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

Guangzhao Chen[1,2], Junyi Hua[3], Yuan Shi[4], Chao Ren[1,*]

1. Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China
2. Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong SAR, China
3. School of International Affairs and Public Administration, Ocean University of China, Qingdao, China
4. Department of Geography & Planning, University of Liverpool, Liverpool, UK

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

Highlight

- We proposed an hourly thermal comfort dataset at 100m resolution for a high-density city during the summer season by machine learning.

- The dataset, which includes air temperature and relative humidity indicators, maintains satisfactory accuracy at different hours of the day, especially during nighttime.

- The dataset reveals that the core urban area cools more slowly at night than the fringes of the urban area.

- The thermal comfort index reveals that people in built-up areas are exposed to more severe thermal stress at night than the actual air temperature.

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

Guangzhao Chen[1,2], Junyi Hua[3], Yuan Shi[4], Chao Ren[1,*]

1. Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China
2. Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong SAR, China
3. School of International Affairs and Public Administration, Ocean University of China, Qingdao, China
4. Department of Geography & Planning, University of Liverpool, Liverpool, UK

*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

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 100m 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 R2, 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

29 comfort index, NET, reveals that people in built-up areas feel hotter than measured by Ta

30 during the night due to the urban heat island effect. We believe this newly developed thermal

31 comfort dataset can provide novel, reliable and fine-grained data support for urban climate

32 research areas such as urban heat islands, heat exposure, heat-related health risk assessment,

33 and urban energy consumption estimation.

34

Highlight

36 ● We proposed an hourly thermal comfort dataset at 100m resolution for a high-density city

37 during the summer season by machine learning.

38 ● The dataset, which includes air temperature and relative humidity indicators, maintains

39 satisfactory accuracy at different hours of the day, especially during nighttime.

40 ● The dataset reveals that the core urban area cools more slowly at night than the fringes of

41 the urban area.

42 ● The thermal comfort index reveals that people in built-up areas are exposed to more severe

43 thermal stress at night than the actual air temperature.

Keywords

45 Hourly mapping; Machine learning; High spatiotemporal resolution; Air temperature; Relative

46 humidity; Thermal comfort

47

1. Introduction

49 With urbanisation[1] and the frequent occurrence of extreme weather events due to climate

50 change[2], including hot extremes in summer [3,4], people living in high-density cities are

2

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 [5]. Typically, thermal comfort can be assessed and measured by combining vital meteorological indicators, such as air temperature (Ta) and relative humidity (RH) [6]. Using these thermal comfort indicators, one can conduct research in areas such as urban heat islands [7-9], compound extreme weather [10,11], energy consumption [12,13], and heat-related health risks [14,15]. 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 [16]. 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 [17]. Land surface temperature (LST) is often used as a proxy for the spatial distribution of air temperature [18], because satellite imagery can provide spatially continuous data for a large spatial extent. However, LST and air temperature still cannot be equated[19]. 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 [20]. 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 [21] and inverse distance weighted (IDW) interpolation [22]. The interpolation methods follow "the first law of geography", which states

3

76 that the near things are more related to the distant things [23]. They require that the sampling

77 points be distributed as evenly as possible across the study area and that only the planar

78 distances between them are considered [21]. Therefore, the uneven distribution of sampling points

79 in reality and the differences in geographical conditions can introduce errors. The second is

80 climate models, which can be subdivided into macro-scale (i.e. global or regional scale),

81 mesoscale and micro-scale climate models [24-26]. Climate models are a class of mechanistic

82 models that simulate the spatial and temporal variability of different elements through the

83 action of physical mechanisms [27]. Climate models can simulate the spatial pattern of Ta at the

84 hourly or minute temporal resolution, while their spatial resolution varies from 50 km to 1 m

85 as the scale of the model varies from large to small [24,25]. Macro-scale climate models with

86 coarse spatial resolution often require simplifying elements such as urban structure [24].

87 Mesoscale climate models have a more complex structure, considering land surface-

88 atmosphere interactions, so they can only cover horizontal scales of tens to hundreds of

89 kilometres with kilometre-level resolution [26]. As the scale is further downscaled, micro-scale

90 climate models with better spatial resolution make strict demands on computational resources,

91 computational time, model complexity and fine historical input data [25]. As a result, micro-scale

92 climate models with high spatial resolutions are difficult to use for city-scale Ta simulations [28].

93 In addition, the higher temporal resolution also means more demands on computing resources.

94 Therefore, it is difficult to combine both high spatial resolution and high temporal resolution

95 in urban-scale Ta mapping through climate models. The third is regression methods, which

96 estimate the distribution of Ta in space and time by establishing quantitative relationships

97 between Ta and the relevant elements. In addition to traditional simple or multiple linear

98 regression [29,30] including geographically weighted regression (GWR) [22], machine learning (ML)

99 is a popular approach. Typical ML include support vector regression (SVR) [31], artificial neural

100 network (ANN) [32], random forest (RF) [33] and gradient boosting decision tree (GBDT) [34].

101 Machine learning can fit non-linear correlations, providing a higher estimation accuracy than

102 traditional linear regression. Moreover, ML is more applicable to mining relationships with

103 multiple variables and large data volumes. In general, satellite remote sensing imagery

104 providing metrics such as reflectance, LST and Normalized Difference Vegetation Index

105 (NDVI) are used as variables to predict the Ta pattern[35,36].

106 RH is another crucial indicator of thermal comfort. Similar to Ta, the spatial distribution

107 of RH can be estimated by three routes: traditional spatial interpolation[37,38], climate models[28,39]

108 and regression[40-42]. Likewise, the advantages and disadvantages of each type of RH mapping

109 are the same as those of Ta mapping before mentioned. We found fewer studies on RH mapping

110 than Ta, especially RH mapping using machine learning. However, a recent study reported an

111 increase in the temporal resolution of RH mapping by machine learning to the daily scale, while

112 most RH mapping studies focus on the monthly scale[41]. This study only used historical RHs as

113 the driving factors for estimating the current RH.

114 Overall, previous studies of thermal comfort datasets have had a few limitations. First, the

115 high spatial resolution of thermal comfort datasets is not well balanced with the high temporal

116 resolution[43]. Second, the effect of LULC, which is a non-negligible factor influencing near-

117 surface local climatic conditions[20], is rarely considered in Ta and RH mapping. Third, the

118 thermal comfort dataset pays insufficient focus on high-density cities and summer seasons,

119 which are closely related to the well-being of large populations in the context of a warming

120 climate but require high spatial and temporal resolution.

121 Therefore, in this study, we aimed to generate a spatiotemporal hourly thermal comfort

122 dataset, including air temperature and relative humidity, at a 100m resolution for Hong Kong,

123 a typical high-density city. This dataset covers the summer season (May to September) from

124 2008 to 2018. To better depict the spatial pattern, we employed ML algorithms and

125 simultaneously considered the effects of meteorological drivers, landscape drivers and

126 topography. Notably, the calculation of the landscape driver is based on a detailed LULC

127 classification, local climate zone (LCZ), which subdivides the LULC types within the city.

