Constructing air temperature and relative humidity-based hourly thermal comfort dataset for a high-density city using machine learning

Guangzhao Chen [a,b], Junyi Hua [c], Yuan Shi [d], Chao Ren [a,*]

[a] Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong Special Administrative Region
[b] Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong Special Administrative Region
[c] School of International Affairs and Public Administration, Ocean University of China, Qingdao, China
[d] Department of Geography & Planning, University of Liverpool, Liverpool, UK

ABSTRACT

Global warming causes new challenges for urban citizens and metropolitan governments in adapting to the changing thermal environment. However, fine-scale spatiotemporal mapping of urban thermal environments has been inadequate. Therefore, this study takes a typical high-density city, Hong Kong, as an example and utilises a machine learning algorithm, the random forest (RF), to carry out 100 m resolution hourly thermal environment mapping, including air temperature (Ta), relative humidity (RH) and the net effective temperature (NET), for the summer season (May to September) of 2008–2018, considering meteorological drivers, topography and local-climate-zone-based landscape drivers. The validation results show that the developed dataset achieves satisfactory accuracy. The mean values of R^2, root mean square error (RMSE) and mean absolute error (MAE) for Ta achieve 0.8723, 1.1160 °C and 0.8227 °C, respectively, while those for RH reach 0.7970, 5.3816% and 3.8641%. In addition, the thermal comfort index, NET, reveals that people in built-up areas feel hotter than measured by Ta during the night due to the urban heat island effect. We believe this newly developed thermal comfort dataset can provide novel, reliable and fine-grained data support for urban climate research areas such as urban heat islands, heat exposure, heat-related health risk assessment, and urban energy consumption estimation.

1. Introduction

With urbanisation (Wu et al., 2020) and the frequent occurrence of extreme weather events due to climate change (Stott, 2016), including hot extremes in summer (Ma and Yuan, 2021; Zhang and Wu, 2011), people living in high-density cities are increasingly challenged. Thus, thermal comfort is one of the current hot topics of interest in urban climate research, which involves studying the links between the outdoor environment and human well-being (Rosenthal, 2010). Typically, thermal comfort can be assessed and measured by combining vital meteorological indicators, such as air temperature (Ta) and relative humidity (RH) (Matzarakis and Amelung, 2008). Using these thermal comfort indicators, one can conduct research in areas such as urban heat islands (de Souza et al.,

---

2014; Khan and Chatterjee, 2016; Liu et al., 2007), compound extreme weather (Dzwonkowski et al., 2020; Li et al., 2020), energy consumption (Braun et al., 2014; Ihara et al., 2008), and heat-related health risks (Bhaskaran et al., 2012; Xu et al., 2019). However, due to technical limitations, thermal comfort datasets with high spatiotemporal resolutions are still scarce. Although meteorological observation networks can provide long-term observations of meteorological indicators with a high temporal resolution, it is still insufficient for high spatial resolution mapping. In addition, particularly in high-density urban areas, the heterogeneity within cities may cause large variations in meteorological indicator values, which are difficult to reflect by data from observation networks (Kloog et al., 2014). This study will focus on mapping the two thermal comfort indicators, Ta and RH.

Ta is one of the most important indicators of what constitutes a measure of thermal comfort (Lau et al., 2019). Land surface temperature (LST) is often used as a proxy for the spatial distribution of air temperature (Lin et al., 2016), because satellite imagery can provide spatially continuous data for a large spatial extent. However, LST and air temperature still cannot be equated (Zhao et al., 2020). LST cannot directly reflect thermal comfort as well as Ta. The spatiotemporal changes of Ta on small scales are largely affected by the landscape pattern of land cover and land use (LULC), because the land surface changes the boundary layer climate condition (Emmanuel, 2021). Therefore, obtaining accurate Ta spatial patterns with high spatiotemporal resolutions is not easy.

Previous studies have carried out many attempts at Ta mapping. There are three broad categories of common Ta mapping methods. The first is traditional spatial interpolation methods, which include Kriging interpolation (Florio et al., 2004) and inverse distance weighted (IDW) interpolation (Wang et al., 2017). The interpolation methods follow "the first law of geography", which states that the near things are more related to the distant things (Tobler, 2004). They require that the sampling points be distributed as evenly as possible across the study area and that only the planar distances between them are considered (Florio et al., 2004). Therefore, the uneven distribution of sampling points in reality and the differences in geographical conditions can introduce errors. The second is climate models, which can be subdivided into macro-scale (i.e. global or regional scale), mesoscale and micro-scale climate models (Oke et al., 2017; Simon, 2016; Yan et al., 2020). Climate models are a class of mechanistic models that simulate the spatial and temporal variability of different elements through the action of physical mechanisms (Lamarque et al., 2013). Climate models can simulate the spatial pattern of Ta at the hourly or minute temporal resolution, while their spatial resolution varies from 50 km to 1 m as the scale of the model varies from large to small (Oke et al., 2017; Simon, 2016). Macro-scale climate models with coarse spatial resolution often require simplifying elements such as urban structure (Oke et al., 2017). Mesoscale climate models have a more complex structure, considering land surface-atmosphere interactions, so they can only cover horizontal scales of tens to hundreds of kilometres with kilometre-level resolution (Yan et al., 2020). As the scale is further downscaled, micro-scale climate models with better spatial resolution make strict demands on computational resources, computational time, model complexity and fine historical input data (Simon, 2016). As a result, micro-scale climate models with high spatial resolutions are difficult to use for city-scale Ta simulations (Acero and Arrizabalaga, 2018). In addition, the higher temporal resolution also means more demands on computing resources. Therefore, it is difficult to combine both high spatial resolution and high temporal resolution in urban-scale Ta mapping through climate models. The third is regression methods, which estimate the distribution of Ta in space and time by establishing quantitative relationships between Ta and the relevant elements. In addition to traditional simple or multiple linear regression (Alvares et al., 2013; Zhao et al., 2005) including geographically weighted regression (GWR) (Wang et al., 2017), machine learning (ML) is a popular approach. Typical ML include support vector regression (SVR) (Chevalier et al., 2011), artificial neural network (ANN) (Chronopoulos et al., 2008), random forest (RF) (Mohsenzadeh Karimi et al., 2020) and gradient boosting decision tree (GBDT) (Friedman, 2001). Machine learning can fit non-linear correlations, providing a higher estimation accuracy than traditional linear regression. Moreover, ML is more applicable to mining relationships with multiple variables and large data volumes. In general, satellite remote sensing imagery providing metrics such as reflectance, LST and Normalized Difference Vegetation Index (NDVI) are used as variables to predict the Ta pattern (Shen et al., 2020; Zhou et al., 2020).

RH is another crucial indicator of thermal comfort. Similar to Ta, the spatial distribution of RH can be estimated by three routes: traditional spatial interpolation (Hodam et al., 2017; Li et al., 2014), climate models (Acero and Arrizabalaga, 2018; Maharjan and Regmi, 2015) and regression (Fries et al., 2012; Hanoon et al., 2021; Li and Zha, 2018). Likewise, the advantages and disadvantages of each type of RH mapping are the same as those of Ta mapping before mentioned. We found fewer studies on RH mapping than Ta, especially RH mapping using machine learning. However, a recent study reported an increase in the temporal resolution of RH mapping by machine learning to the daily scale, while most RH mapping studies focus on the monthly scale (Hanoon et al., 2021). This study only used historical RHs as the driving factors for estimating the current RH.

Overall, previous studies of thermal comfort datasets have had a few limitations. First, the high spatial resolution of thermal comfort datasets is not well balanced with the high temporal resolution (Di Napoli et al., 2021). Second, the effect of LULC, which is a non-negligible factor influencing near-surface local climatic conditions (Emmanuel, 2021), is rarely considered in Ta and RH mapping. Third, the thermal comfort dataset pays insufficient focus on high-density cities and summer seasons, which are closely related to the well-being of large populations in the context of a warming climate but require high spatial and temporal resolution.

Therefore, in this study, we aimed to generate a spatiotemporal hourly thermal comfort dataset, including air temperature and relative humidity, at a 100 m resolution for Hong Kong, a typical high-density city. This dataset covers the summer season (May to September) from 2008 to 2018. To better depict the spatial pattern, we employed ML algorithms and simultaneously considered the effects of meteorological drivers, landscape drivers and topography. Notably, the calculation of the landscape driver is based on a detailed LULC classification, local climate zone (LCZ), which subdivides the LULC types within the city.

2. Data and method

The workflow for generating the thermal comfort dataset in this study is shown in Fig. 1. The detailed data preparation and

implementation methods are illustrated in the subsequent subsections.

2.1. Study area

Hong Kong is located on the eastern side of the Pearl River estuary and is part of a transitional region with a tropical and subtropical monsoon climate. It is bordered to the north by Shenzhen, one of the four first-tier cities in mainland China. Hong Kong is an international financial, shipping and trading centre with a developed economy, one of the core cities of the Guangdong-Hong Kong-Macao Greater Bay Area and one of China's windows to the outside world. Moreover, as Hong Kong is mainly hilly, with only about 20% of its land in the lowlands, it makes Hong Kong one of the world's most famous high-density cities. Meanwhile, the heterogeneity brought about by the complex topography also makes it an ideal study area for applying machine learning for Ta and RH mapping with high spatial and temporal resolution.

2.2. Meteorological data (SY)

Meteorological data are provided by the Hong Kong Observatory (HKO) weather stations. These weather stations are located throughout Hong Kong, providing hour-by-hour meteorological observations from 2008 to 2018. Due to the development and management of weather stations, the number of weather stations from 2008 to 2018 ranged from 36 to 46. Fig. 2 shows the distribution of these weather stations in 2018. These weather stations cover various parts of Hong Kong well, with each station averaging 4005.1 m from its nearest neighbouring station.

The weather stations observed five meteorological metrics at hourly intervals, including air temperature (Ta), relative humidity (RH), precipitation (PRE), barometric pressure (PRS) and wind speed (VV2). Meanwhile, the latitude and longitude coordinates and elevation of the weather station were also provided. The observations were used as meteorological drivers in the subsequent machine-learning-based Ta and RH mapping. Of course, Ta was not used as a meteorological driver when estimating Ta to avoid circular arguments. The same was true for RH when estimating RH.

