exhibited a pronounced landscape divide that air temperatures in contiguous mountainous areas with dense trees were significantly lower than those in the plains.
- Air temperatures tend to fall more slowly in the core of metropolitan areas than in the urban fringe.

Keywords

Machine learning; Hourly air temperature mapping; High spatial resolution; Local climate zone,

1. Introduction

The rise of mega- and high-density compact urban regions is now an irreversible trend of urbanization[1]. Such high-density mega-urban living has caused numerous environmental challenges and problems, such as intensified urban heat islands (UHIs)[2-4] and air pollution[5,6]. Simultaneously, the complex urban morphology poses a great challenge in depicting the near-surface air temperature within and around cities.

Characterising the spatiotemporal variability of the near-surface air temperature at fine resolutions is of importance for investigating the UHI intensity and heat-related risks[7]. It is becoming even more important in the context of climate change. Specifically, human activities are predicted to have caused about 1.0 ℃ of global warming, compared to the pre-industrial period[8]. Moreover, it is projected that without a significant reduction in greenhouse gas emissions, the global near-surface temperature will continue to increase with an increasing number of extreme weather events, like extreme heat waves[8].

Air temperature (*Ta*) is a key variable in the investigation of climate change[8], energy consumption[9], thermal comfort[10] and human health[11]. *Ta* has been widely employed in the fields of epidemiology and public health to explore its relation to morbidity and mortality in vulnerable populations[12,13]. An accurate and in-depth understanding of *Ta* will help scientists conduct subsequent research applications in various fields to provide scientific-evidence-based findings for policymakers to achieve sustainable development. However, this subject is still under-researched in many regions worldwide due to technical limitations.

1.1. Literature review

Typically, meteorological stations measure *Ta* at a reference height of 2 m above the ground[14]. Meteorological stations usually keep a long-term archive of observational weather data. However, their ability to capture the spatial variation of *Ta*, particularly in heterogeneous areas, is limited due to their limited spatial coverage[15]. Specifically, meteorological stations provide long-term observational weather data at fine temporal resolution. However, due to the lack of adequate spatial coverage, their ability to depict small-scale spatial variability in heterogeneous regions (including cities) is limited. Therefore, data from meteorological networks are not often sufficient for studying the impact of extreme hot weather on heat-related health risks, as the air temperature may greatly vary with space and time. To address this issue, statistical methods are applied to map the spatiotemporal pattern of *Ta* based on limited meteorological stations. These methods can be divided into two groups: (1) Spatial interpolation methods, e.g. inverse distance weighting (IDW)[16], Kriging interpolation[17] and geographic weighted regression[16]. These interpolation methods are employed to predict *Ta* in an area surrounding a known meteorological station at a fixed time. A prerequisite of these methods is a relatively homogenous distribution of weather stations, but a study area may have a highly heterogeneous distribution of weather stations. (2) Regression methods that can predict *Ta* at any location and time by establishing a quantitative relation between *Ta* and possible influencing factors. These methods include linear regression with simple or multiple variants[18,19] and nonlinear regression, including machine learning (ML) methods[20]. Through training and testing with considerable input data, ML models learn how to estimate *Ta* with optimal accuracy, even in areas with highly heterogeneous landscape patterns.

Climate model simulation is another choice for mapping the spatiotemporal pattern of *Ta* across different scales, from global, regional, to city scales. Global or regional climate models yield *Ta* with low spatial resolution (approximately 100–250 and 25–50 km) and high temporal resolution (e.g. hourly or minute). Both kinds of climate models provide rough descriptions of climate variables since the urban structure and its influence on climate are both simplified in the model setup and simulation[21]. Mesoscale models, such as the weather research and forecast model, have been developed with additional urban information to simulate climates at the local scale (1–5 km)[21]. However, simulation of *Ta* using mesoscale climate models is timeconsuming and relies on the computation power of the hardware. Furthermore, *Ta* generated via mesoscale models can still not assist in the spatiotemporal pattern analysis of a thermal environment at the district/block scale (e.g. hundreds or tens of meters). Microscale climate models, like ENVI-met, have been further developed for simulating microscale urban climates[22]. Unfortunately, despite the fine spatial and temporal resolution, the spatiotemporal pattern of *Ta* across the entire city is hard to simulate using microscale climate models due to the high time cost and limited computing ability of the model. Generally, the simulation of *Ta* using various climate models is limited by the lack of historical input data, long simulating time, high learning cost, complicated model setup and simulation, a balance between spatial coverage and spatial/temporal resolution and computational power.

Remotely sensed data have the advantage of broad spatial coverage and various spatial and temporal resolutions; the land surface temperature (LST) retrieved from remote sensing images is the most commonly used satellite predictor for mapping the spatiotemporal variation of *Ta*[23]. LST-based *Ta* estimation is mainly achieved via the following ways: (1) Temperature-vegetation index method. This method assumes that the LST of vegetation is similar to its surrounding *Ta*. Hence, the spatial pattern of *Ta* can be interpolated based on the relation between LST and vegetation[24,25]. However, such a method is unsuitable for urban areas, which are mostly covered by unvegetated surfaces[26]. (2) Energy balance model. Both LST and *Ta* are important components of energy fluxes in the energy balance model, i.e. both are essential for calculating the longwave radiation and sensible heat flux[27]. *Ta* can be retrieved by analysing the energy exchanges within an urban canopy layer using LST[28]. This method requires input data that are not measured by satellite sensors and needs prior knowledge to construct energy balance models. (3) Statistical methods. Linear and nonlinear regression models have been implemented for building a relation between *Ta* and LST as well as other auxiliary data, like the land cover, daylight duration and evapotranspiration[29-33]. However, such a *Ta*–LST relation is sensitive to location, background climate and the presence of daylight[34,35].

1.2. Research gaps

The spatiotemporal changes of air temperature at a micro to local climate scale can be largely affected by the landscape pattern of land use/land cover (LU/LC) because the land surface changes the boundary layer climate conditions[36]. The abovementioned *Ta* estimation methods have their own strengths and limitations in the investigation of the spatiotemporal changes in *Ta*. Therefore, developing a time-series *Ta* dataset with both high spatial and temporal resolutions still needs to be explored, especially when focusing on the intra-urban variation of the thermal environment. Most of the existing research is concerned with the daily air temperature characteristics. Furthermore, most resulting spatiotemporal temperature models are usually site-specific.

The local climate zone (LCZ) classification scheme not only enables the investigation of a fine-scale intra-urban variation of *Ta* but also increases the transferability of the resultant models, due to its ten built types classified using building morphology parameters (e.g. building height, building coverage ratio and sky view factor)[37]. Furthermore, the relevance of LCZs to the urban thermal environment has been argued in literature[38-40]. Moreover, information from remote sensing imagery, such as Normalized Difference Vegetation Index (NDVI) and multispectral albedos, has been used as input variables for generating LCZ maps. Hence, landscape patterns of LCZ classes, which can be represented by landscape metrics of LCZs, will help refine the spatial variation of *Ta*, particularly in a complex urban context. As mentioned earlier, ML algorithms exhibit good performance in estimating *Ta* across the city scale because of their strong learning ability from a large number of trials.

1.3. Study objectives

Herein, we aim to estimate the spatiotemporal hourly resolved air temperature on a 1 km grid across the study area (Guangdong province in China as the testbed) by incorporating LCZ-based landscape patterns as predictors, refining both the temporal and spatial coverage. Specifically, this study combines the LCZ-based landscape patterns with an ML method to predict the *Ta* distribution in a highly urbanised region with complex urban morphology, the Guangdong province. The study objectives include (1) developing a 12-year (2008–2019) spatiotemporal distribution map of *Ta* at an hourly resolution and 1-km grid across the Guangdong province, (2) generating averaged hourly *Ta* maps during warm seasons (May–September) for each year and (3) identifying different *Ta* patterns during the nighttime and daytime in urban and rural areas to facilitate an understanding of the spatiotemporal variability of *Ta*.

2. Materials and method

2.1. Study area and time period

Guangdong province is located in the southernmost part of mainland China and faces the South China Sea to the south. The east and west sides of the Pearl River Estuary in the Pearl River Delta region of Guangdong Province are bordered by Hong Kong and Macao Special Administrative Regions, respectively. Additionally, it is a subtropical region with high spatial heterogeneity of LU/LC. The terrain of is high in the north and low in the south and is complex and diverse, including mountains, hills, plains and mesas. Its geographic complexity makes it a suitable study area for testing the applicability of ML algorithms in predicting air temperature with high spatial and temporal resolution.

2.2. Meteorological data

As a part of the national meteorological stations network of China, 86 national meteorological monitoring stations are located and are operational in the Guangdong province (Fig. 1). All stations are operated by the China Meteorological Administration (CMA). The siting, equipment set up and operation strictly follow the World Meteorological Organization (WMO) guidelines[41]. Hourly air temperature has been continuously recorded and managed by CMA data centre (https://data.cma.cn/en) as a dataset, which is ready for scientific and academic use. Herein, to facilitate the development of ML-based prediction models, the hourly air temperature data from 2008 to 2019 was requested from the CMA. The data are quality controlled by the CMA. The observed data missing rate is less than 1%.