2. Data and method

129 The workflow for generating the thermal comfort dataset in this study is shown in Fig.1.

130 The detailed data preparation and implementation methods are illustrated in the subsequent
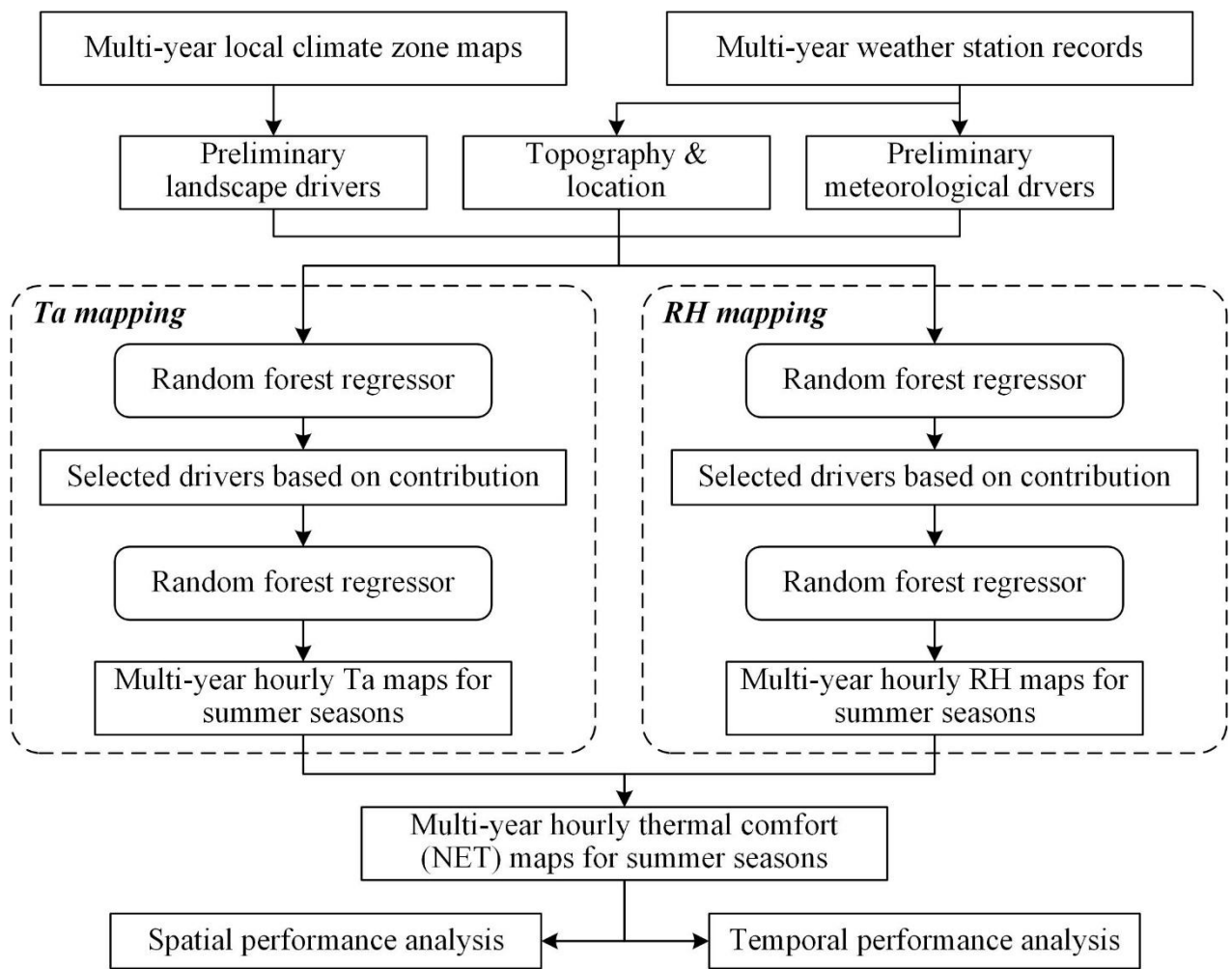
131 subsections.

132

133 Fig. 1 The workflow in this study

2.1. Study area

135 Hong Kong is located on the eastern side of the Pearl River estuary and is part of a

136 transitional region with a tropical and subtropical monsoon climate. It is bordered to the north

137 by Shenzhen, one of the four first-tier cities in mainland China. Hong Kong is an international

138 financial, shipping and trading centre with a developed economy, one of the core cities of the

139 Guangdong-Hong Kong-Macao Greater Bay Area and one of China's windows to the outside

140 world. Moreover, as Hong Kong is mainly hilly, with only about 20% of its land in the lowlands,

141 it makes Hong Kong one of the world's most famous high-density cities. Meanwhile, the

142 heterogeneity brought about by the complex topography also makes it an ideal study area for

143 applying machine learning for Ta and RH mapping with high spatial and temporal resolution.

144 2.2. Meteorological data (SY)

145 Meteorological data are provided by the Hong Kong Observatory (HKO) weather stations.

146 These weather stations are located throughout Hong Kong, providing hour-by-hour

147 meteorological observations from 2008 to 2018. Due to the development and management of

148 weather stations, the number of weather stations from 2008 to 2018 ranged from 36 to 46. Fig.

149 2 shows the distribution of these weather stations in 2018. These weather stations cover various

150 parts of Hong Kong well, with each station averaging 4005.1 metres from its nearest
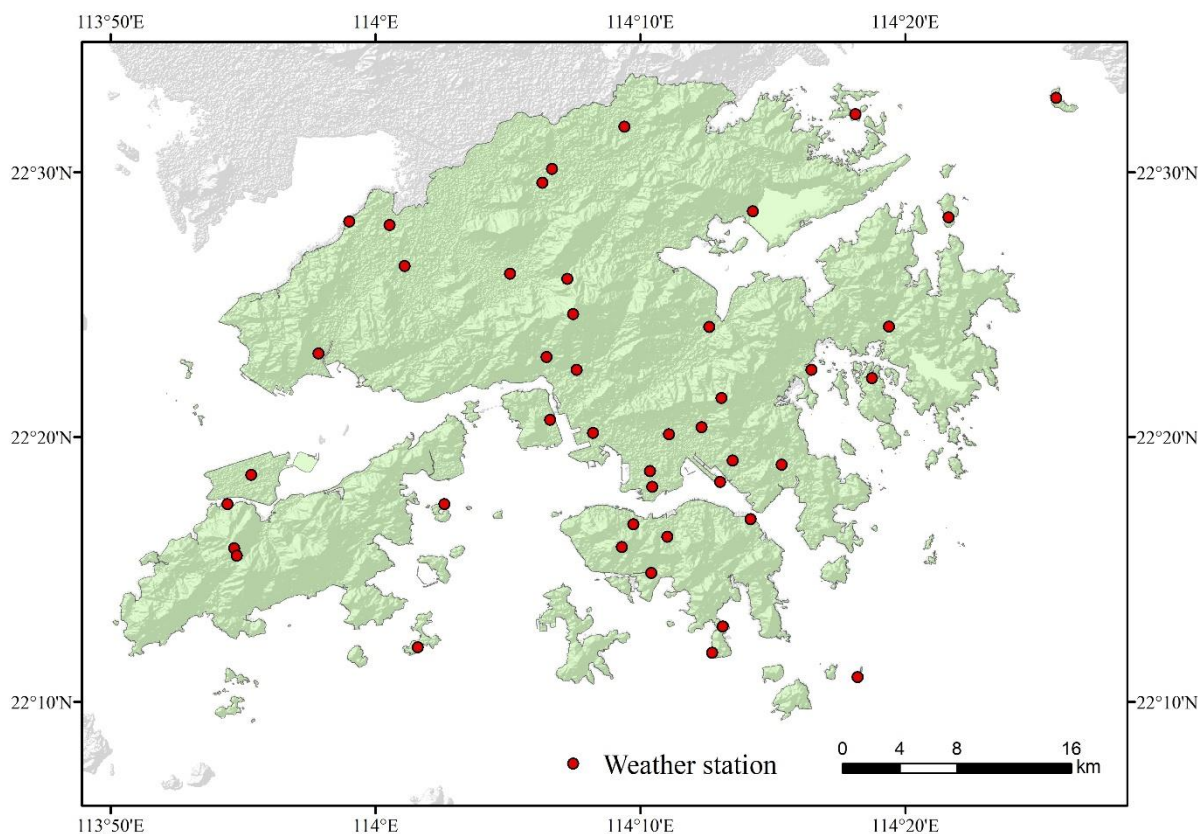
151 neighbouring station.