To execute the well-trained machine learning model, we needed to input the spatial pattern of the meteorological drivers. Thus, we employed the Kriging interpolation to interpolate the observed meteorological drivers hour by hour into 100 m resolution maps.
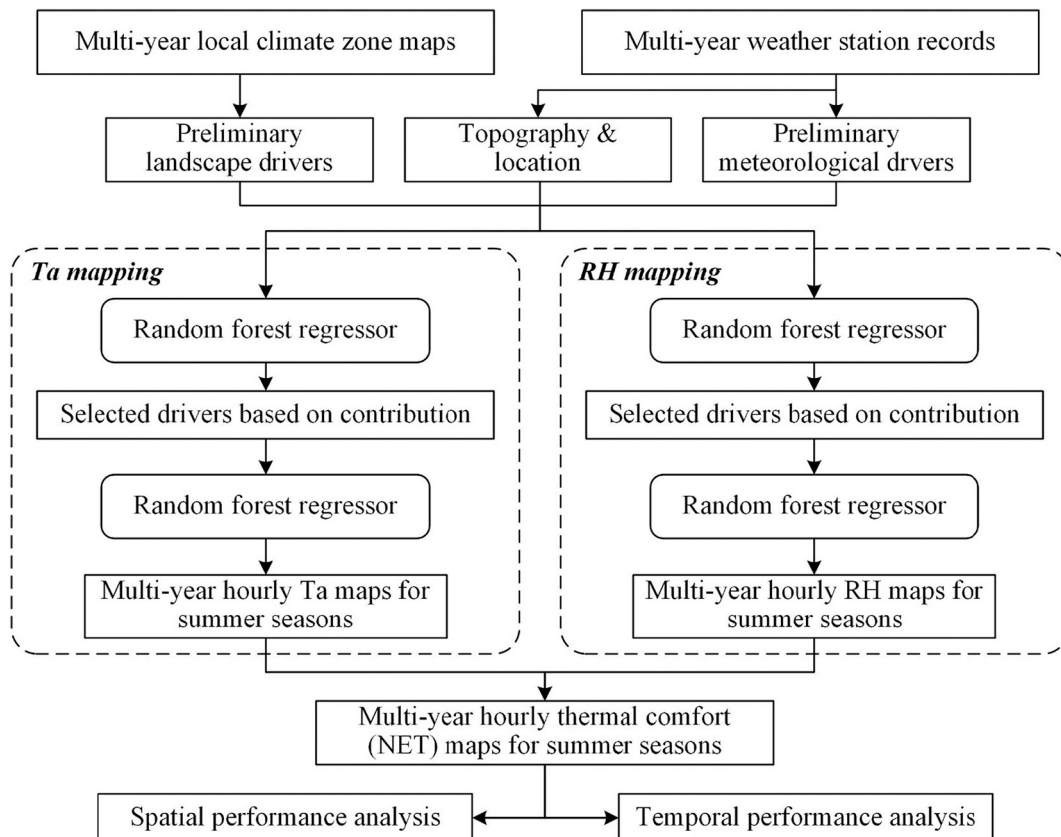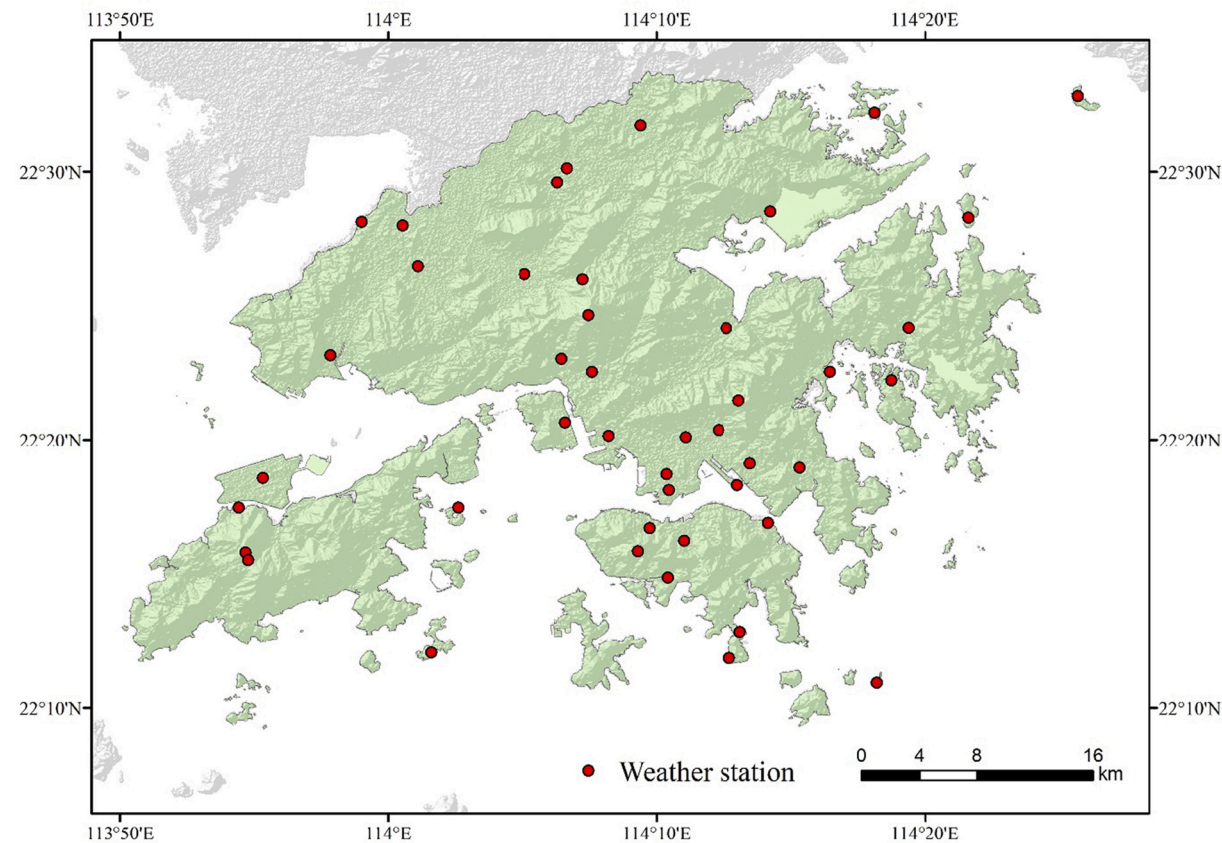

Fig. 1. The workflow in this study.

Fig. 2. Study area and the distribution of weather stations in 2018. Each red dot represents one weather station of Hong Kong Observatory. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. LCZ data and landscape drivers

In previous studies, the LCZ has been demonstrated to be closely related to the urban thermal environment, due to its unique definition of land cover types (Fricke et al., 2022; Li et al., 2022; Ren et al., 2022; Yin et al., 2018; Zhao et al., 2021). The LCZ typically contains 17 land cover types (Stewart and Oke, 2012), sometimes subdivided into an additional wetland type from the water (Chen

Table 1
LCZ types and simplified definitions in Chen et al. (Chen et al., 2022)

LCZ types	Simplified definitions
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[#]

et al., 2022; Chung et al., 2021) (Table 1). It includes both urban-type LCZs and natural-type LCZs. However, unlike traditional land cover classifications, LCZs subdivide urban-type land into ten categories based on physical properties. These physical properties include building density, height, material and area, which depict the urban morphology in detail and are closely linked to the near-surface environment (Stewart and Oke, 2012).

The generation of LCZ datasets is not easy, especially for multi-year, as it requires constructing multi-year sample sets for the numerous LCZ types. Therefore, in this study, we employed the 2008–2018 LCZ dataset with 100 m resolution produced by Chen et al. (Chen et al., 2022), which covers Hong Kong. This LCZ dataset is based on a multi-year comparable LCZ sample set. It performs classification employing the Google Earth Engine (GEE) platform, considering satellite remote sensing images such as Landsat 8, Landsat 5, Sentinel-1 SAR GRD, Sentinel-2 MSI, VIIRS and DMSP OLS, as well as GMTED2010 (Global Multi-resolution Terrain Elevation Data 2010) (Chen et al., 2022). The included remote sensing images are derived from different sensors, including spectral, radar, nighttime light and terrain, providing multi-dimensional information for the LCZ classification. As a result, this LCZ dataset achieves an overall accuracy of 61.61% over a multi-year span, which is a comparable and acceptable accuracy to most current LCZ classification products (Chen et al., 2022).

Moreover, landscape metrics have been proven essential for mapping fine air temperature patterns (Chen et al., 2022; Shojaei et al., 2017). Landscape metrics are developed from the "patch-corridor-matrix" theory in landscape ecology (Forman, 1995). They provide pictures of the landscape's characteristics through three levels of metrics, including patch-, class- and landscape-level. As the names imply, the patch-level metrics reflect the characteristics of a single patch of land covers, such as area, perimeter and shape. Class-level metrics capture the spatial pattern of the characteristics of all patches of a specific land cover type. While the landscape-level metrics provide the picture of how all land cover types are mixed in space. Of course, LCZ can also be combined with landscape metrics to reflect LCZ-based landscape characteristics. Using different search radii when being calculated, landscape metrics can reflect the landscape characteristics within different radii around a location.

In this study, we employed Fragstats software (version 4.2), a well-established and widely used landscape metric software, to calculate LCZ-based landscape metrics for Hong Kong over the study period. Meanwhile, referencing previous studies, we calculated landscape metrics from 100 m to 1000 m search radius. That is, when traversing each grid to calculate the different landscape metrics, we took each grid as a reference point and calculated the different landscape metrics with different search radius from 100 m to 1000 m in 100 m intervals. The selected landscape metrics are listed in Table S1. As there are 18 LCZ types, we finally generated a vast number of landscape drivers, 13,550 in total. However, to reduce the computational burden of subsequent machine learning modelling and to select landscape indices of general interest, those landscape indices that had valid values at less than 50% of the weather stations were excluded. Therefore, only 226 landscape metrics became our preliminary landscape drivers for the subsequent machine learning modelling.