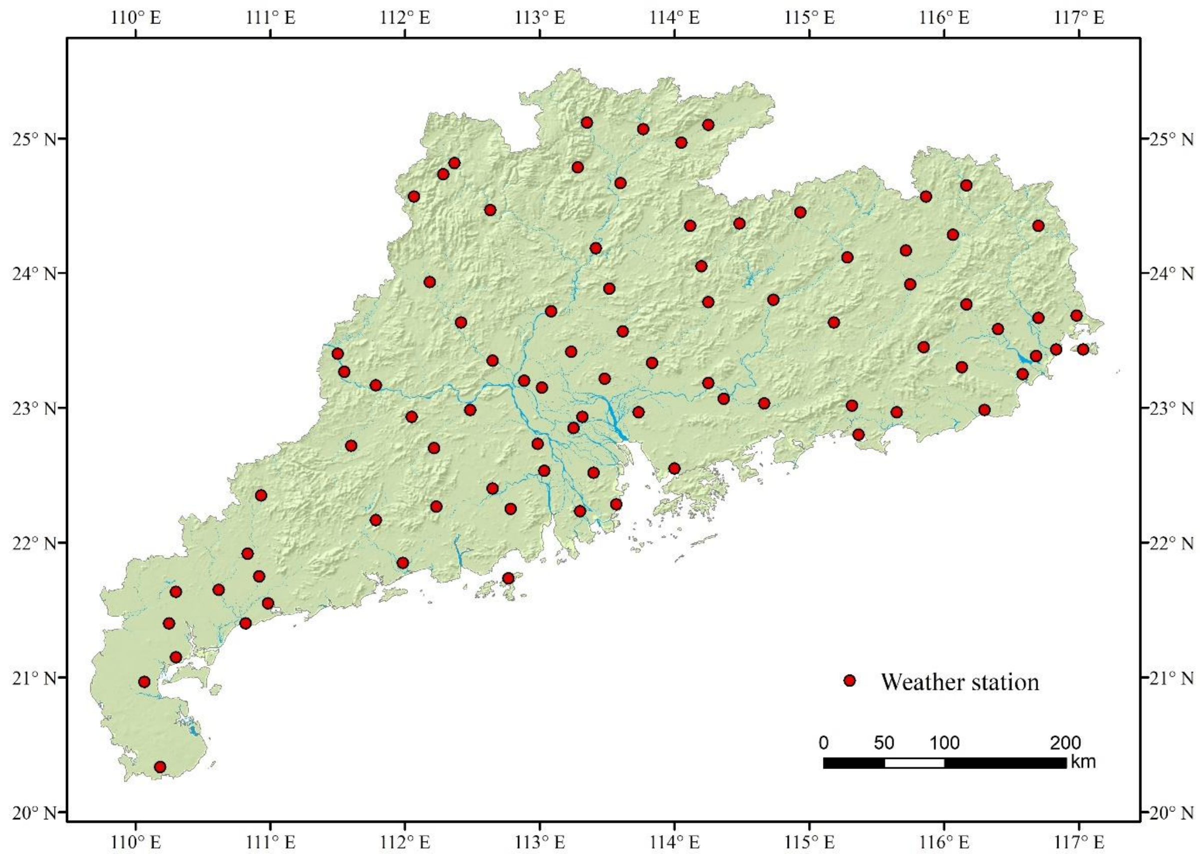


Fig. 1 Weather stations in the study area

In addition to the air temperature, the observed data include meteorological variables such as relative humidity (RHU), precipitation (PRE), barometric pressure (PRS) and wind speed (VV2). These variables are used as meteorological drivers in the subsequent spatial estimation modelling of the hourly air temperature. Furthermore, the geographical coordinates and elevation information of the weather stations are provided.

To drive a well-trained Random Forest (RF) model for spatially estimating the air temperature, the spatial pattern of these meteorological drivers across the study area needs to be obtained. Hence, we performed the Kriging interpolation using the observed data to estimate the spatial patterns.

2.3. LCZ data and Landscape pattern analysis

2.3.1. LCZ mapping

Previous studies have demonstrated that the physical foundations of cities, including building form and building materials, can influence the spatial variations in the air temperature[42,43]. As a widely used land surface classification scheme that defines the land cover types based on the physical characteristics of the land surface (Table 1), LCZ has unique advantages over traditional land cover classifications in depicting landscapes, especially landscapes within cities[37,44,45]. Based on LCZ, urban and natural landscapes have been classified into 18 types. We generated correspondingly categorical maps for each year in the study period (2008–2019), which well represent the landscape diversity and geographic complexity as well as the temporal changes of LU/LC in the study area. Noted that the 2012 LCZ map was not generated due to the quality deficiencies of the 2012 remote sensing images. The LCZ map development process can be divided into three steps: (1) creating a multi-year LCZ sample set, (2) preparing the input data on the Google Earth Engine (GEE) platform and (3) conducting LCZ classification on the GEE platform using an RF classifier, as performed in Chung et al.[46].

Table 1 Categories and definitions of local climate zone (LCZ) simplified from Stewart & Oke[37]

LCZ types	**Built and land cover types**
LCZ 1	Compact high-rise
LCZ 2	Compact mid-rise
LCZ 3	Compact low-rise
LCZ 4	Open high-rise

LCZ 5	Open mid-rise
LCZ 6	Open low-rise
LCZ 7	Lightweight low-rise
LCZ 8	Large low-rise
LCZ 9	Sparsely built
LCZ 10	Heavy industry
LCZ A	Dense trees
LCZ B	Scattered trees
LCZ C	Bush, scrub
LCZ D	Low plants
LCZ E	Bare rock or paved
LCZ F	Bare soil or sand
LCZ G	Water
LCZ H	Wetlands[#]

[#] Wetlands is an additional LCZ type that adapted the land surface properties of coastal cities in the Guangdong province.

First, we selected 2165 LCZ sample polygons through Google Earth Pro based on fine-resolution remote sensing images of 2019, which comprise more than 100 sample polygons per LCZ type. Then, using the historical images provided by Google Earth Pro, we modified the labels of these samples in different years to construct a year-by-year sample set from 2008 to 2019. In the LCZ classification of each year, 70% of the 2165 samples were randomly selected for classifier training, while the remaining 30% were used for accuracy validation.

Second, we selected suitable multi-year images from the multi-source remote sensing images provided by the GEE platform and clipped them to the Guangdong province extent. Data from Landsat 8 (Landsat 8 Surface Reflectance Tier 2), Landsat 5 (Landsat 5 Surface Reflectance Tier 2), Sentinel-1 SAR GRD (C-band Synthetic Aperture Radar Ground Range Detected, log scaling), Sentinel-2 MSI (Multi-Spectral Instrument, Level-1C), VIIRS (Stray Light Corrected Nighttime Day/Night Band Composites Version 1) and DMSP OLS (Nighttime Lights Time Series Version 4) were selected as input data for multi-year LCZ classification since they cover different spectral and nighttime light information. Furthermore, GMTED2010 (Global Multi-resolution Terrain Elevation Data 2010) ware chosen as the input data to provide elevation information. Table S1 provides the descriptions of these input data.

Third, we performed year-by-year LCZ classification by applying the RF classifier provided by the GEE platform using training samples and multi-source remote sensing images as the input data. RF is an ensemble ML algorithm that estimates or classifies objectives by constructing multiple decision trees and aggregating their decision results based on votes[47]. It is a nonlinear algorithm that balances accuracy and computational efficiency and performs stably because errors in a single decision tree are unlikely to affect the voting results[48,49]. Therefore, RF is widely used in land classification based on remote sensing images. Herein, we employed the '.smileRandomForest' package from the GEE platform to perform LCZ classification. We kept the default parameter settings of the package except for the number of trees (i.e. n-tree). We searched for the optimal n-tree from 20 to 120 at 10-tree intervals based on the validation accuracy and finally set n-tree as 80.

2.3.2. LCZ-based landscape pattern

Most previous studies on the spatial estimation of air temperature have usually investigated the LU/LC and landscape types at the exact location of the weather stations[50,51]. Few studies have analysed how the spatial configuration, such as the mixture, evenness, diversity, clustering of different LU/LC and landscape types, affects the variability in the spatiotemporal distribution of air temperature. Herein, based on the generated LCZ maps, we introduced highly quantifiable measures, landscape metrics, to quantify the LU/LC pattern of the study area. Landscape metrics are developed based on the classic ‘patch-corridor-matrix’ theory in the landscape ecology[52]. Corresponding to the above landscape theory, landscape metrics can be divided into three main categories: patch-, class- and landscape-level metrics. Patch-level metrics represent the characteristics of a single patch of a specific type of landscape or LCZ class. Class-level metrics reflect the spatial pattern of all patches with the same LCZ class within a certain spatial extent, while landscape-level metrics provide an understanding of how different LCZ classes spatially mix together. Landscape metrics have been widely used to categorically analyse remote-sensed spatial datasets for two decades[53]. Herein, based on literature[54,55], a set of landscape metrics with radiuses ranging from 1 to 10 km were chosen as candidate predictor variables (Table S2) to quantify the detailed spatial pattern around each of the weather stations and the spatial pattern in the entire study area. Fragstats (program version 4), a widely used software[56], was employed to determine the landscape metrics on the basis of the LCZ categorical map for each year in the study period. Using the above process, a large predictor dataset (with an extensive amount of landscape pattern metrics of 13550 variables, as there are 18 classes of LCZ types reflecting the various landscape in the study area) has been generated. However, to reduce the computational burden on the model, only landscape metrics with more than 80% of the valid values in the sample were included as the preliminary drivers for subsequent modelling. Figures S1~S3 present the patterns of the three landscape metrics with the highest contribution to the model, based on the subsequent importance assessment of the drivers.

2.4. Estimating hourly air temperature spatial patterns using the random forest model

The previously prepared meteorological and landscape drivers were input into the RF model to estimate the spatial hourly air temperature patterns. We selected the RF model as the regressor because it not only has the abovementioned advantages but also allows the importance assessment of each driver to the estimation accuracy[57], which is essential for this study. To estimate the air temperature at a certain hour, we considered real-time-efficient drivers like the current time (hour), meteorological drivers for each of the previous 24 h and environmental drivers like the landscape drivers, longitude, latitude and elevation, yielding a total of 941 preliminary drivers in the RF model. The driving factors need to be considered as comprehensively as possible, but this will increase the computational burden of the model and significantly increase the operation time. Moreover, most of the drivers contribute little to improving accuracy. Therefore, we first built an RF model using the 2019 data to select critical drivers from the 941 preliminary drivers based on the importance assessment. Simultaneously, we tested the optimal n-tree for the RF model. Finally, we identified key drivers and adopted the optimal n-tree for building the RF model for other years.