Fig. 2 Study area and the distribution of weather stations in 2018. Each red dot represents one weather station of Hong Kong Observatory

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.

2.3. LCZ data and landscape drivers

167   In previous studies, the LCZ has been demonstrated to be closely related to the urban

168   thermal environment, due to its unique definition of land cover types [44-48]. The LCZ typically

169   contains 17 land cover types [49], sometimes subdivided into an additional wetland type from the

170   water [50,51] (Table 1). It includes both urban-type LCZs and natural-type LCZs. However, unlike

171   traditional land cover classifications, LCZs subdivide urban-type land into ten categories based

172   on physical properties. These physical properties include building density, height, material and

173   area, which depict the urban morphology in detail and are closely linked to the near-surface

174   environment [49].

175   Table 1 LCZ types and simplified definitions in Chen et al. [50]

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#

176   The generation of LCZ datasets is not easy, especially for multi-year, as it requires

177   constructing multi-year sample sets for the numerous LCZ types. Therefore, in this study, we

178   employed the 2008-2018 LCZ dataset with 100 m resolution produced by Chen et al. [50], which

179   covers Hong Kong. This LCZ dataset is based on a multi-year comparable LCZ sample set. It

180   performs classification employing the Google Earth Engine (GEE) platform, considering

181   satellite remote sensing images such as Landsat 8, Landsat 5, Sentinel-1 SAR GRD, Sentinel-

182   2 MSI, VIIRS and DMSP OLS, as well as GMTED2010 (Global Multi-resolution Terrain

183 Elevation Data 2010) [50]. The included remote sensing images are derived from different sensors,

184 including spectral, radar, nighttime light and terrain, providing multi-dimensional information

185 for the LCZ classification. As a result, this LCZ dataset achieves an overall accuracy of 61.61%

186 over a multi-year span, which is a comparable and acceptable accuracy to most current LCZ

187 classification products [50].

188   Moreover, landscape metrics have been proven essential for mapping fine air temperature

189 patterns [50,52]. Landscape metrics are developed from the "patch-corridor-matrix" theory in

190 landscape ecology [53]. They provide pictures of the landscape's characteristics through three

191 levels of metrics, including patch-, class- and landscape-level. As the names imply, the patch-

192 level metrics reflect the characteristics of a single patch of land covers, such as area, perimeter

193 and shape. Class-level metrics capture the spatial pattern of the characteristics of all patches of

194 a specific land cover type. While the landscape-level metrics provide the picture of how all

195 land cover types are mixed in space. Of course, LCZ can also be combined with landscape

196 metrics to reflect LCZ-based landscape characteristics. Using different search radii when being

197 calculated, landscape metrics can reflect the landscape characteristics within different radii

198 around a location.

199   In this study, we employed Fragstats software (version 4.2), a well-established and widely

200 used landscape metric software, to calculate LCZ-based landscape metrics for Hong Kong over

201 the study period. Meanwhile, referencing previous studies, we calculated landscape metrics

202 from 100 m to 1000 m search radius. That is, when traversing each grid to calculate the different

203 landscape metrics, we took each grid as a reference point and calculated the different landscape

204 metrics with different search radius from 100m to 1000m in 100m intervals. The selected

205 landscape metrics are listed in Table S1. As there are 18 LCZ types, we finally generated a vast

206 number of landscape drivers, 13550 in total. However, to reduce the computational burden of

207 subsequent machine learning modelling and to select landscape indices of general interest,

208 those landscape indices that had valid values at less than 50% of the weather stations were

209 excluded. Therefore, only 226 landscape metrics became our preliminary landscape drivers for

210 the subsequent machine learning modelling.

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

212 In this study, we employed the random forest (RF) model [54], a well-performing, efficient

213 and widely used machine learning algorithm, for hourly Ta and RH mapping in Hong Kong.

214 The RF model consists of a number of decision trees that vote to determine the outcome of

215 regression or classification [55]. With these "weak" but varied decision trees, the RF model can

216 integrate them into a powerful regressor or classifier. Because of this structure, RF models can

217 be used to solve non-linear problems and effectively avoid overfitting [56]. Moreover, the RF

218 model can evaluate the importance of each driving factor [57], which is crucial for the subsequent

219 selection of the key drivers and for improving the efficiency of our models.

220 We used the Python-based "scikit-learn" extension package (Version 0.24.2) to construct

221 annual RF models for estimating Ta and RH, respectively. In the case of the RF model for

222 estimating Ta, the previously prepared meteorological and landscape drivers were fed into the

223 RF model, as well as the current time (hour), longitude, latitude and elevation. The

224 meteorological drivers include RH, PRE, PRS and VV2 for the previous 24 hours. Thus, a total

225 of 326 preliminary driving factors were fed into the RF model. However, the large number of

226 driving factors increases the training and prediction time of the model and takes up more

227 computational resources. Moreover, most of these driving factors can not improve the model

228 accuracy. Therefore, we trained the RF model using the 2018 data and selected key driving

229 factors based on their importance. We employed the permutation importance provided by

230 scikit-learn as a metric for importance assessment, which is suitable in cases where a feature

231 has many unique values. The permutation importance of a feature is defined as the deviation

232 of the metric value from the baseline metric value after the permutation of this feature. We

233 executed the training of the RF model 10 times, taking the mean value of permutation

234 importance as the importance of driving factors. These driving factors were ranked in

235 descending order of permutation importance, and we selected the top driving factors that

236 contributed more than 98% in total importance as key driving factors. These key driving factors

237 were then applied to the modelling of the other years. At the same time, we found the optimal

238 number of decision trees by adjusting the n_tree parameter. The same process was followed in

239 the RF modelling for estimating RH, but the only difference was that Ta replaced RH as a

240 meteorological driver.

241 In constructing an RF model, 70% of the samples were used as training samples, while

242 the remaining 30% were used as test samples. Due to the characteristic of the RF model,

243 approximately 30% of the training samples are not actually involved in the model training in

244 each modelling process, which are called out-of-bag samples. The goodness-of-fit, R^2,

245 calculated using them is called the oob_score, which can be used to measure the model

246 accuracy. Besides, we also calculated three accuracy metrics for the models using the test

247 samples. They are R^2, the root mean square error (RMSE) and the mean absolute error (MAE).