2.4. Mapping hourly air temperatures and relative humidity using machine learning

In this study, we employed the random forest (RF) model (Breiman, 2001), a well-performing, efficient and widely used machine learning algorithm, for hourly Ta and RH mapping in Hong Kong. The RF model consists of a number of decision trees that vote to determine the outcome of regression or classification (Kamusoko and Gamba, 2015). With these "weak" but varied decision trees, the RF model can integrate them into a powerful regressor or classifier. Because of this structure, RF models can be used to solve non-linear problems and effectively avoid overfitting (Lee et al., 2013). Moreover, the RF model can evaluate the importance of each driving factor (Zhang et al., 2020), which is crucial for the subsequent selection of the key drivers and for improving the efficiency of our models.

We used the Python-based "scikit-learn" extension package (Version 0.24.2) to construct annual RF models for estimating Ta and RH, respectively. In the case of the RF model for estimating Ta, the previously prepared meteorological and landscape drivers were fed into the RF model, as well as the current time (hour), longitude, latitude and elevation. The meteorological drivers include RH, PRE, PRS and VV2 for the previous 24 h. Thus, a total of 326 preliminary driving factors were fed into the RF model. However, the large number of driving factors increases the training and prediction time of the model and takes up more computational resources. Moreover, most of these driving factors can not improve the model accuracy. Therefore, we trained the RF model using the 2018 data and selected key driving factors based on their importance. We employed the permutation importance provided by scikit-learn as a metric for importance assessment, which is suitable in cases where a feature has many unique values. The permutation importance of a feature is defined as the deviation of the metric value from the baseline metric value after the permutation of this feature. We executed the training of the RF model 10 times, taking the mean value of permutation importance as the importance of driving factors. These driving factors were ranked in descending order of permutation importance, and we selected the top driving factors that contributed more than 98% in total importance as key driving factors. These key driving factors were then applied to the modelling of the other years. At the same time, we found the optimal number of decision trees by adjusting the n_tree parameter. The same process was followed in the RF modelling for estimating RH, but the only difference was that Ta replaced RH as a meteorological driver.

In constructing an RF model, 70% of the samples were used as training samples, while the remaining 30% were used as test samples. Due to the characteristic of the RF model, approximately 30% of the training samples are not actually involved in the model training in each modelling process, which are called out-of-bag samples. The goodness-of-fit, R (Stott, 2016), calculated using them is called the oob_score, which can be used to measure the model accuracy. Besides, we also calculated three accuracy metrics for the models using the test samples. They are R (Stott, 2016), the root mean square error (RMSE) and the mean absolute error (MAE). These four accuracy metrics provide a comprehensive picture of the accuracy of the RF models.

2.5. The thermal index - Net effective temperature (NET)

Previous studies have developed different thermal indices to reflect the combined or certain aspects of the thermal environment (Matzarakis and Amelung, 2008; Di Napoli et al., 2021; Hentschel, 1986). To reflect the thermal environment under the combined influence of Ta and RH, we employed a thermal index, Net Effective Temperature (NET) (Hentschel, 1986). In addition, NET is adopted by the Hong Kong Observatory (HKO), so the NET evaluation result can directly serve the HKO (Yip et al., 2007). NET was first introduced in 1937 with the name Effective Temperature (Hentschel, 1986). It was initially proposed to measure the effect of RH in hot weather. Later it was improved to consider the effect of wind on human thermal comfort (Hentschel, 1986). That is, the stressful feeling can be exacerbated by calm wind and humidity in hot weather and strong wind and humidity in cold weather (Yip et al., 2007). Therefore, this improvement makes it applicable to human thermal comfort in both hot and cold weather (Li and Chan, 2000). NET is determined by Ta, RH and wind speed and can be described by the following equation (Hentschel, 1986):

$$NET = 37 - \frac{37 - Ta}{0.68 - 0.0014 \times RHU + \frac{1}{1.76 + 1.4v^{0.75}}} - 0.29 \times Ta \times (1 - 0.01 \times RHU) \tag{1}$$

where the unit of NET is °C, Ta represents the air temperature (in °C), RH is the relative humidity (in %), and v means the wind speed (in m/s). The setting of the constants in Eq. (1) was applied in a previous study of Hong Kong (Yip et al., 2007). Since this study only generates Ta and RH mapping, v will be set to 0 when calculating NET. Hence, the NET calculated in this study reflects the most stressful thermal environment felt in the summer in a calm wind condition.
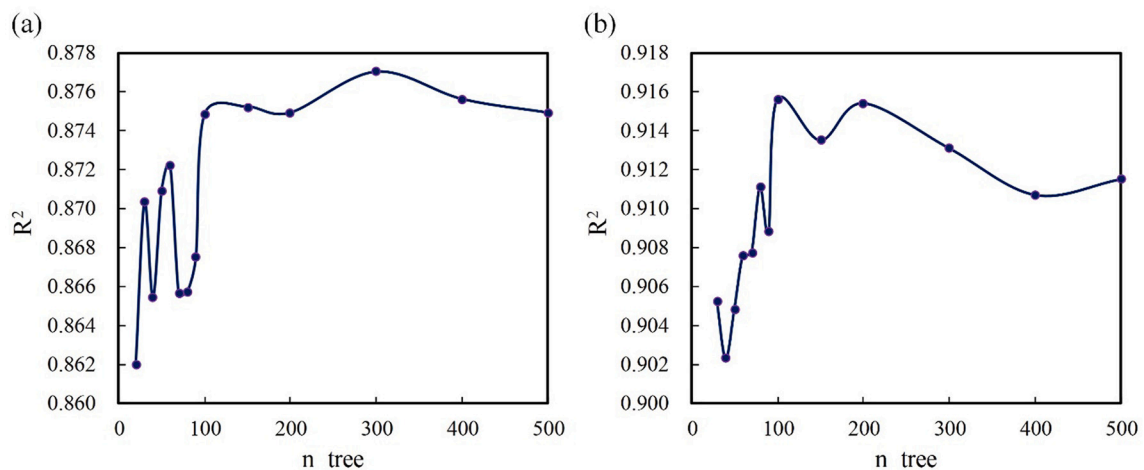
3. Results

3.1. Validation of the hourly Ta maps

By inputting 2018 data, we used the 326 preliminary driving factors to find the optimal n_tree from 20 to 500 for the RF model for estimating Ta, as shown in Fig. 3(a). The results show no significant improvement in the R^2 value of the model after the n_tree exceeds 100. Therefore, we used the top 90 drivers that contributed the most in the importance assessment when n_tree was 100 as the key drivers. The importance assessment showed that the importance score of these 90 key drivers (0.717) contributed 98.3% of the total importance score of the preliminary driving factors (0.729). Further, we searched from 30 to 500 for the optimal n_tree when modelling the Ta estimation using these 90 key drivers (Fig. 3(b)). The results show that the R^2 value of the model also reached the highest when the n_tree was 100. Therefore, we set n_tree to 100 when using the 90 key drivers for RF modelling for Ta estimation for each year.

Among the 90 key drivers for estimating Ta, 78 were meteorological drivers, while 10 were landscape drivers (see Table S2). Meanwhile, current time and elevation were also selected as key drivers. When estimating the Ta using these 90 drivers, the five most important drivers and their contributions were 38.0% for RH_1Hours_Before, 10.9% for Current_time (hour), 5.0% for mw03_ai (Aggregation Index at a radius of 300 m), 4.4% for elevation and 3.8% for RH_22Hours_Before. In total, they contributed 62.1% of the importance. It is evident from this that meteorology, landscape, time and topography all play a non-negligible role in estimating the spatial distribution of Ta.

We then conducted RF modelling to estimate Ta for each year using the determined key drivers and n_tree. The accuracy of the RF models and the number of weather stations for each year are shown in Table 2. The R^2 and oob_score, calculated using the test samples

Fig. 3. The relationship between n_tree and R^2 in RF modelling for the Ta estimation by (a) modelling with all preliminary drivers and (b) modelling with the 90 key drivers.

and out-of-bag samples, respectively, show similar accuracies, with mean values of 0.8723 and 0.8652. The values of the RMSE for each year range from 0.9648 °C to 1.3028 °C, with a mean value of 1.1160 °C. At the same time, MAE values ranged from 0.7147 °C to 0.9838 °C, with a mean value of 0.8227 °C. Overall, the accuracy metrics show the RF models for estimating Ta achieved good accuracy.

Further, we validated the accuracy of the RF models for estimating Ta at different hours. Fig. 4 shows the R^2 values of the RF models for estimating Ta for different years at different hours. The RF models of different years showed similar performance. In general, the RF models achieved satisfactory R^2 at different hours. Still, relatively, the models maintained a more stable and higher R^2 during nighttime (i.e. 19:00 to 5:00). A decrease in R^2 occurred during the sunrise period (i.e. 6:00 to 9:00), and a steady recovery followed. Scatter plots of observed versus predicted Ta values for each hour based on the test samples for each year are shown in Figs. S1-S11. They all exhibit excellent linear relationships distributed around the 1:1 diagonal.

Fig. 5 shows the RMSE of the RF models for estimating Ta for different years at different hours. Similar to R (Stott, 2016), the RF model's RMSE performed relatively better at night. Moreover, from 6:00 to 9:00, RMSE showed an increase. However, the difference is that the RMSE remains steadily at relatively high values for most of the subsequent daytime hours. It was until dusk (i.e. 17:00 to 19:00) that the RMSE fell back to a lower value. An almost identical trend is seen in the MAE, shown in Fig. S12.