In Python, we used the '.RandomForestRegressor' class provided by the 'scikit-learn' extension package (Version 0.24.2) to build the RF model. The default values are employed for all parameters except the n-tree. Additionally, we employed the permutation importance provided by scikit-learn as the metric to assess the importance of the drivers as it is applicable in cases where there are many unique values of the features. The permutation importance of a feature is defined as the deviation of the metric value from the baseline metric value after permutation of this feature column. We performed ten evaluations of the permutation importance of the drivers and took their average value as the importance of the drivers.

To build the RF model, 70% of the samples were randomly selected for training the model. We used four accuracy metrics to measure the model accuracy. One is to calculate the goodness-of-fit, R^2, of the trained model using the remaining 30% samples. The second is to estimate the R^2 of the model using the out-of-bag samples (oob_score) during model training. Further, the root mean square error (RMSE) and mean absolute error (MAE) were calculated using the test samples to evaluate the model's bias. These four metrics provide a comprehensive picture of the model's generalisation ability.

3. Result

3.1. Accuracy of the LCZ mapping

Table S3 presents the assessment table for LCZ mapping in the study area from 2008 to 2019. Moreover, we used user accuracy (UA) and producer accuracy (PA) to assess the performance of each LCZ type and used the overall accuracy (OA) and Kappa coefficient to measure the overall performance of LCZ maps for each year. The results showed that the average value of the OA of the LCZ maps reached 61.64% and that of the Kappa coefficient reached 0.594; the best performance was observed in 2019, where OA and the Kappa coefficient reached 71.86% and 0.702, respectively. According to Bechtel et al.[44], the accuracy of our LCZ maps is comparable to that of most current LCZ mapping and is therefore acceptable.

3.2. Accuracy of the hourly air temperature estimation

We selected 90 drivers from the 941 preliminary drivers for subsequent model training and estimation with the permutation importance. The sum of the importance scores of the 90 drivers (1.772) represents 97.0% of the total importance score of all the preliminary drivers (1.826). Therefore, the selected drivers are sufficiently representative. Among the 90 drivers, 74 meteorological and 12 landscape drivers are present, and current time, latitude, longitude and elevation drivers are also present. The five most important drivers are RHU_1Hours_Before (0.725), Current_time (0.471), mw09_shdi (Shannon's Diversity Index at a radius of 9 km, 0.072), latitude (0.072) and RHU_10Hours_Before (0.065).

We performed tests to search for the optimal n-tree from 50 to 400. The results showed that the R^2 calculated using the out-of-bag samples (oob_score) logarithmically grew with increasing n-tree value (Fig. 2). Furthermore, the oob_score significantly improved with increasing n-tree increased from 50 to 200. With increasing n-tree from 200 to 400, the oob_score still displayed a slight improvement. Therefore, for better accuracy, we set the n-tree value in the RF modelling to 400.

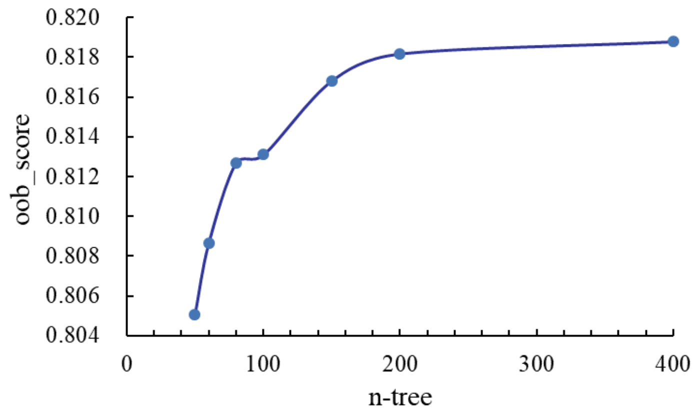


Fig. 2 Relation between the number of trees (n-tree) in the RF modelling and R2 calculated using the out-of-bag samples (oob_score)

After determining the drivers and n-tree value, we executed the RF modelling for each year. Table 2 shows the performance of the RF models for each year. The R^2, RMSE and MAE calculated using the 30% validation samples and the oob_score calculated using the out-of-bag samples exhibit similar accuracies. The RF models exhibited good accuracy in different years, with the mean values of R^2 and oob_score reaching 0.8001 and 0.7960, respectively. Additionally, the mean values of RMSE and MAE were 1.4821℃ and 1.0872℃, respectively. The results indicate that the RF models constructed to estimate the hourly air temperatures from 2008 to 2019 are acceptable and reliable.

Table 2 Accuracy of the RF models for each year

Year	**R^2**	**oob_score**	**RMSE (℃)**	**MAE(℃)**
2008	0.8036	0.7992	1.5112	1.0940
2009	0.8127	0.8084	1.4592	1.0879
2010	0.7684	0.7652	1.6049	1.1953
2011	0.8202	0.8132	1.5062	1.0951
2013	0.7685	0.7648	1.5498	1.1163
2014	0.8272	0.8252	1.4336	1.0323

2015	0.8197	0.8153	1.3368	0.9886
2016	0.7725	0.7697	1.5097	1.1133
2017	0.7810	0.7762	1.5145	1.1259
2018	0.8041	0.7997	1.3607	1.0002
2019	0.8234	0.8188	1.5165	1.1100
Mean	0.8001	0.7960	1.4821	1.0872

3.3. Performance of the air temperature estimation in various hours

Furthermore, we explored the performance of the estimated air temperature at different hours. We merged the temperatures for all dates at a particular hour and assessed the model performance for that hour by comparing the observed and estimated mean air temperature. We selected three metrics to measure the hourly model performance: R^2, RMSE and a deviation ratio.

Figure 3 shows the R^2 of the models for different hours in different years. Figures S4–S14 display the scatter plots of the estimated versus observed values for different hours in different years. Clearly, the performances show consistency and stability across the years. For example, the models maintained stable high R^2 during the nighttime (i.e. 20:00–07:00), while during the daytime hours, the models did not perform well overall, except for the period from 14:00 to 16:00 when they reached a high R^2 level. Note that herein we directly calculated R^2 using the estimated and observed air temperatures, rather than calculating R^2 after fitting a linear regression to them; thus, R^2 affords a maximum value of 1 and it could be negative. However, when R^2 is negative, the estimated and observed air temperatures may still exhibit a good linear relation, as shown in Figure S4–S14.

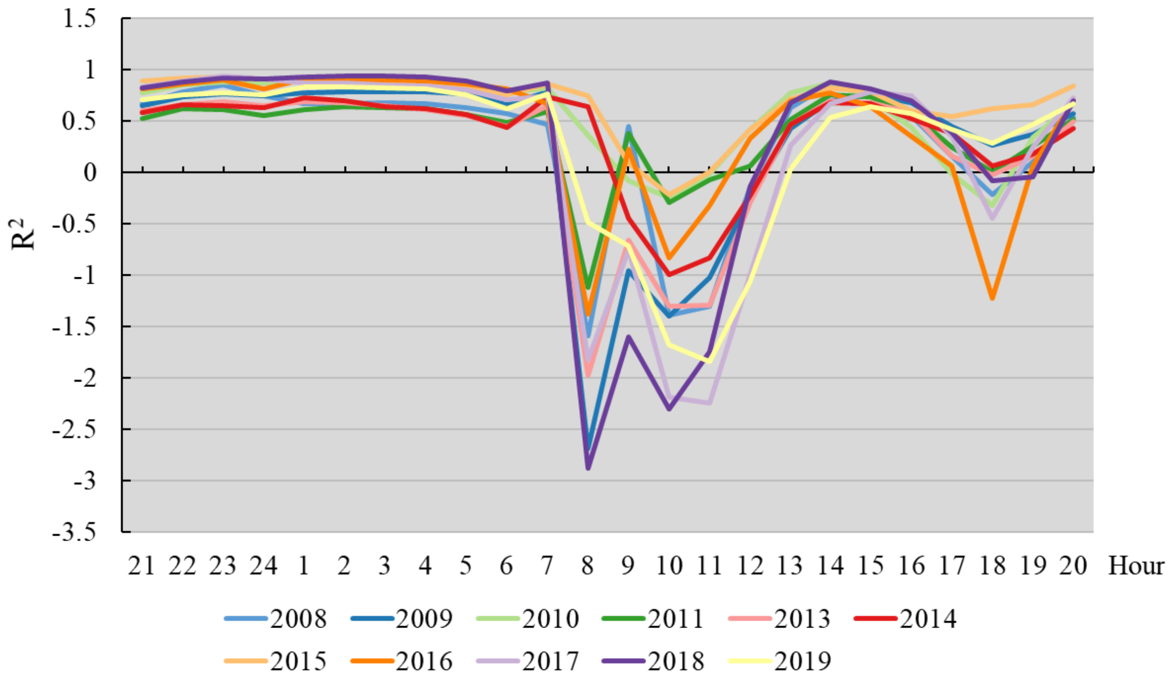


Fig. 3 R^2 of the air temperature estimation models for different hours in different years

RMSE is a metric reflecting the absolute error between the estimated and observed values; thus, a smaller RMSE value denotes a higher estimation accuracy. Fig. 4 presents the RMSE of the air temperature estimation models for different hours in different years. The RMSE distribution is similar to the R^2 distribution. The performance of the models for the same hour was essentially stable across the years. Better RMSE performance was obtained from 20:00 to 07:00 at night and for a short period in the afternoon. Larger RMSE values were afforded in the morning (8:00–11:00), but the RMSE values slightly increased in the late afternoon (around 18:00). During the periods when the model performed well, RMSE did not exceed 0.6°C overall, even reaching 0.2°C.