248 These four accuracy metrics provide a comprehensive picture of the accuracy of the RF models.

2.5. The thermal index - Net Effective Temperature (NET)

250 Previous studies have developed different thermal indices to reflect the combined or

251 certain aspects of the thermal environment [6,43,58]. To reflect the thermal environment under the

252 combined influence of Ta and RH, we employed a thermal index, Net Effective Temperature

253 (NET)[58]. In addition, NET is adopted by the Hong Kong Observatory (HKO), so the NET

254 evaluation result can directly serve the HKO [59]. NET was first introduced in 1937 with the

255 name Effective Temperature[58]. It was initially proposed to measure the effect of RH in hot

256 weather. Later it was improved to consider the effect of wind on human thermal comfort[58].

257 That is, the stressful feeling can be exacerbated by calm wind and humidity in hot weather and

258 strong wind and humidity in cold weather[59]. Therefore, this improvement makes it applicable

259 to human thermal comfort in both hot and cold weather[60]. NET is determined by Ta, RH and

260 wind speed and can be described by the following equation[58]:

261
$$NET = 37 - \frac{37 - Ta}{0.68 - 0.0014 \times RHU + \frac{1}{1.76 + 1.4v^{0.75}}} - 0.29 \times Ta \times (1 - 0.01 \times RHU) \quad \text{(Eq. 1)}$$

262 where the unit of NET is °C, Ta represents the air temperature (in °C), RH is the relative

263 humidity (in %), and v means the wind speed (in m/s). The setting of the constants in Eq. 1

264 was applied in a previous study of Hong Kong[59]. Since this study only generates Ta and RH

265 mapping, v will be set to 0 when calculating NET. Hence, the NET calculated in this study

266 reflects the most stressful thermal environment felt in the summer in a calm wind condition.

267 **3. Results**

268 **3.1. Validation of the hourly Ta maps**

269 By inputting 2018 data, we used the 326 preliminary driving factors to find the optimal

270 n_tree from 20 to 500 for the RF model for estimating Ta, as shown in Fig. 3(a). The results

271 show no significant improvement in the R^2 value of the model after the n_tree exceeds 100.

272 Therefore, we used the top 90 drivers that contributed the most in the importance assessment

273 when n_tree was 100 as the key drivers. The importance assessment showed that the

274 importance score of these 90 key drivers (0.717) contributed 98.3% of the total importance

275 score of the preliminary driving factors (0.729). Further, we searched from 30 to 500 for the

276 optimal n_tree when modelling the Ta estimation using these 90 key drivers (Fig. 3(b)). The

277 results show that the R^2 value of the model also reached the highest when the n_tree was 100.

278 Therefore, we set n_tree to 100 when using the 90 key drivers for RF modelling for Ta

279 estimation for each year.

(a)

(b)
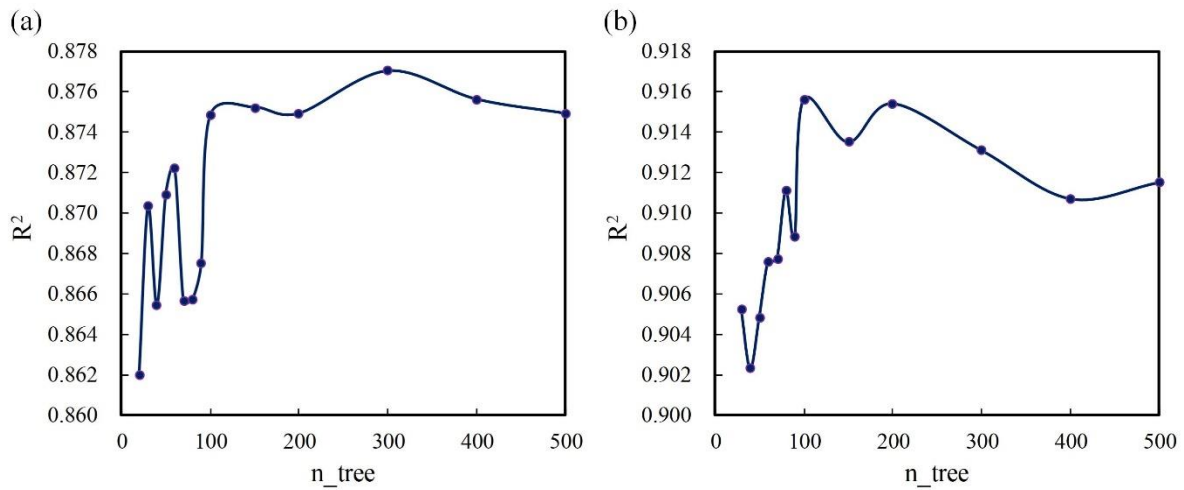
280

281 Fig. 3 The relationship between n_tree and R^2 in RF modelling for the Ta estimation by (a) modelling

282 with all preliminary drivers and (b) modelling with the 90 key drivers.

283

284 Among the 90 key drivers for estimating Ta, 78 were meteorological drivers, while 10

285 were landscape drivers (see Table S2). Meanwhile, current time and elevation were also

286 selected as key drivers. When estimating the Ta using these 90 drivers, the five most important

287 drivers and their contributions were 38.0% for RH_1Hours_Before, 10.9% for Current_time

288 (hour), 5.0% for mw03_ai (Aggregation Index at a radius of 300 m), 4.4% for elevation and

289 3.8% for RH_22Hours_Before. In total, they contributed 62.1% of the importance. It is evident

290 from this that meteorology, landscape, time and topography all play a non-negligible role in

291 estimating the spatial distribution of Ta.

292 We then conducted RF modelling to estimate Ta for each year using the determined key

293 drivers and n_tree. The accuracy of the RF models and the number of weather stations for each

294 year are shown in Table 2. The R^2 and oob_score, calculated using the test samples and out-of-

295 bag samples, respectively, show similar accuracies, with mean values of 0.8723 and 0.8652.

296 The values of the RMSE for each year range from 0.9648°C to 1.3028°C, with a mean value

14

297 of 1.1160°C. At the same time, MAE values ranged from 0.7147°C to 0.9838°C, with a mean

298 value of 0.8227°C. Overall, the accuracy metrics show the RF models for estimating Ta

299 achieved good accuracy.