3.2. Validation of the hourly RH maps

As with the modelling process for estimating Ta, in the RF modelling for estimating RH, we also conducted a key driver selection based on importance assessment and a search for the optimal n_tree using 2018 data. The modelling results for estimating RH using all preliminary driving factors showed that the R^2 of the RF model no longer improved significantly when the n_tree exceeded 100 (Fig. 6 (a)). Therefore, we selected the top 90 drivers in the importance assessment when n_tree was 100 as key drivers. The importance scores of these 90 key drivers (1.083) accounted for 98.4% of the scores of all preliminary drivers (1.100). Then, we found the optimal n_tree based on these 90 key drivers. Then, we find the optimal n_tree based on these 90 key drivers. Fig. 6(b) shows that after the n_tree exceeds 100, the R^2 value of the model shows an oscillation without significant improvement. Therefore, we set the n_tree of the RF model to 100 and applied it to the modelling for estimating RH for other years using these 90 key drivers.

Of these 90 key drivers for RH modelling, 83 were meteorological drivers, and 6 were landscape drivers, while current time was also included (see Table S3). The top five most important drivers were Ta_1Hours_Before, Ta_2Hours_Before, Ta_22Hours_Before, Current_time(hour) and mw02_frac_mn (Mean of Fractal Dimension Index at a radius of 200 m), which contributed 55.3%, 5.3%, 3.0%, 3.0% and 2.6% importance, respectively. The top five drivers show that temperature, time and landscape are critical for estimating RH, as they together contribute 69.3% importance. However, unlike the key drivers for Ta, elevation did not feature as a key driver for RH.

The selected key drivers were used to estimate RHs for different years. Table 3 exhibits the accuracy of the models in different years. The R^2 values calculated using the test samples ranged from 0.7660 to 0.8258 with a mean value of 0.7970, while the oob_score calculated using the out-of-bag samples was similar with a mean value of 0.7870. In addition, the RMSE values for these RF models ranged from 4.8207% to 5.7257%, with a mean value of 5.3816%. The MAE values ranged from 3.4239% to 4.1171%, with a mean value of 3.8641%. The accuracy validation results reflect the reliability of the RF models used to estimate RH across years.

Furthermore, we validated the performance of the RF model for estimating RH at different hours. Fig. 7 shows the R^2 of the RF models for estimating RH for different years at different hours. It can be seen that the R^2 curves for the RH modelling show a consistent trend across the different years but with more fluctuations than Ta's. During the late night hours (i.e. 23:00 to 6:00), R^2 remained more stable at relatively high values, but during the sunrise period (i.e. 7:00 to 8:00), R^2 showed a decrease. However, the R^2 values then climbed up to reach a maximum at 16:00 to 17:00, even higher than those during the nighttime. At dusk and in the subsequent period (i.e. 18:00 to 22:00), the R^2 values again decreased and fluctuated. In addition, scatter plots of the hourly observed versus predicted RH values for each year of the test samples are shown in Figs. S13-S23. They all exhibit a good linear relationship distributed around the 1:1 diagonal.

Regarding error metrics, the RF models for estimating RH had similar RMSE and MAE values and trends at different hours. Fig. 8 presents the RMSE of the RF models for estimating RH for different years at different hours, while the performance of the MAE is shown

Table 2
The accuracy of the RF models for estimating Ta for each year.

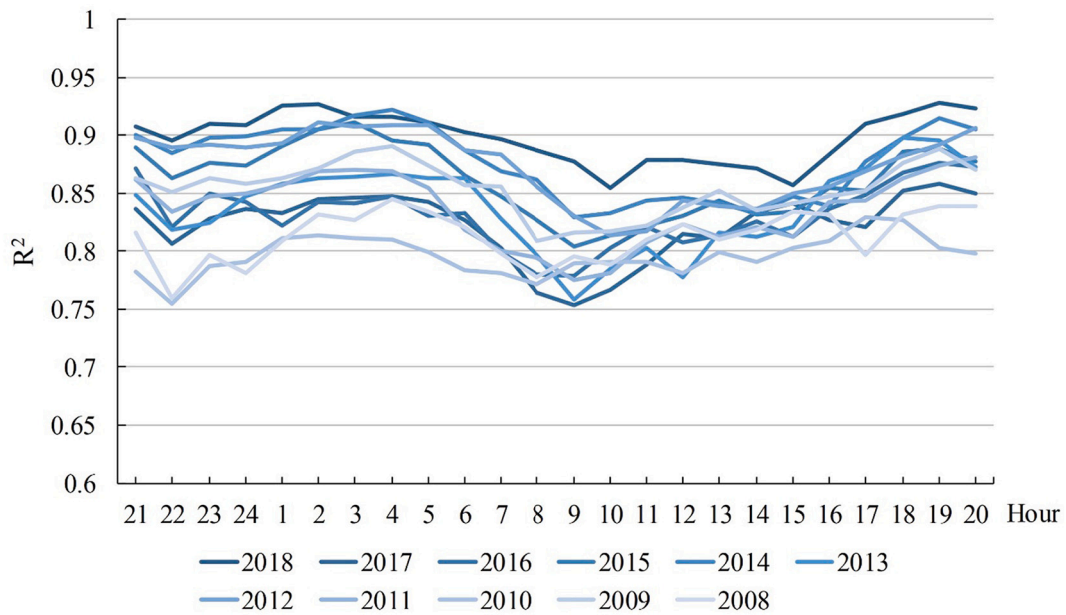| Year	R^2	oob_score	RMSE (°C)	MAE (°C)	Numbers of weather station
2018	0.9156	0.9070	0.9676	0.7147	44
2017	0.8528	0.8456	1.2028	0.8791	45
2016	0.8635	0.8530	1.1379	0.8421	46
2015	0.8854	0.8816	1.0088	0.7413	46
2014	0.8926	0.8881	1.0522	0.7589	44
2013	0.8602	0.8521	1.1589	0.8473	44
2012	0.8916	0.8862	0.9648	0.7298	44
2011	0.8672	0.8630	1.1475	0.8275	44
2010	0.8363	0.8214	1.3028	0.9838	42
2009	0.8815	0.8781	1.1205	0.8356	39
2008	0.8490	0.8415	1.2125	0.8898	36
Mean	0.8723	0.8652	1.1160	0.8227	–

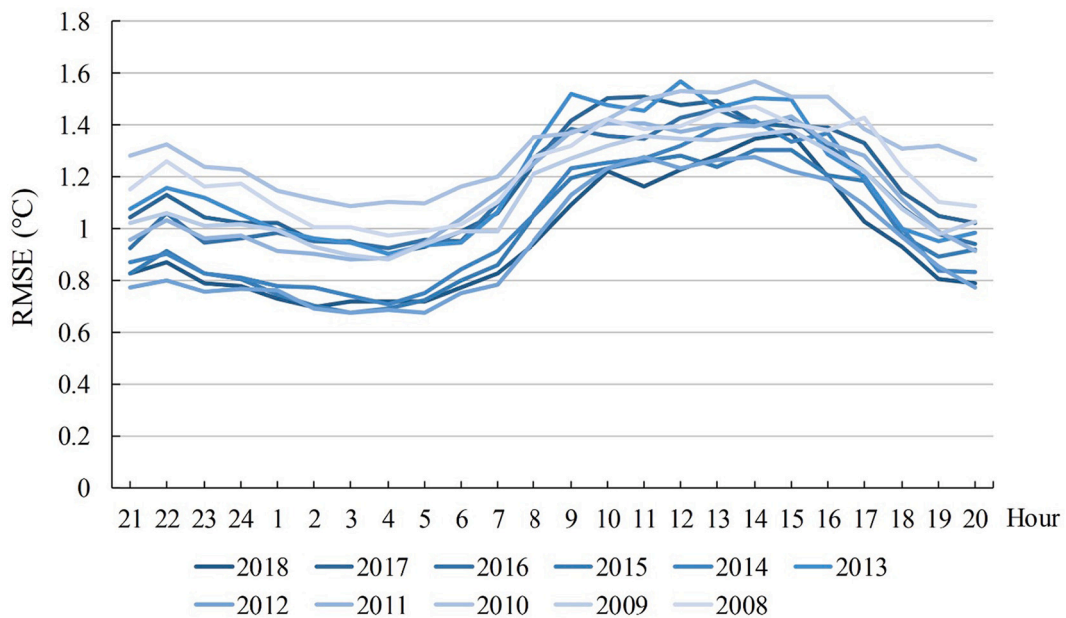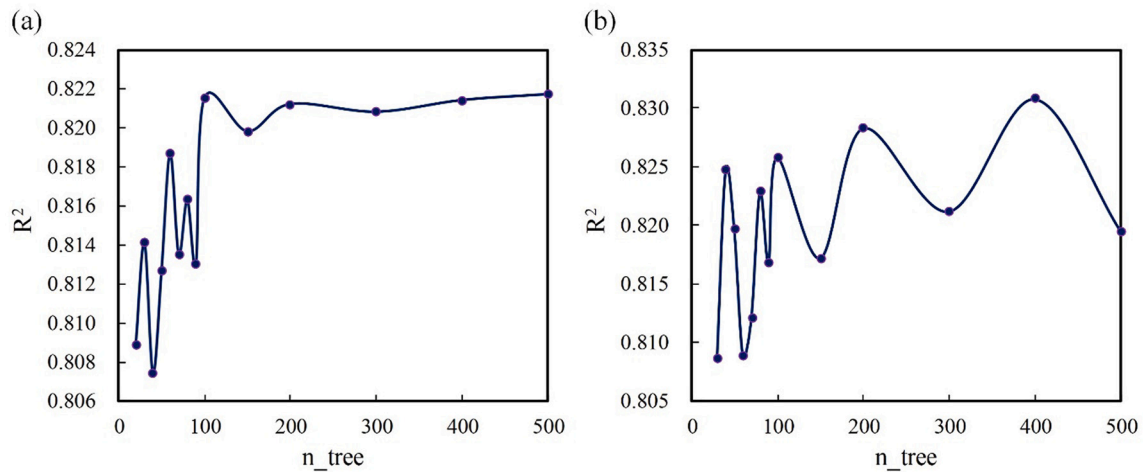**Fig. 4.** R^2 of the RF models for estimating Ta for different years at different hours.