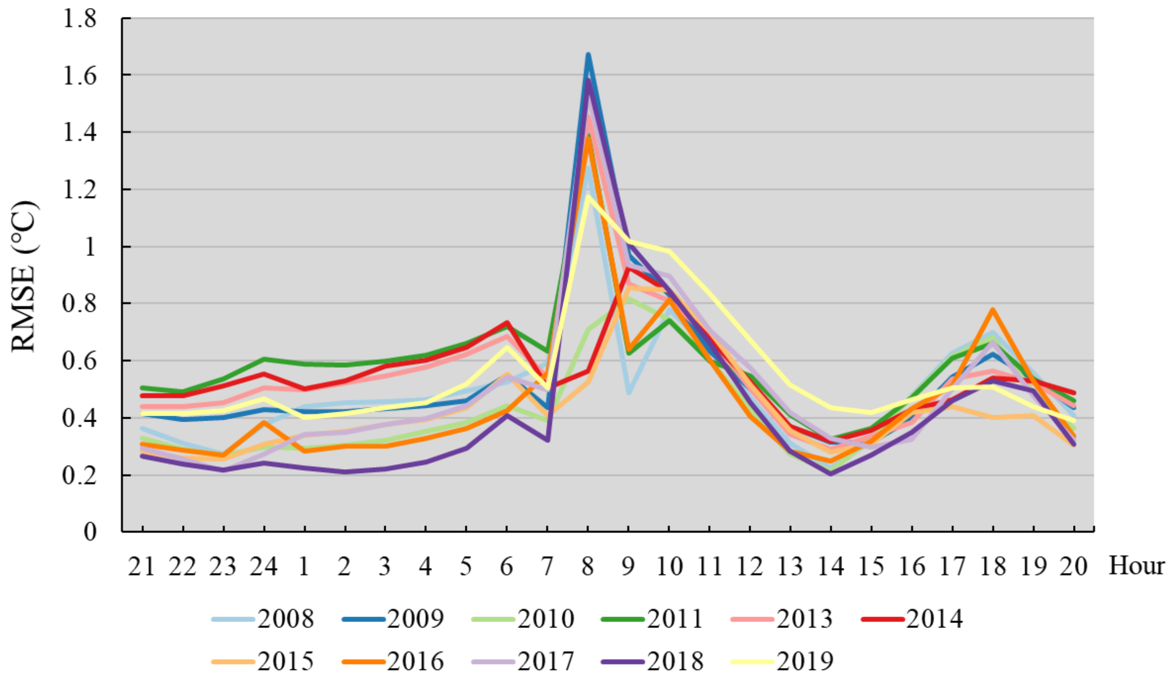


Fig. 4 RMSE of the air temperature estimation models for different hours in different years

In addition to using RMSE to measure the absolute error of the estimation results, we defined a deviation ratio to reflect the relative error of the estimation results. The deviation ratio is the ratio of RMSE to the difference between the air temperatures observed at the middle 50% of the weather stations. Fig. 5 shows the deviation ratio of the estimated air temperatures for different hours in different years. Notably, the trajectory of the deviation ratio is similar to that of the R^2 and RMSE. In most years, the deviation ratios were generally below 0.5 and even below 0.2 during the night (20:00–07:00) and afternoon (14:00–16:00).

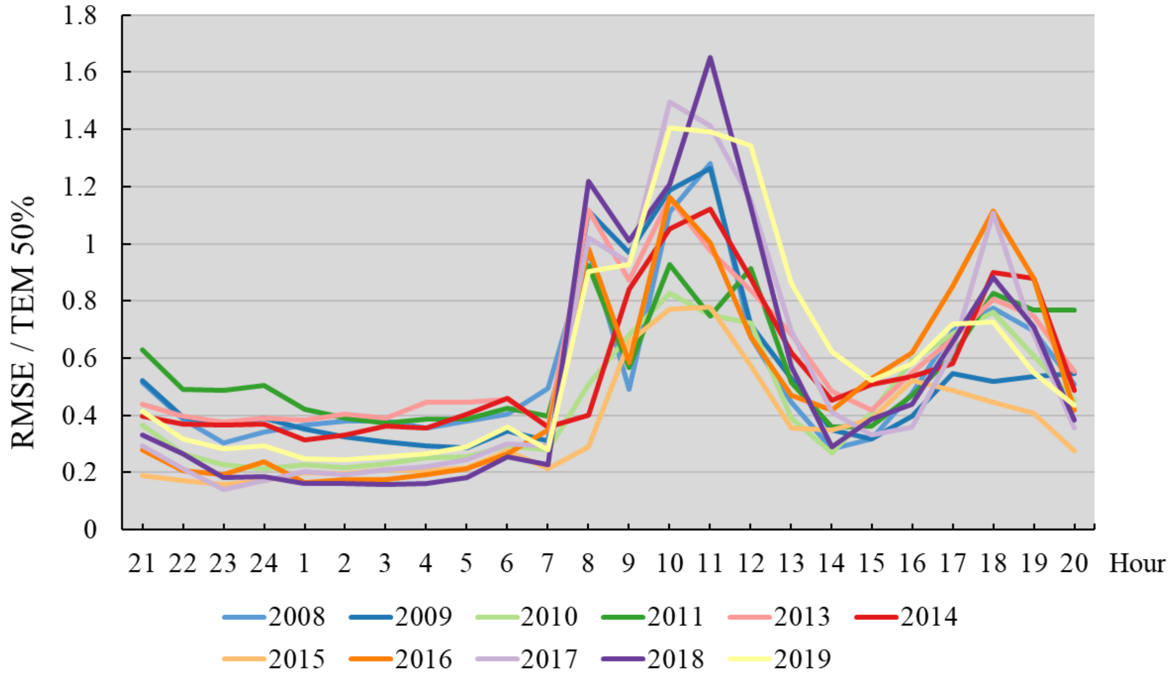
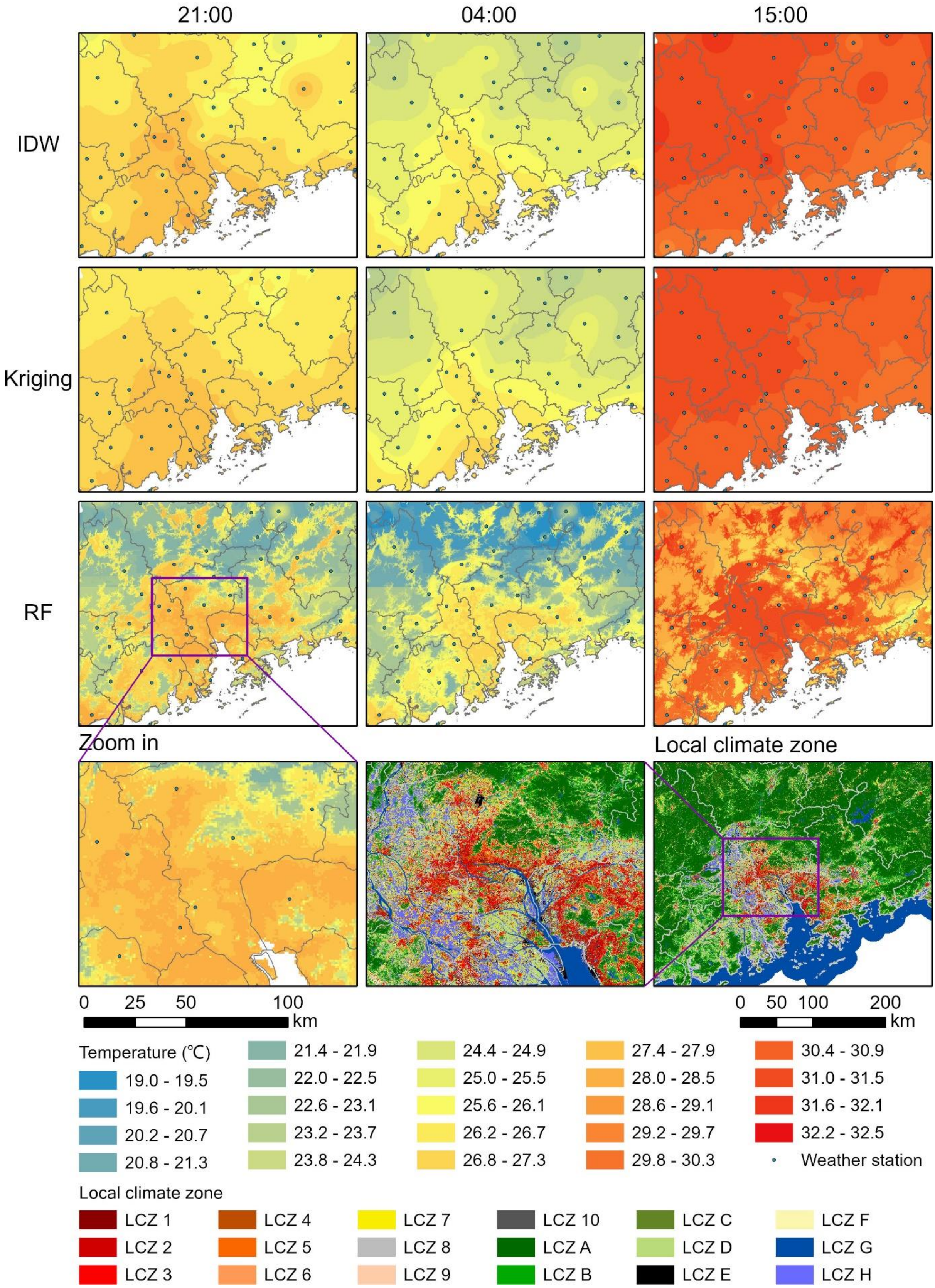


Fig. 5 Deviation ratio of the estimated hourly air temperatures in different years. Here, 'TEM 50%' denotes the difference between the air temperatures observed at the middle 50% of the weather stations.