300

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	—

301

302 Further, we validated the accuracy of the RF models for estimating Ta at different hours.

303 Fig. 4 shows the R^2 values of the RF models for estimating Ta for different years at different

304 hours. The RF models of different years showed similar performance. In general, the RF

305 models achieved satisfactory R^2 at different hours. Still, relatively, the models maintained a

306 more stable and higher R^2 during nighttime (i.e. 19:00 to 5:00). A decrease in R^2 occurred

307 during the sunrise period (i.e. 6:00 to 9:00), and a steady recovery followed. Scatter plots of

308 observed versus predicted Ta values for each hour based on the test samples for each year are

309 shown in Figures S1-S11. They all exhibit excellent linear relationships distributed around the
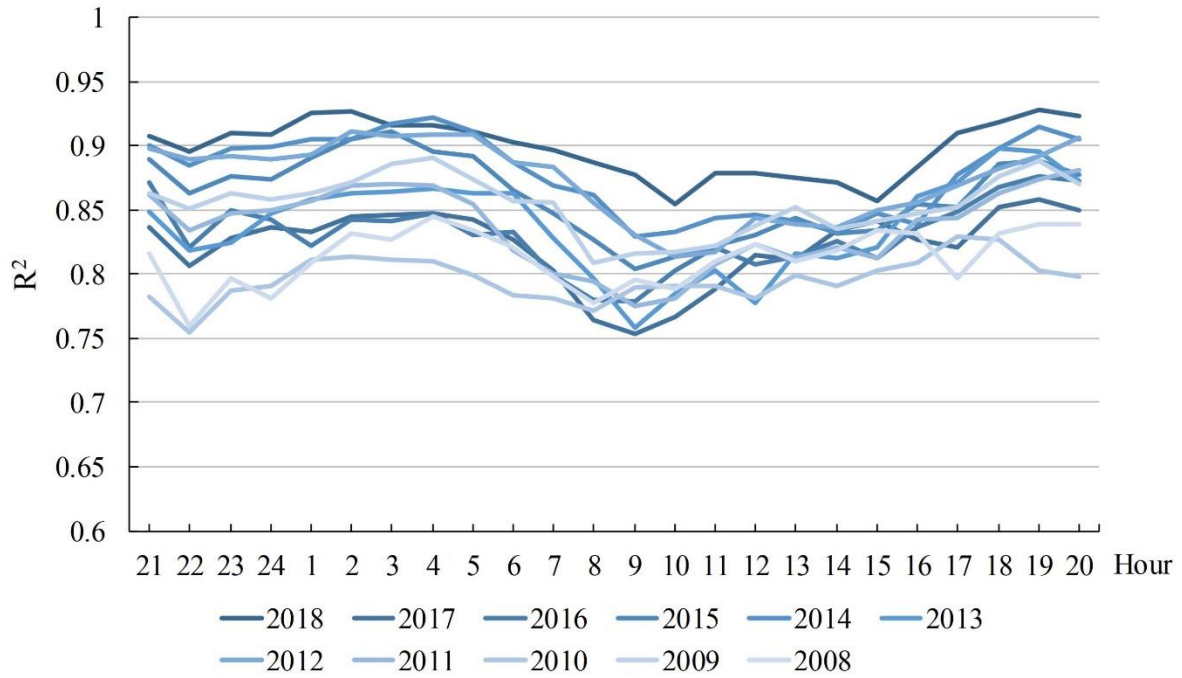
310 1:1 diagonal.

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

Fig. 5 shows the RMSE of the RF models for estimating Ta for different years at different hours. Similar to R^2, 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 Figure S12.
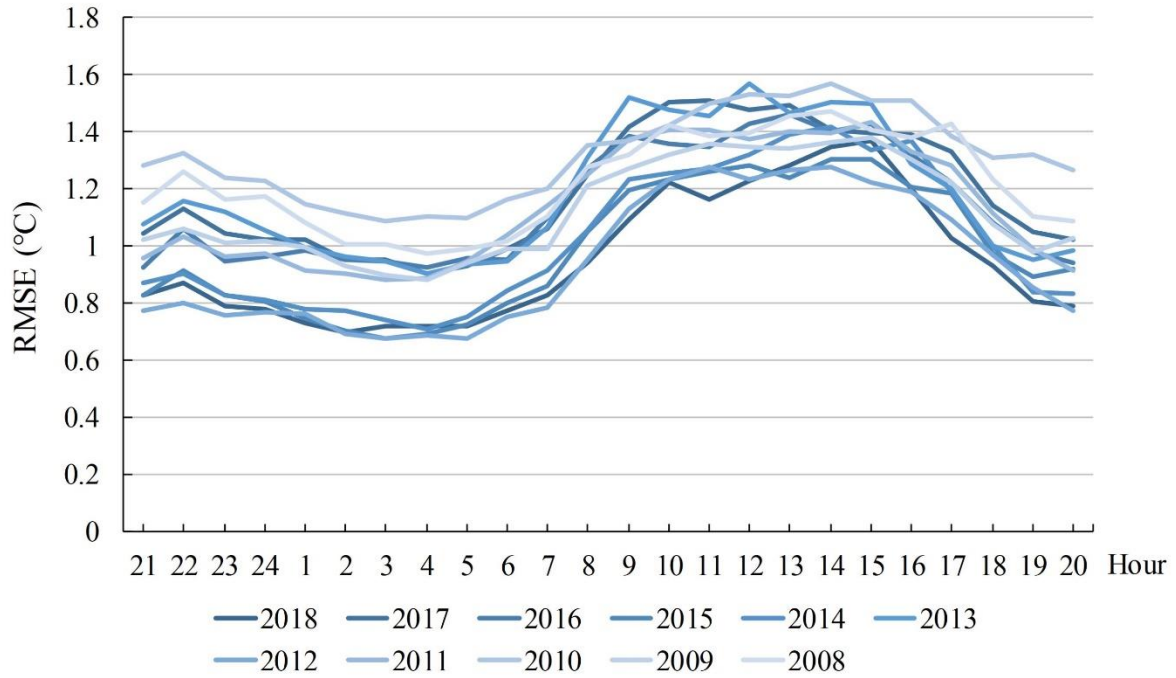
Fig. 5 RMSE of the RF models for estimating Ta for different years at different hours.

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.

17

(a)

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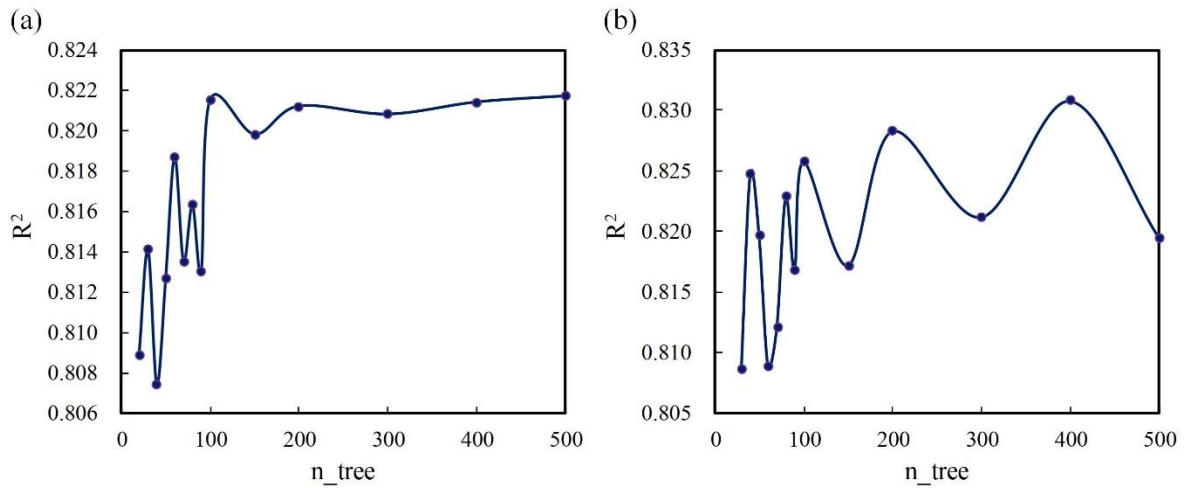
Fig. 6 The relationship between n_tree and R^2 in RF modelling for the RH estimation by (a) modelling with all preliminary drivers and (b) modelling with the 90 key drivers.