Fig. 5. RMSE of the RF models for estimating Ta for different years at different hours.

in Fig. S24. Here we take the performance of RMSE as an example. Similar to R (Stott, 2016), the RMSE of the RF models for estimating RH performs better and keeps stable during the late night hours (i.e. 23:00 to 6:00). However, during the sunrise period (i.e. 7:00 to 9:00), the RMSE value gradually increased. In addition, it remained high for most hours in the daytime (i.e. 9:00 to 15:00). As dusk approaches and beyond(i.e. 16:00 to 22:00), the RMSE value gradually decreases and fluctuates somewhat.

3.3. Spatial performance of the hourly Ta maps

In this study, we also explored the spatial performance of the RF model for estimating Ta. For presentation purposes, we aggregated all estimated hourly Ta maps by different hours at mean values. In addition, we compared the Ta pattern using the conventional Kriging interpolation, which is based on the mean Ta values of each weather station at different hours. For the 24 h of the day, we chose

Fig. 6. The relationship between n_tree and R^2 in RF modelling for the RH estimation by (a) modelling with all preliminary drivers and (b) modelling with the 90 key drivers.

Table 3
The accuracy of the RF models for estimating RH for each year.

Year	R^2	oob_score	RMSE (%)	MAE (%)
2018	0.8258	0.8148	5.3915	3.8074
2017	0.7937	0.7804	5.2699	3.7796
2016	0.7749	0.7644	5.4954	4.0460
2015	0.7910	0.7836	5.5663	3.9594
2014	0.7660	0.7514	5.5258	3.8667
2013	0.8011	0.7947	4.8690	3.5539
2012	0.7994	0.7831	5.4583	4.0021
2011	0.8048	0.7986	5.4245	3.9064
2010	0.8155	0.811	4.8207	3.4239
2009	0.7916	0.7837	5.6502	4.1171
2008	0.8031	0.7915	5.7257	4.0428
Mean	0.7970	0.7870	5.3816	3.8641

three representative hours for the presentation: 4:00, which is the hour before sunrise, when temperatures and human activity drop to their lowest level after a night; 14:00, which is generally the hottest hour in a day; and 21:00, which is shortly into the night, when temperatures have dropped, but people are usually still active.

Fig. 9 shows the difference between the Ta spatial patterns estimated by the RF model and the Kriging interpolation in the 2018 summer season at the representative hours. It can be seen that the Ta maps estimated by the RF model provide more spatial detail, whereas the Ta maps assessed by kriging interpolation do not. Furthermore, as the topography effect is not considered, the Ta spatial pattern with Kriging interpolation only correlates with the geospatial distribution of weather stations, i.e. there is more Ta spatial detail where the weather stations are denser. Nevertheless, comparing the LCZ maps, we can see that the Ta spatial pattern estimated by the RF model shows a correlation out of topography. That is, Ta is lower in hilly areas covered by dense trees (LCZ A) and vegetation (LCZ D) than in lowlands.

Fig. 9 also zooms in on the Kowloon Peninsula, one of the core urban areas of Hong Kong. Comparing the LCZ maps reveals that Ta is higher in the core urban area, with a mixture of different types of buildings. Furthermore, comparing the Ta spatial pattern at 21:00 and 4:00, it can be seen that the core urban area cools more slowly at night than the fringes of the urban area.

3.4. Spatial performance of the hourly RH maps

We also explored the spatial performance of the hourly RH maps estimated by the RF model. As with the hourly Ta maps, we calculated the mean values of the RH maps at different hours for presentation by aggregation. The RH maps evaluated with Kriging interpolation using the mean RH values for each weather station at different hours were still used as controls. Similarly, 21:00, 4:00 and 14:00 have been chosen as representative hours.

Fig. 10 shows the difference between the RF model and the Kriging interpolation in estimating the spatial pattern of RH in the 2018 summer season. Overall, the RH maps ars not as spatially heterogeneous as the Ta maps. Also, the RH spatial patterns estimated by the RF model and the Kriging interpolation are generally similar. However, because it only considers RH values at weather station locations, the Kriging interpolation shows some linear transitional features in the spatial distribution of RH where weather stations are
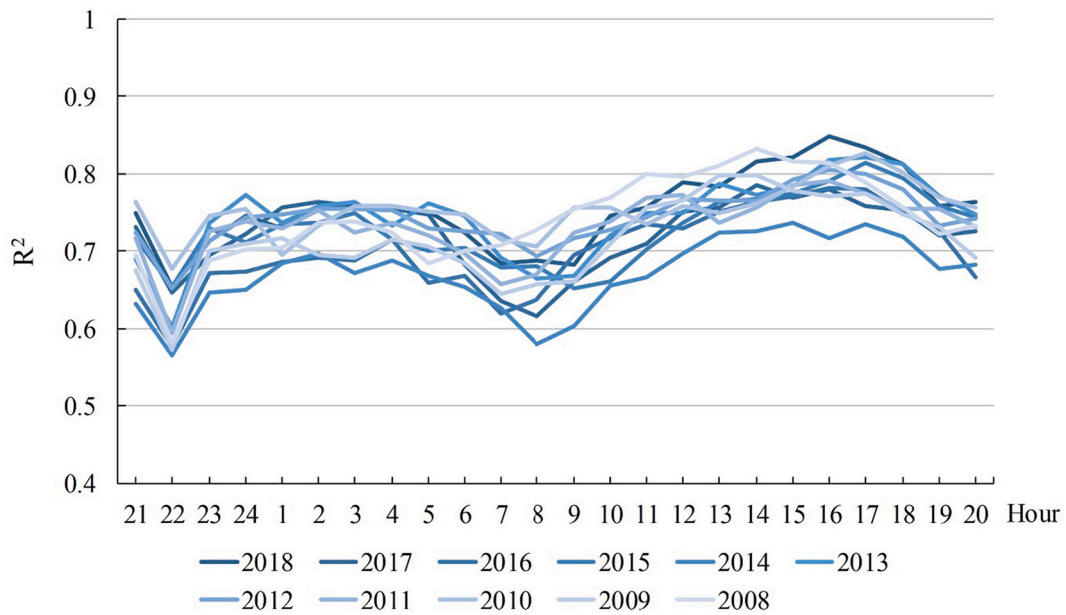
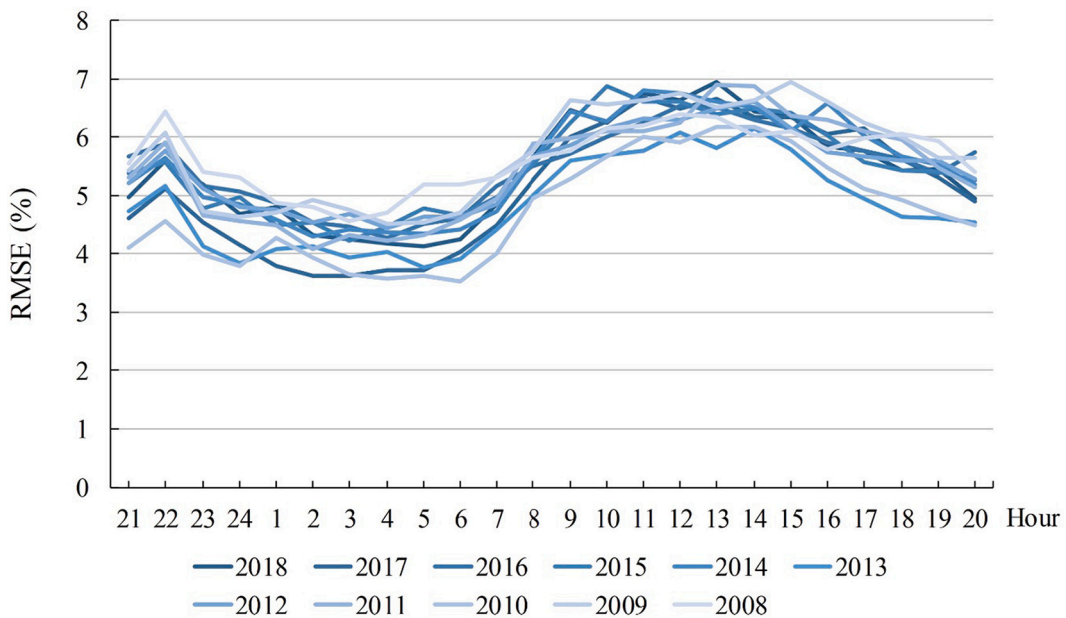**Fig. 7.** R^2 of the RF models for estimating RH for different years at different hours.

Fig. 8. RMSE of the RF models for estimating RH for different years at different hours.

relatively sparse. In contrast, the RF model for estimating RH provides more spatial detail due to considering landscape drivers. Moreover, the north-western part of Hong Kong becomes the area where the daily variation in RH is more pronounced, i.e. it is wetter there at 4:00 and drier at 14:00.