3.4. Spatial performance of the air temperature estimation

In addition to the overall and temporal perspective, we explored how the models performed in space. We compared the performance of the RF model with traditional spatial interpolation methods, such as IDW and Kriging interpolation, for estimating the spatial distribution of the air temperature. Fig. 6 shows the comparison result for the mean air temperature in warm seasons in 2019. Clearly, the air temperature distribution estimated by the RF model was generally consistent with that estimated by IDW and Kriging interpolation. Although humidity strongly contributes to the predictions of the air temperature distribution, landscape metrics add considerable spatial detail to the air temperature distribution mapping, which cannot be obtained by directly interpolating air temperature using almost any other methods. Moreover, the difference in air temperature between urban and rural areas was more evident in the results of the RF model than in the those of IDW and Kriging interpolation. Rural areas cooled faster than the urban areas at night. Moreover, comparing the air temperature distribution at 21:00 and 04:00, the temperature dropped more slowly in the urban core than in the urban fringe.

394



395 Fig. 6 Comparing IDW and Kriging interpolation with the RF model in terms of the spatial performance

of the air temperature estimation for the mean air temperature in warm seasons in 2019.

Furthermore, to demonstrate the role of landscape drivers in enhancing the spatial detail of the air temperature estimation, we added a control experiment without LCZ-based landscape drivers in the modelling. Fig. 7 shows the role of LCZ-based landscape drivers in the air temperature estimation, taking the example of 21:00 in the 2019 warm season. When modelling without the LCZ-based landscape drivers (Fig. 7(a)), elevation enhanced the spatial detail by making the air temperatures cooler in mountainous places and hotter at lower elevations near the sea. However, the effect of urban morphology on the air temperature distribution could not be reflected. When the LCZ-based landscape drivers were considered (Fig. 7(b)), the effect of urban agglomerations on the air temperature distribution was revealed.

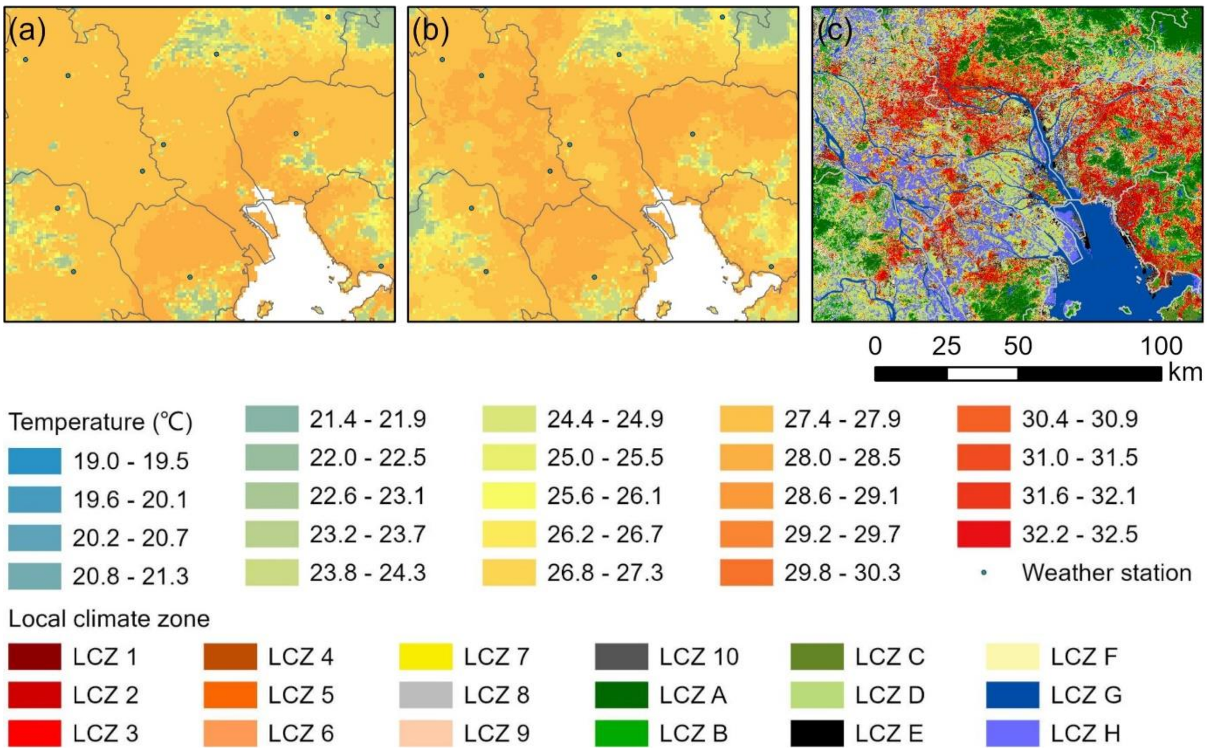


Fig. 7 Comparing the impact of modelling with and without LCZ-based landscape drivers on the spatial detail of the air temperature estimation. (a) Mean air temperature at 21:00 for the 2019 warm season modelling without LCZ-based landscape drivers; (b) Mean air temperature at 21:00 for the 2019 warm season modelling

with LCZ-based landscape drivers; (c) LCZ in 2019.

Since the RF model spatially demonstrated the difference in air temperature between urban and rural areas, we analysed the difference in the RF model performance for estimating urban and rural air temperatures. Therefore, we first selected urban and rural stations from the 86 weather stations. To exclude changes in the station types due to urbanisation, we counted the major land types around a station within a radius of 500 m. LCZs 1–10 are urban and LCZs A–H are rural. If more than 50% of the land around a station was urban LCZs, it was denoted as an urban station; otherwise, it was denoted as a rural station. Ultimately, only the stations whose station type remained constant throughout 2008–2009 were included in the subsequent urban–rural analysis. Consistent with Section 3.3, we selected the gaps in R^2, RMSE and the deviation ratio between urban and rural areas to measure the difference in the RF model performance in urban and rural areas.

Fig. 8 shows the differences in R^2 between urban and rural areas in different years for the hourly air temperature estimations. In the figure, $R^2_{urban–rural}$ greater than zero denotes that R^2 is better for air temperature estimation in urban areas than rural areas. The results show that at night (20:00–07:00), which is also the period that continuously maintains good overall R^2, urban areas afforded better R^2 than rural areas. However, during the daytime period, when the overall R^2 was good (14:00–16:00), R^2 in urban areas was generally lower than that in rural areas. In contrast, during the remaining periods, when the overall R^2 was relatively low, the urban and rural areas did not exhibit a general advantage or disadvantage in R^2 across the years.

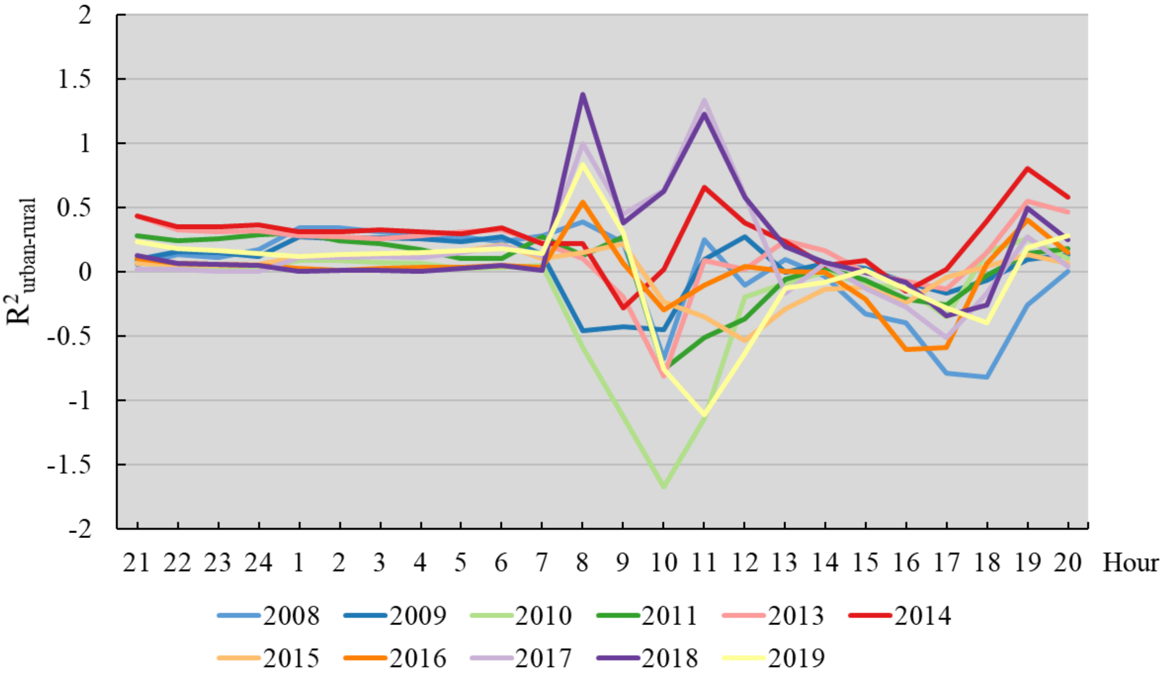


Fig. 8 Differences in R^2 between urban and rural areas in different years for hourly air temperature estimations

A similar comparison was applied for RMSE. Fig. 9 shows the differences in RMSE between urban and rural areas in different years for the hourly air temperature estimations. Since RMSE measures the absolute error between the estimated and observed air temperatures, an $RMSE_{urban–rural}$ less than zero indicates that the estimated temperature in urban areas is closer to the observed temperature than that in rural areas, and vice versa. Unlike R^2, RMSE was consistently smaller in urban areas than in rural areas throughout the day, indicating better performances in urban areas.