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.

18

356

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

357

358   Furthermore, we validated the performance of the RF model for estimating RH at different

359   hours. Fig. 7 shows the R^2 of the RF models for estimating RH for different years at different

360   hours. It can be seen that the R^2 curves for the RH modelling show a consistent trend across

361   the different years but with more fluctuations than Ta's. During the late night hours (i.e. 23:00

362   to 6:00), R^2 remained more stable at relatively high values, but during the sunrise period (i.e.

363   7:00 to 8:00), R^2 showed a decrease. However, the R^2 values then climbed up to reach a

364   maximum at 16:00 to 17:00, even higher than those during the nighttime. At dusk and in the

365   subsequent period (i.e. 18:00 to 22:00), the R^2 values again decreased and fluctuated. In

366   addition, scatter plots of the hourly observed versus predicted RH values for each year of the

367   test samples are shown in Figures S13-S23. They all exhibit a good linear relationship
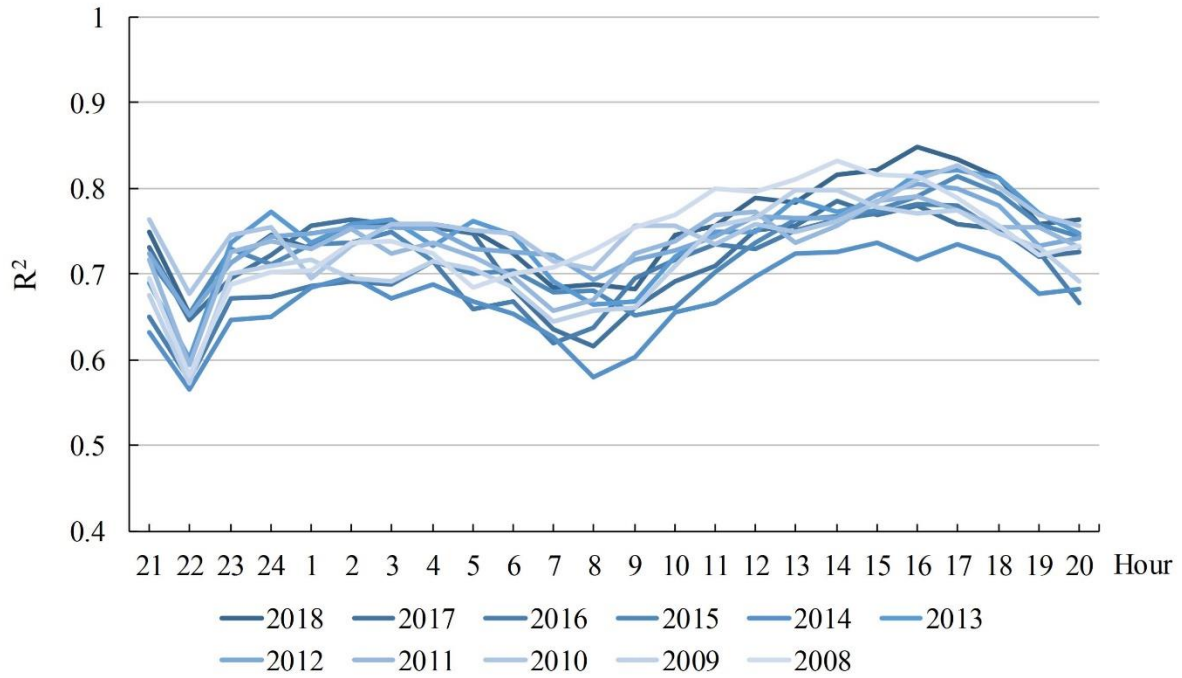
368   distributed around the 1:1 diagonal.

Fig. 7 R^2 of the RF models for estimating RH for different years at different hours.

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 in Figure S24. Here we take the performance of RMSE as an example. Similar to R^2, 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.
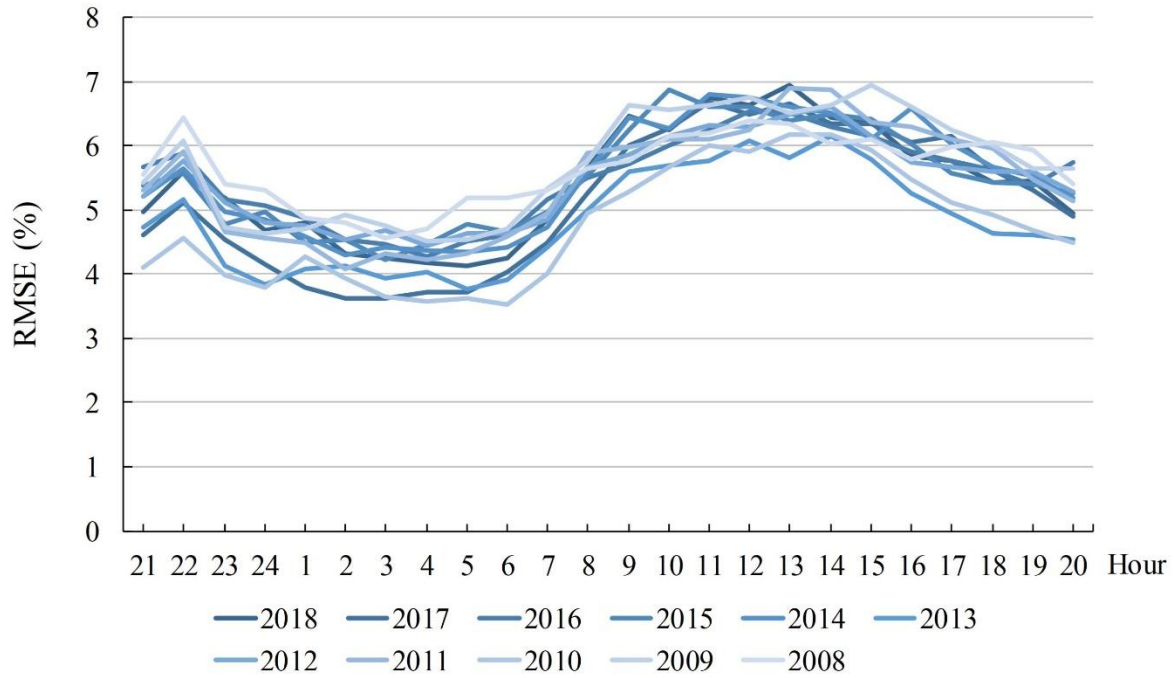
Fig. 8 RMSE of the RF models for estimating RH for different years at different hours.

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 hours of the day, we chose 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

397 distribution of weather stations, i.e. there is more Ta spatial detail where the weather stations

398 are denser. Nevertheless, comparing the LCZ maps, we can see that the Ta spatial pattern

399 estimated by the RF model shows a correlation out of topography. That is, Ta is lower in hilly

400 areas covered by dense trees (LCZ A) and vegetation (LCZ D) than in lowlands.