Furthermore, Fig. 10 zooms in on the central part of Hong Kong, where there are large areas of vegetation-covered mountains and densely populated towns. Comparing the RH map with the LCZ map shows that the RH values tend to be lower in the mountainous areas where the landscape patch pattern is simpler and where there are mainly continuous patches of dense trees (LCZ A) and low plants (LCZ D).
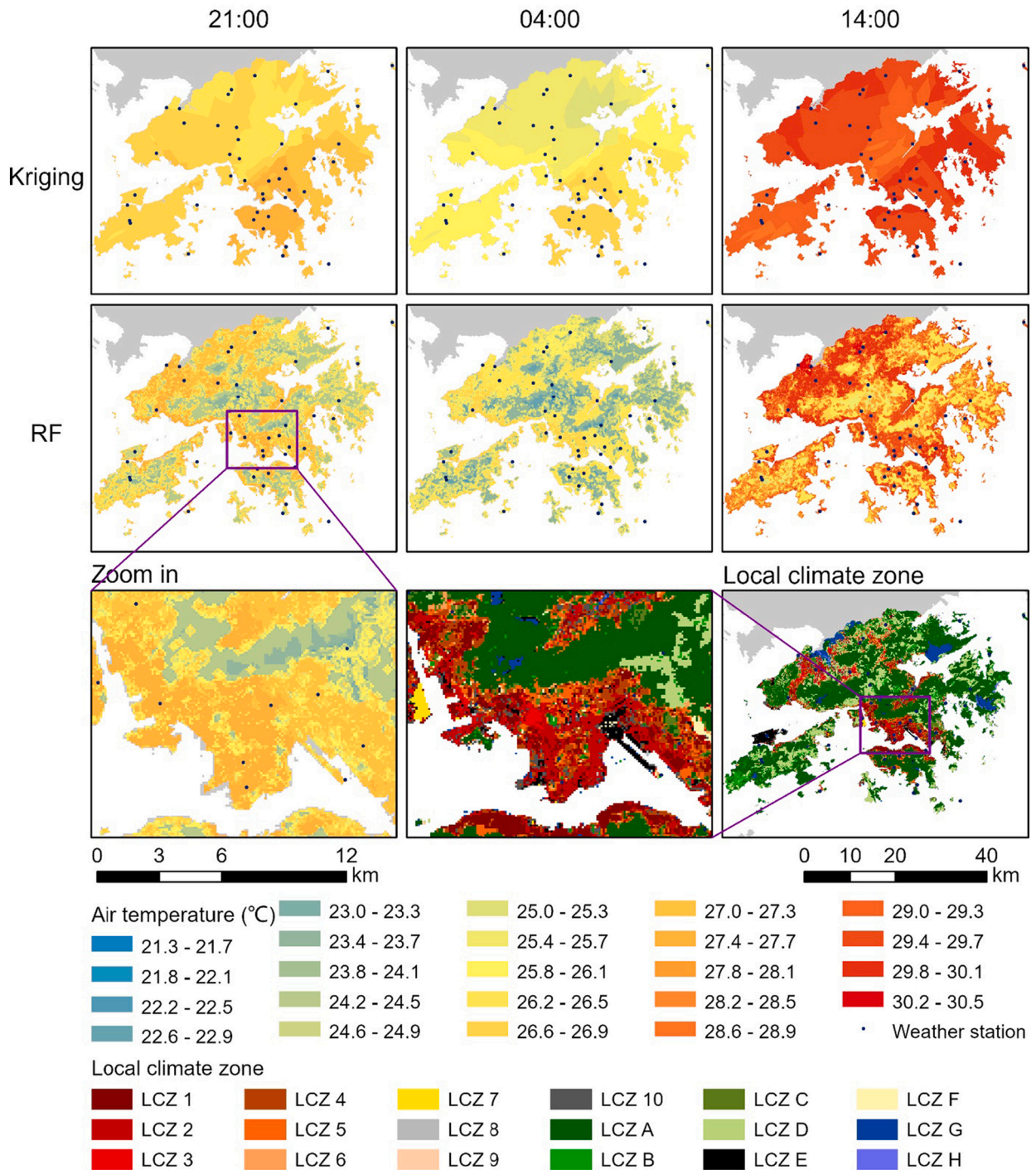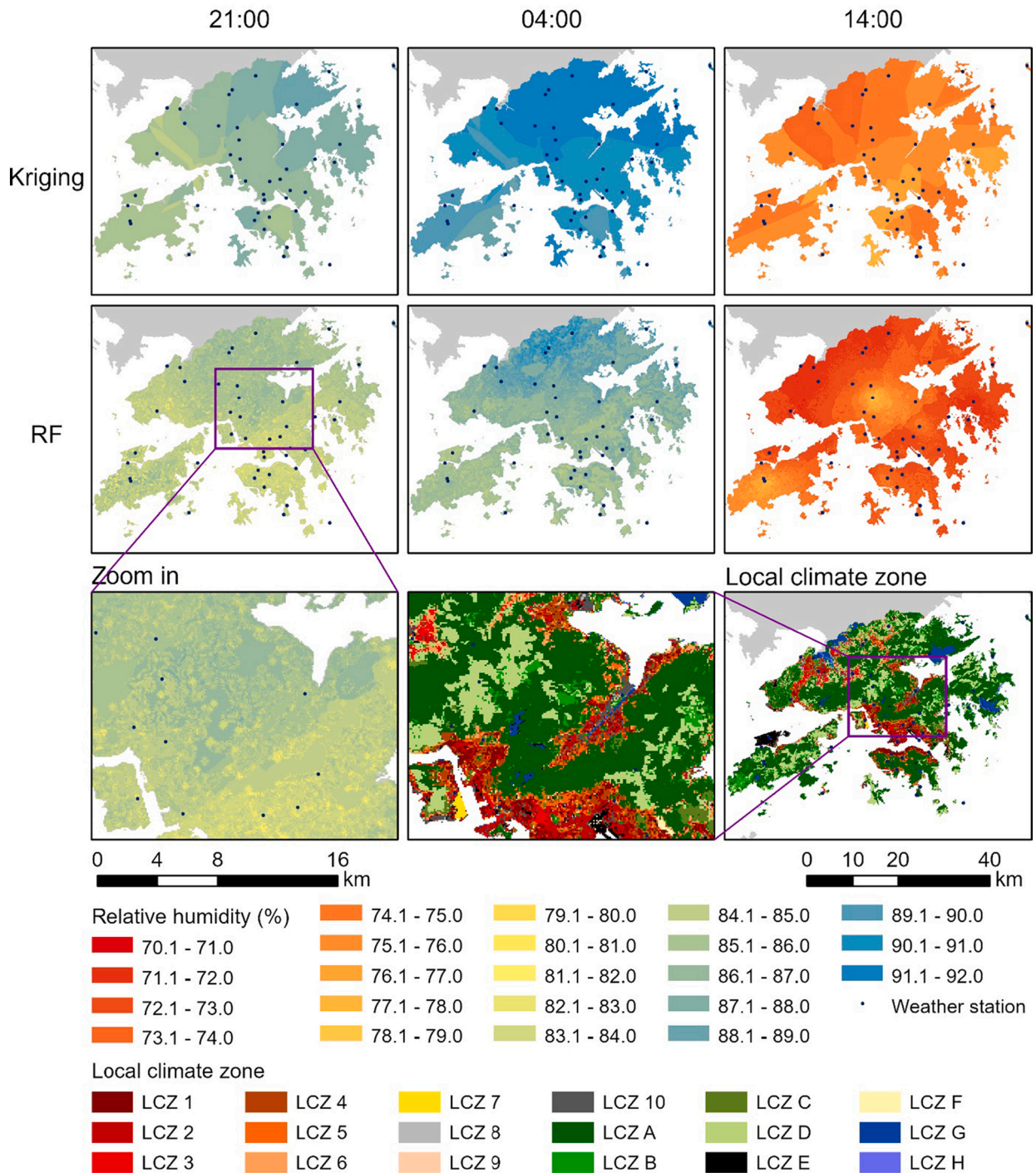
Fig. 9. The difference between the RF model and the Kriging interpolation in estimating the spatial pattern of Ta in the 2018 summer season.

3.5. The thermal index distribution in Hong Kong

Ta and RH provide different dimensions to portray the thermal environment, but we still need a thermal index as a composite indicator of the thermal environment. In this study, we employed the NET as the thermal index. Furthermore, NET is adopted by the HKO (Yip et al., 2007), so the NET-based thermal comfort map could be a useful reference to HKO. NET reflects the human thermal comfort considering the combined effect of Ta, RH and wind speed. As wind speed maps are not provided in this study, the NET map obtained here can be considered the most stressful level of human thermal comfort felt in a calm wind condition during the summer
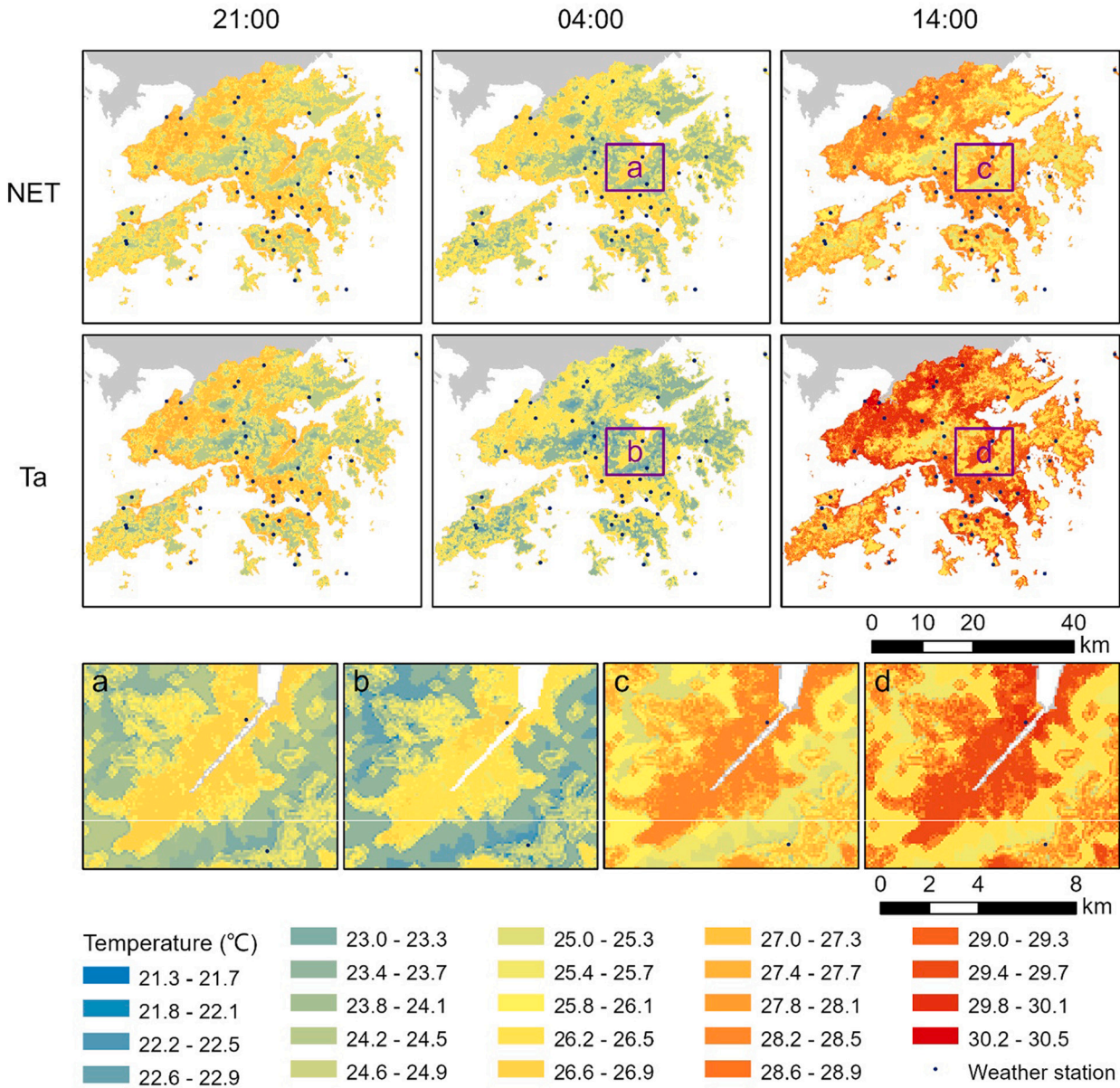
Fig. 10. The difference between the RF model and the Kriging interpolation in estimating the spatial pattern of RH in the 2018 summer season.