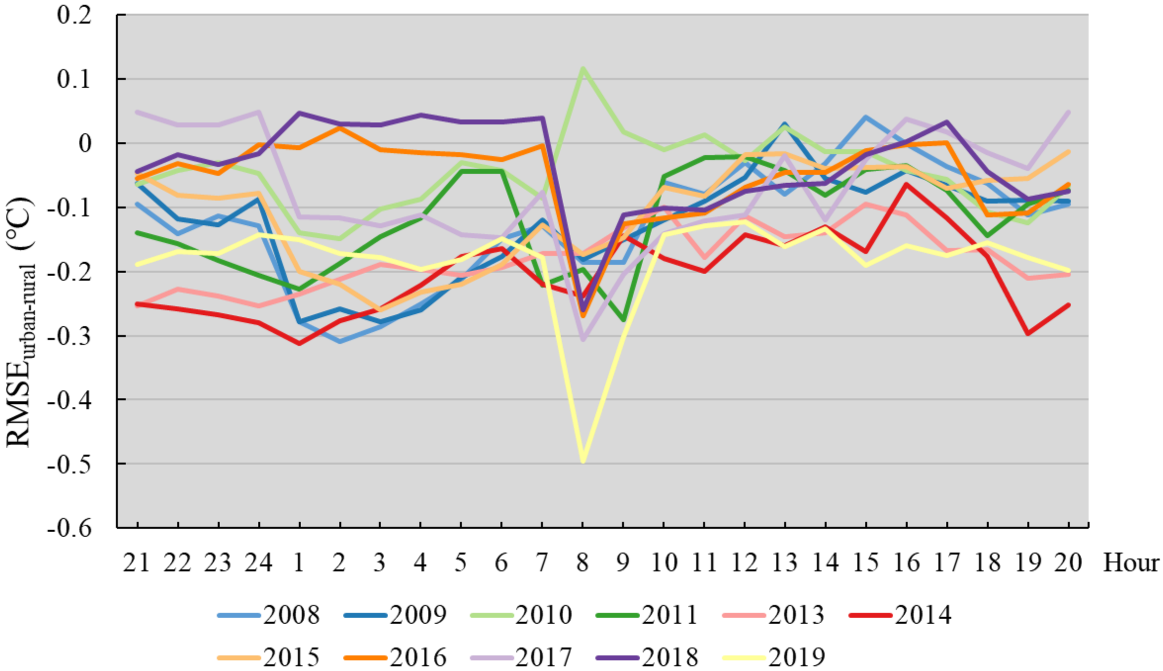


Fig. 9 Differences in RMSE between urban and rural areas in different years for the hourly air temperature estimations.

Figure 10 shows the urban–rural difference in the performance of the RF models in terms of the relative error by comparing the deviation ratios. A value of less than zero on the Y-axis signifies that urban areas afford a smaller deviation ratio than rural areas, signifying better model performance. The results show that the deviation ratio was consistently slightly lower in urban areas than rural areas for most nighttime hours. In contrast, the difference was insignificant during the daytime, or rural areas performed marginally better than urban areas.

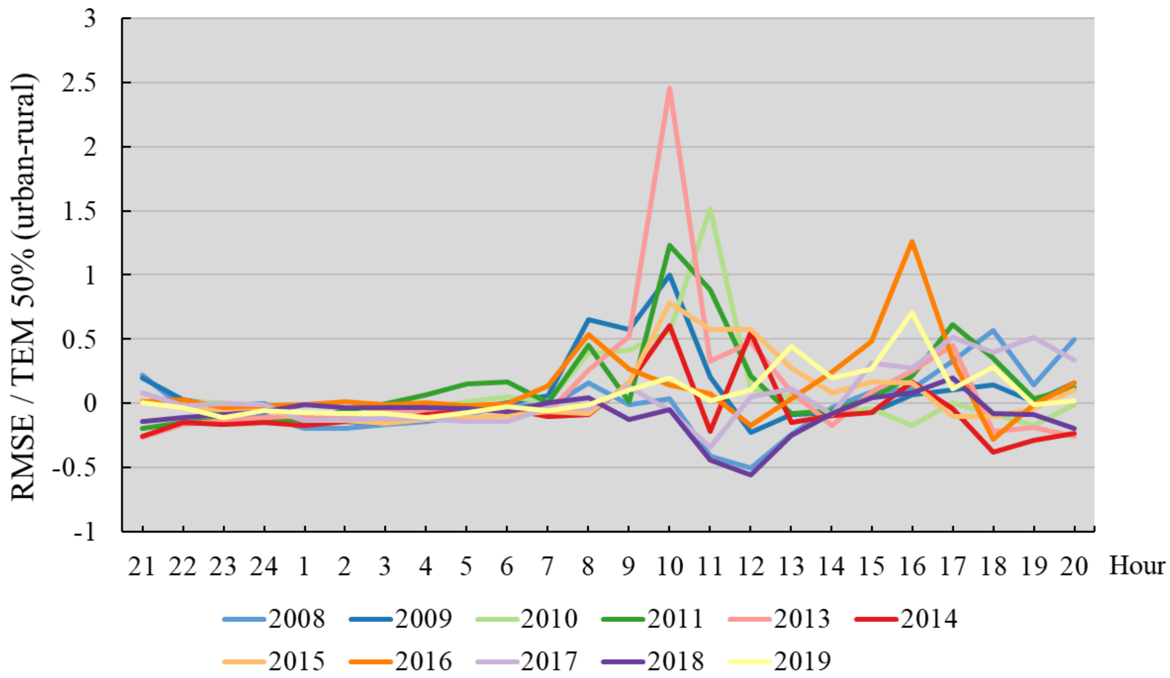


Fig. 10 Differences in the deviation ratio of the estimated hourly air temperatures between urban and rural areas in different years. 'TEM 50%' denotes the difference between the air temperatures observed at the middle 50% urban/rural weather stations.

4. Discussion

4.1. Nighttime vs daytime estimation

Overall, the results show that the RF models for estimating hourly air temperatures performed better at nighttime than daytime. This suggests that the dataset we created is appropriate for urban climate studies, such as UHI, which have been demonstrated to be typically more pronounced at nighttime than daytime[58-60]. Note that the overall R^2 of the RF models was satisfactory, although R^2 was negative for some hours, mainly since we directly calculated R^2 using the estimated and observed temperatures instead of linearly regressing them before calculating R^2. On the other hand, the estimated and observed temperatures maintained a high linear correlation (Pearson's correlation coefficient, R) in almost all hourly periods (Figure S4–S14).

To improve the relatively low accuracy of air temperature estimation during the daytime, we tried modelling adjustment. We separated the 7:00–21:00 period from the whole day for RF modelling. However, the adjusted daytime models did not significantly improve the estimation accuracy during the daytime and presented the same hourly accuracy trajectories as the whole-day models in different years. Furthermore, we determined that the estimation accuracy always started decreasing in the morning after the sun rose and the fog gradually dissipated, it recovered in the early afternoon when the solar radiation was stable and then decreased again when the sun went down and the solar radiation decreased. The decrease in evaluation accuracy always occurred when there was a significant change in solar radiation. A similar situation has been observed in some other studies on spatial air temperature estimation, where the accuracy was lower in the daytime than in the nighttime[61]. Therefore, we infer that the variation in solar radiation due to the Earth's rotation likely decreases the temperature estimation accuracy as it is the primary source of surface heat, subsequently causing a minor air temperature difference during the daytime than the nighttime[62]. However, due to the lack of local observation data, it is not included in the driving factors. Thus, we currently recommend using the nighttime portion of our dataset.

4.2. Importance assessment of drivers

According to the importance assessment of the drivers, the top importance drivers are mainly the meteorological drivers, 74 of the 90 selected drivers. Among them, RHU was the most important driver. The RHUs for each hour within the last 24 h were input into the 90 drivers, totally contributing 52.9% importance. RHU from 1 h prior was the most important driver, contributing 40.9% importance, while RHUs from 10 h, 24 h and 16 h prior were also selected as the top 10 most important drivers. The current time (h) is the second most important driver (26.6%), demonstrating the inherent characteristics of air temperature at different times of the day. Additionally, PRSs for each hour within the last 24 h contributed a total of 7.2% importance. Simultaneously, the landscape and geographic (elevation, latitude and longitude) factors also evidently influence the final spatial pattern of temperature, contributing 5.9% and 5.8% importance, respectively. Therefore, considering more landscape and physical drivers to finely depict the hourly air temperature pattern should be helpful.

4.3. Landscape vs temperature pattern

The spatial pattern of the air temperature estimations exhibited a pronounced landscape divide, which was associated with landscape drivers. Comparing the spatial air temperature patterns, the LCZ maps and Digital Elevation Model (DEM), we determined that the landscape divide appeared in the contiguous area of LCZ A (dense trees). In other words, air temperatures tend to be cooler in the mountainous regions with contiguous dense trees than in the areas of other land types, such as plains. Some users may be concerned about the accuracy of this hourly air temperature dataset in mountainous regions. However, since none of the weather stations are located in mountainous regions with continuous dense trees, we cannot specifically verify the air temperature estimation accuracy there. Therefore, we recommend that these users consider the factors of mountainous regions and plains when using this dataset.