401 Fig. 9 also zooms in on the Kowloon Peninsula, one of the core urban areas of Hong Kong.

402 Comparing the LCZ maps reveals that Ta is higher in the core urban area, with a mixture of

403 different types of buildings. Furthermore, comparing the Ta spatial pattern at 21:00 and 4:00,

404 it can be seen that the core urban area cools more slowly at night than the fringes of the urban

405 area.
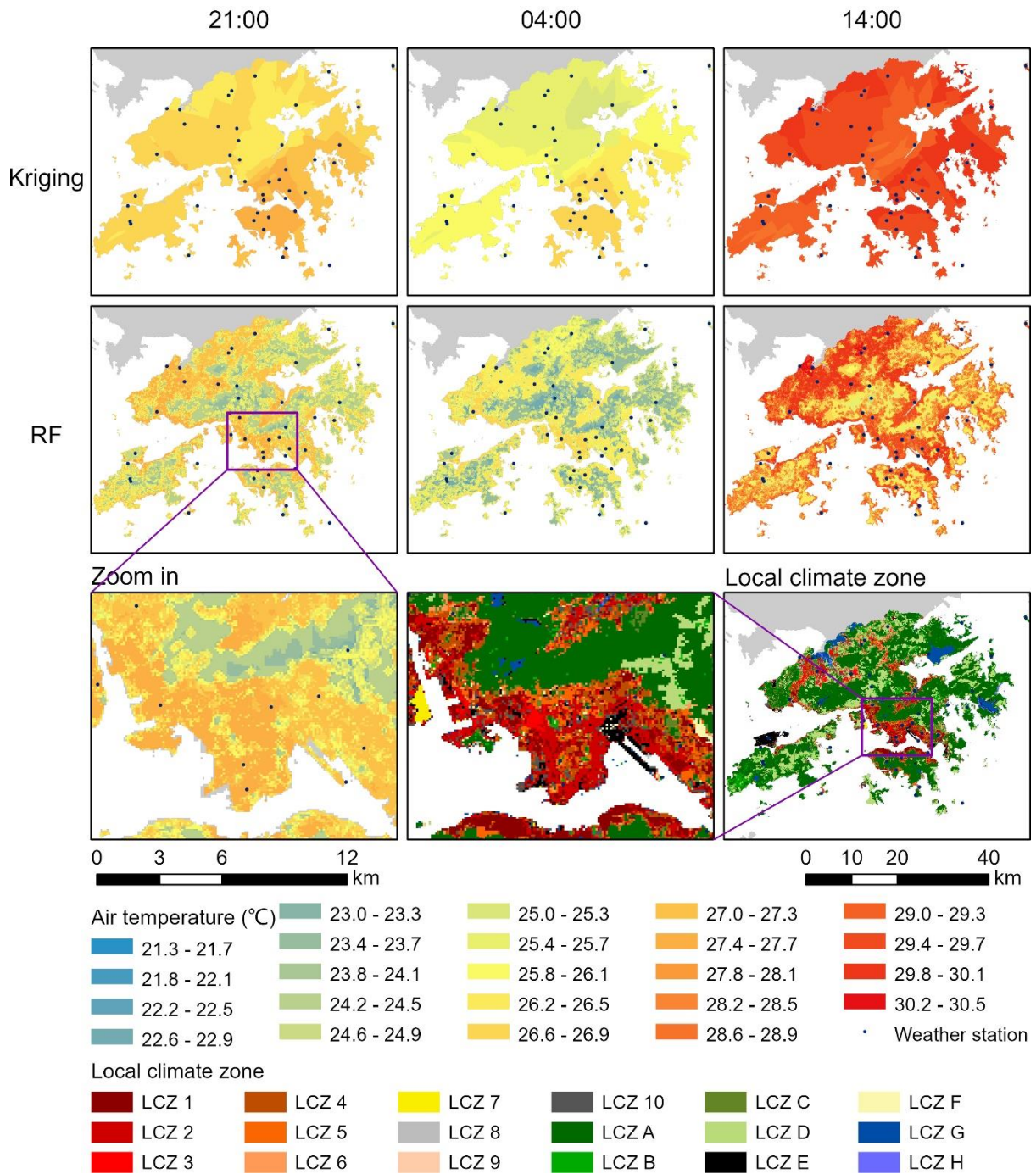
Fig. 9 The difference between the RF model and the Kriging interpolation in estimating the spatial pattern of Ta in the 2018 summer season.

3.4. 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 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
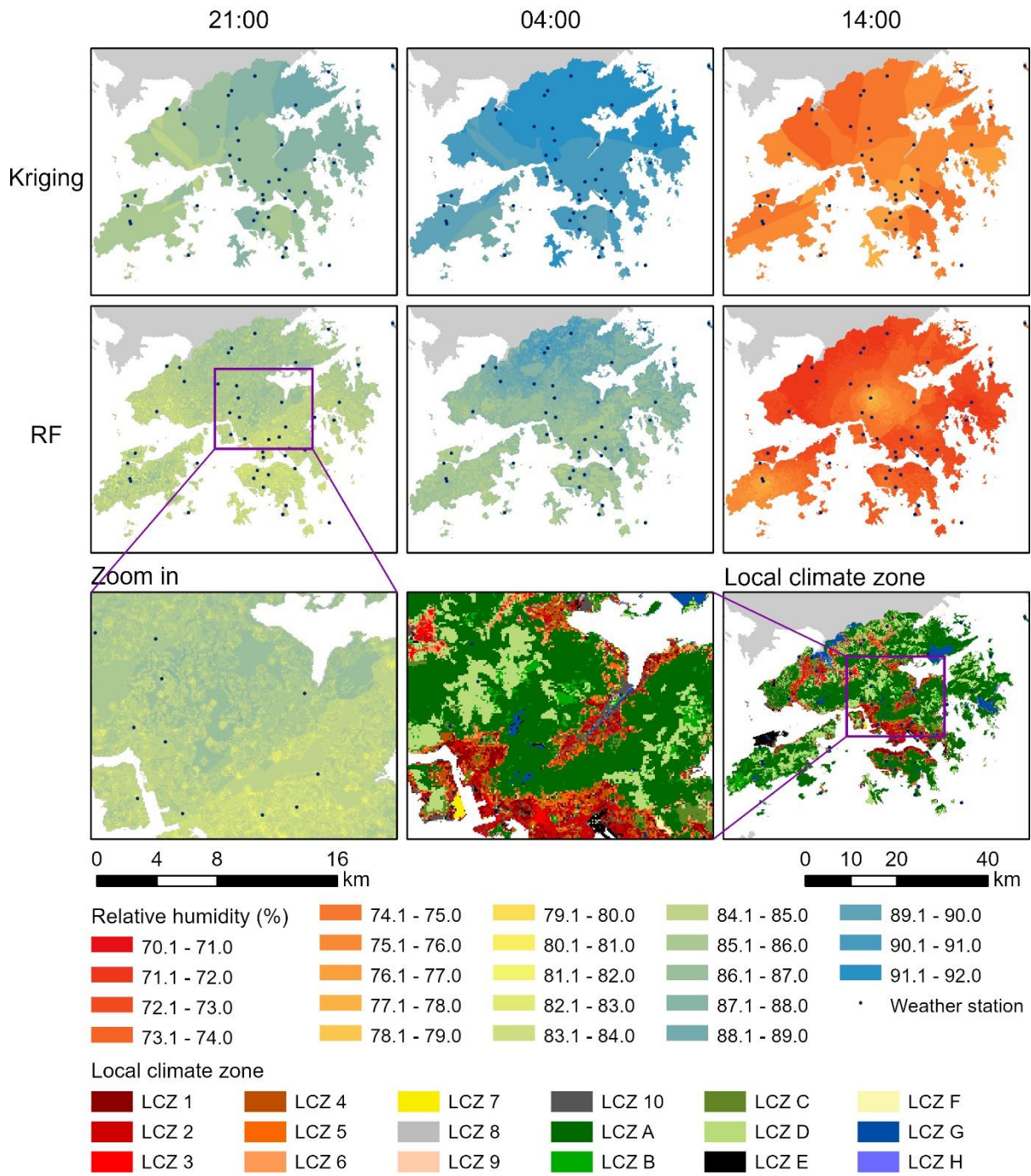
trees (LCZ A) and low plants (LCZ D).