season.

Fig. 11 shows the difference between NET and Ta in depicting the spatial pattern of the thermal environment in the 2018 summer season. As it considers the role of RH, NET shows a subtle difference from Ta at late night (4:00) and midday (14:00). In the north-western part of Hong Kong, for example, NET is higher than Ta at 4:00 and lower than Ta at 14:00 because it is more humid at late night and drier during the day there compared to other areas. It means that people there will feel hotter than the actual air temperature at night, while the thermal comfort will be relatively relieved during the daytime.

Fig. 11 also zooms in on the Shatin District, which has a long, narrow built-up area surrounded by hills on three sides. Here, we can

Fig. 11. The difference between NET and Ta in depicting the spatial pattern of the thermal environment in the 2018 summer season.

still find the pattern that the NET of the built-up area is higher than Ta at night (Fig. 11 (a) and (b)), while the opposite is observed during the daytime (Fig. 11 (c) and (d)). Moreover, the difference between the NET of the hilly area and the built-up area is more significant than that of Ta at night (Fig. 11 (a) and (b)), while it is smaller than Ta during daytime (Fig. 11 (c) and (d)). In other words, at night, people feel the built-up area hotter than the hilly area more strongly than the difference in actual air temperature. In contrast, during the daytime, people perceive the difference in temperature between the built-up area and the hilly area to be smaller than it actually is.

4. Discussion

4.1. Differences in driving factors in estimating Ta vs RH

As aforementioned, the meteorological drivers played the most important role in the RF model for estimating Ta, accounting for 78 of the 90 key drivers (Table S2). Of these meteorological drivers, RH, PRS and VV2 for the previous 24 h were all included, while PRE was selected for only 6 h. Regarding importance, RH, PRS, VV2 and PRE contributed 52.4%, 12.3%, 4.0% and 1.6%, respectively.

Besides, the landscape drivers, current time (hour), and elevation contributed 14.4%, 10.9% and 4.4%, respectively.

The meteorological drivers also played the most critical role in the RF model for estimating RH (Table S3). Ta, PRS and VV2 for the previous 24 h were all included, while PRE was also included for 11 h. In the importance assessment, Ta, PRS, VV2 and PRE contributed 70.8%, 9.7%, 6.4% and 2.4%, respectively. Simultaneously, the landscape drivers and current time contributed 7.7% and 3.0%, respectively.

We can find that Ta and RH are the most critical drivers of each other, which reflects their clear correlation. However, Ta has a greater influence on RH than RH has on Ta. In addition, the landscape has different influence degrees on Ta and RH. Ta is more influenced by the landscape, including a non-negligible influence of Ta by altitude. This is reflected in the higher correlation of the hourly Ta distribution with the LCZ-based landscape and topography.

4.2. The performance difference between daytime and nighttime of the models

Overall, the RF models for estimating Ta and RH both perform better during nighttime, especially at late night (23:00 to 5:00). Then, they both show a decrease in accuracy during the sunrise period (6:00 to 9:00). Moreover, their estimation errors remain at a high level for most of the subsequent daytime period (10:00 to 15:00). However, the R^2 of the model for estimating RH gradually increases during this period as RH differences between regions become larger during daytime. The models' errors decrease progressively until after the sun gradually sets (after 16:00). Similar better model performance during nighttime has been seen in previous studies of hourly Ta mapping (Zhang and Du, 2022). Thus, we infer that solar radiation is one of the main factors affecting Ta and RH mapping accuracy, as it is the most important source of energy driving weather variability.

4.3. Pattern variability of Ta and RH

In general, the spatial pattern of Ta shows a higher correlation with the LCZ-based landscape. Ta is lower in hilly areas than in built-up areas in the lowlands. Whereas the spatial pattern of RH is less sensitive to the response of the landscape, only slightly tending to be wetter in hilly areas with simpler landscapes. In addition, the value range of RH is greater during daytime than nighttime. Nevertheless, both Ta and RH show the most remarkable influence by the landscape within a 200–300 m radius. This suggests that the scale effects affecting the thermal environment can be explored in depth in the future, facilitating more effective urban planning to improve the thermal environment.

Furthermore, the NET index provides a good overview of the effects of Ta and RH on human thermal comfort. It reflects some subtle differences that are difficult to perceive when looking at Ta or RH alone. For example, people in built-up areas feel more stressful human thermal comfort than actual air temperature during nighttime in the summer season. This suggests that combining the hourly Ta and RH maps can provide additional information for assessing human thermal comfort in cities.

4.4. Comparison to other Ta and RH mapping

The accuracy of our hourly Ta and RH mapping is comparable to that of other products. For Ta mapping, the ML-based hourly Ta mapping was reported to achieve RMSE of 0.8–1.9 °C and MAE of 0.6–1.5 °C (Zhou et al., 2020; Zhang and Du, 2022), while on a daily scale, these two metrics were reported to be 2.0 °C and 1.5 °C (Shen et al., 2020). In contrast, our hourly Ta mapping has RMSE and MAE of 1.1160 °C and 0.8227 °C, respectively, which achieves comparable or even better performance. For RH mapping, on the other hand, an hourly RH mapping reported RMSEs of 11.3–19.7% (Bregaglio et al., 2010); an ML-based daily RH mapping could achieve an R^2 of 0.71 (Hanoon et al., 2021), while a summer RH mapping reported RMSE and MAE of 7.4% and 2.4%, respectively (Li and Zha, 2018). In contrast, our hourly RH mapping has R (Stott, 2016), RMSE and MAE of 0.7970, 5.3816% and 3.8641%, respectively. Overall, both our Ta and RH mapping achieved comparable and satisfactory accuracy.

4.5. Potential application

The multi-year summer season hourly thermal comfort dataset proposed in this study contains Ta, RH and NET maps, which allows it to be applied to many relevant areas. For example, it can be used to explore the long-term relationship between urban morphology and the thermal environment and to capture the trend of urban heating (Zhao et al., 2020). Moreover, this dataset has fine spatio-temporal resolution and focuses on high-density urban areas. Thus, it can be combined with big data on human activity trajectories to explore real-time heat exposure (Epstein and Moran, 2006; Varghese et al., 2019) and heat-related health risk assessment (Bhaskaran et al., 2012; Xu et al., 2019) at a fine spatiotemporal scale. In addition, this dataset can be used for thermal comfort assessment, which can further guide urban planning, building design, and the assessment of cooling energy consumption (Braun et al., 2014; Ihara et al., 2008).

4.6. Limitations and future work

Despite all our efforts, there are still some limitations in this study. First, the spatial pattern of meteorological drivers was obtained by Kriging interpolation. In the future, finer meteorological spatial data could reduce the estimation error due to interpolation. Second, the LCZ maps that calculated the landscape drivers introduced their errors. The increasing accuracy of LCZ classification in the future will reduce this part of the error. Third, the previous inferred that solar radiation might be an essential driver of hourly Ta and RH

mapping, but this study did not consider the drivers directly related to solar radiation. Because Hong Kong is located in a monsoonal climate zone with clouds during most of the summer season, multispectral remote sensing images that observe the land surface at a near-hourly scale are difficult to obtain. In addition, the machine learning model for thermal comfort mapping could be applied to the whole year and other regions in the future to explore the robustness and generalisation of the model.

5. Conclusions

In this study, we present an hourly thermal comfort dataset for Hong Kong, including Ta and RH using a machine learning algorithm (RF regression model), which covers the summer season (May to September) from 2008 to 2018. Since considering the meteorological driver and the LCZ-based landscape driver, the thermal comfort dataset provides fine spatial details at 100 m resolution, especially the spatial pattern of hourly Ta presents a closer relationship with the landscape. Moreover, the validation results show excellent accuracy of the thermal comfort dataset in both hourly Ta and RH estimations. The mean values of R (Stott, 2016), RMSE and MAE for Ta estimation achieved 0.8723, 1.1160 °C and 0.8227 °C, respectively, while those for RH estimation reached 0.7970, 5.3816% and 3.8641%. Further, the NET calculated by combining Ta and RH gives a more relevant picture of human thermal comfort. For example, it can reveal that people in built-up areas feel hotter than the actual temperature at night, thus presenting a more serious challenge to human thermal comfort. In the future, the hourly thermal comfort dataset can be combined in depth with studies on big data, health, energy consumption, urban planning and more to explore the socio-economic and human impact of the thermal environment.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.uclim.2022.101400.