Furthermore, the urban–rural comparison showed that the models generally had better accuracy in urban areas. Moreover, the nighttime temperature pattern showed some correlation with urban morphology. The tracking of the early- and late-night temperature patterns revealed that air temperatures tend to fall more slowly in the core of metropolitan areas than in the urban fringe. Therefore, we believe that this product will be useful for urban-temperature-related studies.

4.4. Comparison to other studies

Using Fig. 6, we have demonstrated the advantages of ML over conventional interpolation methods in depicting the hourly air temperature distributions in terms of presenting spatial details. Simultaneously, our air temperature mapping accuracy is comparable to that of other studies. On the one hand, hourly air temperature mapping is not well practised. The existing hourly air temperature mapping studies[35,61] typically achieve RMSE and MAE of 0.8–1.9 °C and 0.6–1.5 °C, respectively. On the other hand, the accuracy of our hourly air temperature mapping can be even better than that of the daily air temperature mapping. For example, a national-scale daily air temperature mapping using deep learning[63] affords RMSE and MAE of 2.0 and 1.5 °C, respectively. Overall, our hourly air temperature mapping achieves comparable or even better accuracy.

Additionally, the previous hourly and daily air temperature estimation studies are mainly driven by multi-source remote sensing imagery; however, this study focused on integrating meteorological station data and remote sensing techniques for air temperature estimation. In the future, to improve the air temperature estimation accuracy, more available near real-time remote sensing imagery along with meteorological data and remote sensing techniques could be included.

4.5. Potential applications

Our proposed hourly temperature dataset has the potential for application in various fields. For example, this dataset provides air temperature maps with more spatial detail than traditional air temperature maps obtained by station interpolation, providing better weather service for relevant studies such, as UHI and heat wave. Additionally, the hourly air temperature maps can strongly support health-related heat exposure risk studies, such as blood pressure and myocardial infarction[64,65]. Moreover, air temperature is closely related to energy consumption[66], precipitation[67,68] and air pollution[69]. Therefore, the hourly air temperature maps can contribute towards affording an accurate assessment of urban environmental studies on a fine scale, such as at a building or community level[70].

4.6. Study limitations and future work

Despite several benefits of this dataset, some limitations still exist. First, the meteorological spatial drivers used to predict air temperatures were obtained via Kriging interpolation. In the future, with more efficient interpolation methods, meteorological drivers with more spatially detailed information could further improve the accuracy of the air temperature maps. Second, the accuracy of this dataset is relatively low during the daytime, especially in the morning and at dusk. We believe that this is related to the rapid changes in solar radiation effected by the sun's rising and setting. Therefore, hourly solar radiation could be added to the driving factor in future work. Third, the RF modelling herein only focused on the 1-km scale, and the optimal scale for RF models in air temperature estimation is a topic worth exploring in the future. Forth, although comparable to existing LCZ classification studies, the accuracy of the LCZ maps herein is still not flawless. In future work, improvements in LCZ map accuracy could help enhance the air temperature mapping performance. Furthermore, in the future, if hourly air temperature mapping is extended to cover the whole year, the effect of seasonal differences may need to be considered in the model.

5. Conclusion

Herein, we presented an hourly air temperature mapping method at 1-km resolution by adopting the ML (RF algorithm) technology. The method considered topography and LCZ-based landscape drivers; consequently, the air temperature mapping maintained a satisfactory accuracy while affording a more detailed air temperature pattern than spatial interpolation methods. The generated hourly air temperature maps exhibited particularly outstanding accuracy during the nighttime and showed a pattern of slower cooling processes in the urban core during the nighttime than that in the urban fringe, which can help improve studies such as UHI. Moreover, the importance assessment of the driving factors revealed the essential contribution of relative humidity to air temperature mapping, while landscape drivers played a nonnegligible role. Furthermore, given the high spatiotemporal resolution, the generated air temperature mapping can remarkably contribute towards understanding the spatial patterns of urban climate and health-related heat exposure risk studies.

Contributors

GC, YS and CR conceptualised this paper. CR led the team. GC, YS, RW and CR completed the original draft. All authors edited and revised the final manuscript.

Declaration of interests

The authors declare no competing interests.

Acknowledgments

This work was supported by the supported by the Research Impact Fund (Project No. R4046-18F), the Theme Based Research Fund (Project No.T22-504/21-R) of the Hong Kong Research Grant Council, Hong Kong SAR, China, and the Guangdong Basic and Applied Basic Research Foundation (Project No. 2020A1515011235).

1. McGregor, G. A. R. C. Urban Climate Science for Planning Healthy Cities. In: Ren, C. A. M. G., editor. Cham: Springer International Publishing; 2021. pp. 3-16.
2. Portela, C. I., Massi, K. G., Rodrigues, T. & Alcântara, E. Impact of urban and industrial features on land surface temperature: Evidences from satellite thermal indices. *Sustainable Cities and Society*. **56**, 102100 (2020).
3. Yang, J. et al. Contribution of urban ventilation to the thermal environment and urban energy demand: Different climate background perspectives. *Science of the Total Environment*. **795**, 148791 (2021).
4. Yang, J., Yang, Y., Sun, D., Jin, C. & Xiao, X. Influence of urban morphological characteristics on thermal environment. *Sustainable Cities and Society*. **72**, 103045 (2021).
5. Santamouris, M. Recent progress on urban overheating and heat island research. Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. *Energy and Buildings*. **207**, 109482 (2020).
6. Wang, Y. & Zacharias, J. Landscape modification for ambient environmental improvement in central

business districts--a case from Beijing. *Urban forestry \& urban greening*. **14**, 8-18 (2015).

7. Rosenthal, J. K. *Evaluating the impact of the urban heat island on public health: Spatial and social determinants of heat-related mortality in New York City*. (Columbia University, 2010).
8. Masson-Delmotte, V. et al. Global warming of 1.5 C. *An IPCC Special Report on the impacts of global warming of*. **1**, (2018).
9. Savić, S., Selakov, A. & Milošević, D. Cold and warm air temperature spells during the winter and summer seasons and their impact on energy consumption in urban areas. *Natural Hazards*. **73**, 373-387 (2014).
10. Lau, K. K., Chung, S. C. & Ren, C. Outdoor thermal comfort in different urban settings of sub-tropical high-density cities: An approach of adopting local climate zone (LCZ) classification. *Building and Environment*. **154**, 227-238 (2019).
11. Macintyre, H. L. et al. Assessing urban population vulnerability and environmental risks across an urban area during heatwaves – Implications for health protection. *Science of the Total Environment*. **610**, 678-690 (2018).
12. Basu, R., Feng, W. & Ostro, B. D. Characterizing temperature and mortality in nine California counties. *Epidemiology*, 138-145 (2008).
13. Ostro, B. D., Roth, L. A., Green, R. S. & Basu, R. Estimating the mortality effect of the July 2006 California heat wave. *Environmental Research*. **109**, 614-619 (2009).
14. Landsberg, H. E. *The urban climate*. (Academic press, 1981).
15. Kloog, I., Nordio, F., Coull, B. A. & Schwartz, J. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. *Remote Sensing of Environment*. **150**, 132-139 (2014).
16. Wang, M. et al. Comparison of Spatial Interpolation and Regression Analysis Models for an Estimation of Monthly Near Surface Air Temperature in China. *Remote Sensing*. **9**, 1278 (2017).
17. Florio, E. N., Lele, S. R., Chi Chang, Y., Sterner, R. & Glass, G. E. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. *International Journal of Remote Sensing*. **25**, 2979-2994 (2004).
18. Alvares, C. A., Stape, J. L., Sentelhas, P. C. & de Moraes Gonçalves, J. L. Modeling monthly mean air temperature for Brazil. *Theoretical and Applied Climatology*. **113**, 407-427 (2013).
19. Zhao, C., Nan, Z. & Cheng, G. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains, China. *Ecological Modelling*. **189**, 209-220 (2005).
20. Salcedo-Sanz, S., Deo, R. C., Carro-Calvo, L. & Saavedra-Moreno, B. Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theoretical and Applied Climatology*. **125**, 13-25 (2016).
21. Oke, T. R., Mills, G., Christen, A. & Voogt, J. A. *Urban climates*. (Cambridge University Press, 2017).
22. Simon, H. Modeling urban microclimate: development, implementation and evaluation of new and improved calculation methods for the urban microclimate model ENVI-met., Mainz, Univ., Diss., 2016, 2016.
23. Dos Santos, R. S. Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. *International Journal of Applied Earth Observation and Geoinformation*. **88**, 102066 (2020).
24. Nieto, H., Sandholt, I., Aguado, I., Chuvieco, E. & Stisen, S. Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula. *Remote Sensing of Environment*. **115**, 107-116 (2011).