References

Acero, J.A., Arrizabalaga, J., 2018. Evaluating the performance of ENVI-met model in diurnal cycles for different meteorological conditions. Theor. Appl. Climatol. 131, 455–469.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., de Moraes Gonçalves, J.L., 2013. Modeling monthly mean air temperature for Brazil. Theor. Appl. Climatol. 113, 407–427.

Bhaskaran, K., et al., 2012. Heat and risk of myocardial infarction: hourly level case-crossover analysis of MINAP database. BMJ 345.

Braun, M.R., Altan, H., Beck, S., 2014. Using regression analysis to predict the future energy consumption of a supermarket in the UK. Appl. Energy 130, 305–313.

Bregaglio, S., Donatelli, M., Confalonieri, R., Acutis, M., Orlandini, S., 2010. An integrated evaluation of thirteen modelling solutions for the generation of hourly values of air relative humidity. Theor. Appl. Climatol. 102, 429–438.

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

Chen, G., et al., 2022. Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning. Sci. Total Environ. 841, 156737.

Chevalier, R.F., Hoogenboom, G., McClendon, R.W., Paz, J.A., 2011. Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. Neural Comput. & Applic. 20, 151–159.

Chronopoulos, K.I., Tsiros, I.X., Dimopoulos, I.F., Alvertos, N., 2008. An application of artificial neural network models to estimate air temperature data in areas with sparse network of meteorological stations. J. Environ. Sci. Health Part A 43, 1752–1757.

Chung, L.C.H., Xie, J., Ren, C., 2021. Improved machine-learning mapping of local climate zones in metropolitan areas using composite earth observation data in Google earth engine. Build. Environ. 199, 107879.

Di Napoli, C., Barnard, C., Prudhomme, C., Cloke, H.L., Pappenberger, F., 2021. ERA5-HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. Geosci. Data J. 8, 2–10.

Dzwonkowski, B., et al., 2020. Compounding impact of severe weather events fuels marine heatwave in the coastal ocean. Nat. Commun. 11, 1–10.

Emmanuel, R., 2021. Urban Heat Island Mitigation Technologies. MDPI.

Epstein, Y., Moran, D.S., 2006. Thermal comfort and the heat stress indices. Ind. Health 44, 388–398.

Florio, E.N., Lele, S.R., Chi Chang, Y., Sterner, R., Glass, G.E., 2004. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. Int. J. Remote Sens. 25, 2979–2994.

Forman, R.T.T., 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press.

Fricke, C., Pongr Cz, R., Unger, J.A.N., 2022. Comparison of daily and monthly intra-urban thermal reactions based on LCZ classification using surface and air temperature data. Geogr. Pannon. 26.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189-1232.

Fries, A., Rollenbeck, R.U.T., Nau Ss, T., Peters, T., Bendix, J.O.R., 2012. Near surface air humidity in a megadiverse Andean mountain ecosystem of southern Ecuador and its regionalization. Agric. For. Meteorol. 152, 17–30.

Hanoon, M.S., et al., 2021. Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia. Sci. Rep. 11.

Hentschel, G., 1986. A human biometeorology classification of climate for large and local scales. In: Proc. WMO/HMO/UNEP Symposium on Climate and Human Health, pp. 22–26. Leningrad.

Hodam, S., Sarkar, S., Marak, A.G., Bandyopadhyay, A., Bhadra, A., 2017. Spatial interpolation of reference evapotranspiration in India: comparison of IDW and Kriging Methods. J. Inst. Eng. (India) Ser. A 98, 511–524.

Ihara, T., Genchi, Y., Sato, T., Yamaguchi, K., Endo, Y., 2008. City-block-scale sensitivity of electricity consumption to air temperature and air humidity in business districts of Tokyo, Japan. Energy 33, 1634–1645.

Kamusoko, C., Gamba, J., 2015. Simulating urban growth using a random Forest-cellular automata (RF-CA) model. ISPRS Int. J. Geo Inf. 4, 447–470.

Khan, A., Chatterjee, S., 2016. Numerical simulation of urban heat island intensity under urban–suburban surface and reference site in Kolkata, India. Model. Earth Syst. Environ. 2, 1–11.

Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2014. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. Remote Sens. Environ. 150, 132–139.

Lamarque, J., et al., 2013. The atmospheric chemistry and climate model intercomparison project (ACCMIP): overview and description of models, simulations and climate diagnostics. Geosci. Model Dev. 6, 179–206.

Lau, K.K., Chung, S.C., Ren, C., 2019. Outdoor thermal comfort in different urban settings of sub-tropical high-density cities: an approach of adopting local climate zone (LCZ) classification. Build. Environ. 154, 227–238.

Lee, S., Choi, H., Cha, K., Chung, H., 2013. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: gasoline and naphtha. Microchem. J. 110, 739–748.

Li, P.W., Chan, S.T., 2000. Application of a weather stress index for alerting the public to stressful weather in Hong Kong. Meteorol. Appl. 7, 369–375.

Li, L., Zha, Y., 2018. Mapping relative humidity, average and extreme temperature in hot summer over China. Sci. Total Environ. 615, 875–881.

Li, T., et al., 2014. Mapping near-surface air temperature, pressure, relative humidity and wind speed over Mainland China with high spatiotemporal resolution. Adv. Atmos. Sci. 31, 1127–1135.

Li, D., Yuan, J., Kopp, R.E., 2020. Escalating global exposure to compound heat-humidity extremes with warming. Environ. Res. Lett. 15, 64003.

Li, L., et al., 2022. Variabilities of land surface temperature and frontal area index based on local climate zone. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 2166–2174.

Lin, X., et al., 2016. Empirical estimation of near-surface air temperature in China from MODIS LST data by considering physiographic features. Remote Sens. 8, 629.

Liu, W., Ji, C., Zhong, J., Jiang, X., Zheng, Z., 2007. Temporal characteristics of the Beijing urban heat island. Theor. Appl. Climatol. 87, 213–221.

Ma, F., Yuan, X., 2021. More persistent summer compound hot extremes caused by global urbanization. Geophys. Res. Lett. 48, e2021G–e93721G.

Maharjan, S., Regmi, R.P., 2015. Grid based temperature and relative humidity distribution map of the Kathmandu Valley. J. Inst. Sci. Technol. 19, 7–13.

Matzarakis, A., Amelung, B., 2008. Physiological equivalent temperature as indicator for impacts of climate change on thermal comfort of humans. In: Seasonal Forecasts, Climatic Change and Human Health. Springer, pp. 161–172.

Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F., Shiri, J., 2020. Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. Ish J. Hydraul. Eng. 26, 376–386.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press.

Ren, J., et al., 2022. Exploring thermal comfort of urban buildings based on local climate zones. J. Clean. Prod. 340, 130744.

Rosenthal, J.K., 2010. 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.

Shen, H., et al., 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens. Environ. 240, 111692.

Shojaei, P., et al., 2017. Effect of different land cover/use types on canopy layer air temperature in an urban area with a dry climate. Build. Environ. 125, 451–463.

Simon, H., 2016. Modeling Urban Microclimate: Development, Implementation and Evaluation of New and Improved Calculation Methods for the Urban Microclimate Model ENVI-Met. Mainz, Univ. Diss.

de Souza, D.O., Dos Santos, Alval A., R. C. E. L., 2014. Observational evidence of the urban heat island of Manaus City, Brazil. Meteorol. Appl. 21, 186–193.

Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93, 1879–1900.

Stott, P., 2016. How climate change affects extreme weather events. Science 352, 1517–1518.

Tobler, W., 2004. On the first law of geography: A reply. Ann. Assoc. Am. Geogr. 94, 304–310.

Varghese, B.M., et al., 2019. Characterising the impact of heatwaves on work-related injuries and illnesses in three Australian cities using a standard heatwave definition-excess heat factor (EHF). J. Expos. Sci. Environ. Epidemiol. 29, 821–830.

Wang, M., et al., 2017. Comparison of spatial interpolation and regression analysis models for an estimation of monthly near surface air temperature in China. Remote Sens. 9, 1278.

Wu, X., et al., 2020. Quantitatively evaluating the effect of urbanization on heat waves in China. Sci. Total Environ. 731, 138857.

Xu, D., et al., 2019. Acute effects of temperature exposure on blood pressure: an hourly level panel study. Environ. Int. 124, 493–500.

Yan, D., et al., 2020. Integrating remote sensing data with WRF model for improved 2-m temperature and humidity simulations in China. Dyn. Atmos. Oceans 89, 101127.

Yin, C., Yuan, M., Lu, Y., Huang, Y., Liu, Y., 2018. Effects of urban form on the urban heat island effect based on spatial regression model. Sci. Total Environ. 634, 696–704.

Yip, K.M., Leung, Y.K., Yeung, K.H., 2007. Long-Term Trend in Thermal Index and its Impact on Mortality in Hong Kong. Hong Kong Observatory Reprint.

Zhang, Z., Du, Q., 2022. Hourly mapping of surface air temperature by blending geostationary datasets from the two-satellite system of GOES-R series. ISPRS J. Photogramm. Remote Sens. 183, 111–128.

Zhang, J., Wu, L., 2011. Land-atmosphere coupling amplifies hot extremes over China. Chin. Sci. Bull. 56, 3328–3332.

Zhang, D., Liu, X., Lin, Z., Zhang, X., Zhang, H., 2020. The delineation of urban growth boundaries in complex ecological environment areas by using cellular automata and a dual-environmental evaluation. J. Clean. Prod. 256, 120361.

Zhao, C., Nan, Z., Cheng, G., 2005. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains, China. Ecol. Model. 189, 209–220.

Zhao, J., et al., 2020. Exploring difference in land surface temperature between the city centres and urban expansion areas of China's major cities. Int. J. Remote Sens. 41, 8965–8985.

Zhao, Z., Sharifi, A., Dong, X., Shen, L., He, B., 2021. Spatial variability and temporal heterogeneity of surface urban Heat Island patterns and the suitability of local climate zones for land surface temperature characterization. Remote Sens. 13.

Zhou, B., et al., 2020. Estimating near-surface air temperature across Israel using a machine learning based hybrid approach. Int. J. Climatol. 40, 6106–6121.