25. Prihodko, L. & Goward, S. N. Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*. **60**, 335-346 (1997).
26. Agam, N., Kustas, W. P., Anderson, M. C., Li, F. & Neale, C. M. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sensing of Environment*. **107**, 545-558 (2007).
27. Manoli, G. et al. Magnitude of urban heat islands largely explained by climate and population. *Nature*. **573**, 55-60 (2019).
28. Hou, P. et al. Near-surface air temperature retrieval from satellite images and influence by wetlands in urban region. *Theoretical and Applied Climatology*. **111**, 109-118 (2013).
29. Good, E. Daily minimum and maximum surface air temperatures from geostationary satellite data. *Journal of Geophysical Research: Atmospheres*. **120**, 2306-2324 (2015).
30. Huang, F. et al. Air temperature estimation with MODIS data over the Northern Tibetan Plateau. *Advances in Atmospheric Sciences*. **34**, 650-662 (2017).
31. Meyer, H. et al. Mapping daily air temperature for Antarctica based on MODIS LST. *Remote Sensing*. **8**, 732 (2016).
32. Yoo, C., Im, J., Park, S. & Quackenbush, L. J. Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data. *Isprs Journal of Photogrammetry and Remote Sensing*. **137**, 149-162 (2018).
33. Zhang, H., Zhang, F., Ye, M., Che, T. & Zhang, G. Estimating daily air temperatures over the Tibetan Plateau by dynamically integrating MODIS LST data. *Journal of Geophysical Research: Atmospheres*. **121**, 11, 411-425, 441 (2016).
34. Vancutsem, C., Ceccato, P., Dinku, T. & Connor, S. J. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment*. **114**, 449-465 (2010).
35. Zhou, B. et al. Estimating near - surface air temperature across Israel using a machine learning based hybrid approach. *International Journal of Climatology*. **40**, 6106-6121 (2020).
36. Emmanuel, R. *Urban Heat Island Mitigation Technologies*. (MDPI, 2021).
37. Stewart, I. D. & Oke, T. R. Local Climate Zones for Urban Temperature Studies. *Bulletin of the American Meteorological Society*. **93**, 1879-1900 (2012).
38. Li, L. et al. Variabilities of Land Surface Temperature and Frontal Area Index Based on Local Climate Zone. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. **15**, 2166-2174 (2022).
39. Ren, J. et al. Exploring thermal comfort of urban buildings based on local climate zones. *Journal of Cleaner Production*. **340**, 130744 (2022).
40. Zhao, Z., Sharifi, A., Dong, X., Shen, L. & He, B. Spatial Variability and Temporal Heterogeneity of Surface Urban Heat Island Patterns and the Suitability of Local Climate Zones for Land Surface Temperature Characterization. *Remote Sensing*. **13**, (2021).
41. WMO. *Guide to meteorological instruments and methods of observation (WMO-No. 8)*. (Secretariat of the World Meteorological Organization, Geneva, Switzerland, 2008).
42. Konarska, J., Holmer, B., Lindberg, F. & Thorsson, S. Influence of vegetation and building geometry on the spatial variations of air temperature and cooling rates in a high-latitude city. *International Journal of Climatology*. **36**, 2379-2395 (2016).
43. Yin, C., Yuan, M., Lu, Y., Huang, Y. & Liu, Y. Effects of urban form on the urban heat island effect based on spatial regression model. *Science of the Total Environment*. **634**, 696-704 (2018).

44. Bechtel, B. et al. Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. *Isprs International Journal of Geo-Information*. **4**, 199-219 (2015).
45. Wang, R. et al. Detecting multi-temporal land cover change and land surface temperature in Pearl River Delta by adopting local climate zone. *Urban Climate*. **28**, 100455 (2019).
46. Chung, L. C. H., Xie, J. & Ren, C. Improved machine-learning mapping of local climate zones in metropolitan areas using composite Earth observation data in Google Earth Engine. *Building and Environment*. **199**, 107879 (2021).
47. Breiman, L. Random forests. *Machine Learning*. **45**, 5-32 (2001).
48. Lee, S., Choi, H., Cha, K. & Chung, H. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: Gasoline and naphtha. *Microchemical Journal*. **110**, 739-748 (2013).
49. Kamusoko, C. & Gamba, J. Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model. *Isprs International Journal of Geo-Information*. **4**, 447-470 (2015).
50. Katpatal, Y. B., Kute, A. & Satapathy, D. R. Surface- and Air-Temperature Studies in Relation to Land Use/Land Cover of Nagpur Urban Area Using Landsat 5 TM Data. **134**, 110-118 (2008).
51. Shojaei, P. et al. Effect of different land cover/use types on canopy layer air temperature in an urban area with a dry climate. *Building and Environment*. **125**, 451-463 (2017).
52. Forman, R. T. T. *Land Mosaics: The Ecology of Landscapes and Regions*. (Cambridge University Press, 1995).
53. Southworth, J., Nagendra, H. & Tucker, C. Fragmentation of a Landscape: Incorporating landscape metrics into satellite analyses of land-cover change. *Landscape Research*. **27**, 253-269 (2002).
54. Neel, M. C., McGarigal, K. & Cushman, S. A. Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landscape Ecology*. **19**, 435-455 (2004).
55. Roy, H. & Mark, C. Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. *Progress in Physical Geography*. **20**, 418-445 (1996).
56. McGarigal, K., Cushman, S. A. & Ene, E. FRAGSTATS: spatial pattern analysis program for categorical maps. The University of Massachusetts, Amherst, Massachusetts, USA; 2012.
57. Zhang, D., Liu, X., Lin, Z., Zhang, X. & Zhang, H. The delineation of urban growth boundaries in complex ecological environment areas by using cellular automata and a dual-environmental evaluation. *Journal of Cleaner Production*. **256**, 120361 (2020).
58. Bohnenstengel, S. I., Evans, S., Clark, P. A. & Belcher, S. E. Simulations of the London urban heat island. *Quarterly Journal of the Royal Meteorological Society*. **137**, 1625-1640 (2011).
59. Dialesandro, J. M., Wheeler, S. M. & Abunnasr, Y. Urban heat island behaviors in dryland regions. *Environmental Research Communications*. **1**, 81005 (2019).
60. Kim, Y. & Baik, J. Maximum Urban Heat Island Intensity in Seoul. Boston MA, USA: American Meteorological Society; 2002. pp. 651-659.
61. Zhang, Z. & Du, Q. Hourly mapping of surface air temperature by blending geostationary datasets from the two-satellite system of GOES-R series. *Isprs Journal of Photogrammetry and Remote Sensing*. **183**, 111-128 (2022).
62. Bernard, J., Musy, M., Calmet, I., Bocher, E. & Keravec, P. Urban heat island temporal and spatial variations: Empirical modeling from geographical and meteorological data. *Building and Environment*. **125**, 423-438 (2017).

63. Shen, H. et al. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. *Remote Sensing of Environment*. **240**, 111692 (2020).
64. Xu, D. et al. Acute effects of temperature exposure on blood pressure: An hourly level panel study. *Environment International*. **124**, 493-500 (2019).
65. Bhaskaran, K. et al. Heat and risk of myocardial infarction: hourly level case-crossover analysis of MINAP database. *Bmj*. **345**, (2012).
66. Fumo, N. & Biswas, M. R. Regression analysis for prediction of residential energy consumption. *Renewable and sustainable energy reviews*. **47**, 332-343 (2015).
67. Li, L., Zha, Y. & Wang, R. Relationship of surface urban heat island with air temperature and precipitation in global large cities. *Ecological Indicators*. **117**, 106683 (2020).
68. Mishra, V., Wallace, J. M. & Lettenmaier, D. P. Relationship between hourly extreme precipitation and local air temperature in the United States. *Geophysical Research Letters*. **39**, (2012).
69. Kalisa, E., Fadlallah, S., Amani, M., Nahayo, L. & Habiyaremye, G. Temperature and air pollution relationship during heatwaves in Birmingham, UK. *Sustainable Cities and Society*. **43**, 111-120 (2018).
70. Gonzalez, P. A. & Zamarreno, J. M. Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. *Energy and Buildings*. **37**, 595-601 (2005).

Table

Table 1 Categories and definitions of local climate zone (LCZ) simplified from Stewart & Oke[36]

LCZ types	Built and land cover types
LCZ 1	Compact high-rise
LCZ 2	Compact mid-rise
LCZ 3	Compact low-rise
LCZ 4	Open high-rise
LCZ 5	Open mid-rise
LCZ 6	Open low-rise
LCZ 7	Lightweight low-rise
LCZ 8	Large low-rise
LCZ 9	Sparsely built
LCZ 10	Heavy industry
LCZ A	Dense trees
LCZ B	Scattered trees
LCZ C	Bush, scrub
LCZ D	Low plants
LCZ E	Bare rock or paved
LCZ F	Bare soil or sand
LCZ G	Water
LCZ H	Wetlands#

Wetlands is an additional LCZ type that adapted the land surface properties of coastal cities in the Guangdong province.

Table 2 The accuracy of the RF models for each year

Year	R^2	oob_score	RMSE (℃)	MAE(℃)
2008	0.8036	0.7992	1.5112	1.0940
2009	0.8127	0.8084	1.4592	1.0879
2010	0.7684	0.7652	1.6049	1.1953
2011	0.8202	0.8132	1.5062	1.0951
2013	0.7685	0.7648	1.5498	1.1163
2014	0.8272	0.8252	1.4336	1.0323
2015	0.8197	0.8153	1.3368	0.9886
2016	0.7725	0.7697	1.5097	1.1133
2017	0.7810	0.7762	1.5145	1.1259
2018	0.8041	0.7997	1.3607	1.0002
2019	0.8234	0.8188	1.5165	1.1100
Mean	0.8001	0.7960	1.4821	1.0872

References:

Stewart, I. D. and T. R. Oke (2012). "Local Climate Zones for Urban Temperature Studies." Bulletin of the American Meteorological Society **93** (12): 1879-1900.

Integrating weather observations and local climate zone-based landscape patterns for regional hourly air temperature mapping using machine learning

Guangzhao Chen[1,4], Yuan Shi[2], Ran Wang[3], Chao Ren[4]*, Edward Ng[5], Xiaoyi Fang[6], Zhihua Ren[7]

1. Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong, SAR, China
2. University of Liverpool
3. Nankai University
4. Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China
5. School of Architecture, The Chinese University of Hong Kong, Hong Kong, SAR, China
6. Chinese Academy of Meteorological Sciences, China
7. National Meteorological Information Center, China

*corresponding author: Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China (renchao@hku.hk)

Contributions

GC, YS and CR conceptualized this paper. CR led the team. GC, YS, RW and CR completed the original draft. All authors edited and revised the final manuscript.