Fig. 10 The difference between the RF model and the Kriging interpolation in estimating the spatial pattern of RH 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 [59], so the

437 NET-based thermal comfort map could be a useful reference to HKO. NET reflects the human

438 thermal comfort considering the combined effect of Ta, RH and wind speed. As wind speed

439 maps are not provided in this study, the NET map obtained here can be considered the most

440 stressful level of human thermal comfort felt in a calm wind condition during the summer

441 season.

442   Fig. 11 shows the difference between NET and Ta in depicting the spatial pattern of the

443 thermal environment in the 2018 summer season. As it considers the role of RH, NET shows a

444 subtle difference from Ta at late night (4:00) and midday (14:00). In the north-western part of

445 Hong Kong, for example, NET is higher than Ta at 4:00 and lower than Ta at 14:00 because it

446 is more humid at late night and drier during the day there compared to other areas. It means

447 that people there will feel hotter than the actual air temperature at night, while the thermal

448 comfort will be relatively relieved during the daytime.

449   Fig. 11 also zooms in on the Shatin District, which has a long, narrow built-up area

450 surrounded by hills on three sides. Here, we can still find the pattern that the NET of the built-

451 up area is higher than Ta at night (Fig. 11 (a) and (b)), while the opposite is observed during

452 the daytime (Fig. 11 (c) and (d)). Moreover, the difference between the NET of the hilly area

453 and the built-up area is more significant than that of Ta at night (Fig. 11 (a) and (b)), while it

454 is smaller than Ta during daytime (Fig. 11 (c) and (d)). In other words, at night, people feel the

455 built-up area hotter than the hilly area more strongly than the difference in actual air

456 temperature. In contrast, during the daytime, people perceive the difference in temperature

457 between the built-up area and the hilly area to be smaller than it actually is.
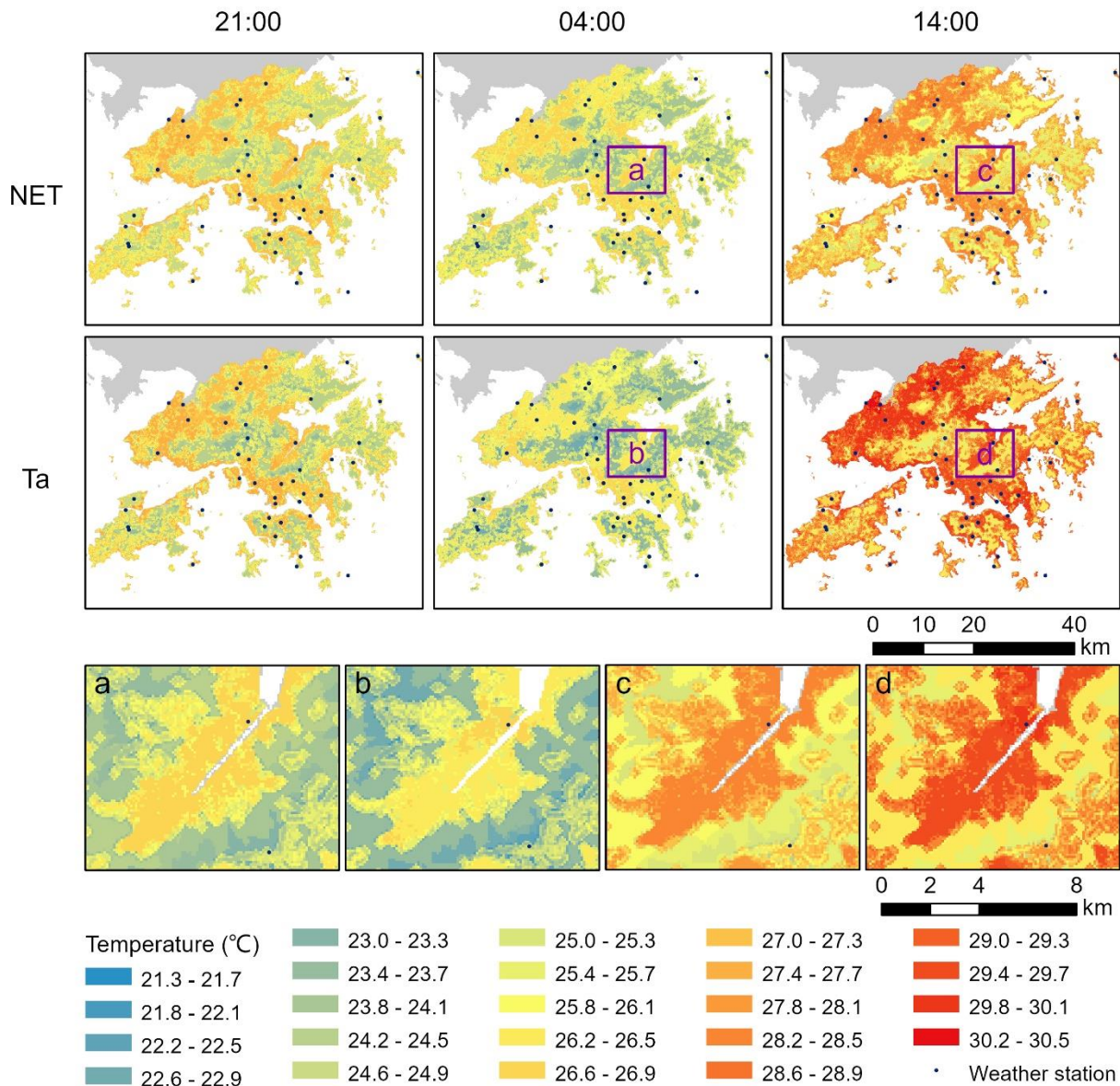
26

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

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 hours were all included, while PRE was selected for only 6 hours. Regarding importance, RH, PRS, VV2 and PRE contributed

27

467   52.4%, 12.3%, 4.0% and 1.6%, respectively. Besides, the landscape drivers, current time (hour),

468   and elevation contributed 14.4%, 10.9% and 4.4%, respectively.

469   The meteorological drivers also played the most critical role in the RF model for

470   estimating RH (Table S3). Ta, PRS and VV2 for the previous 24 hours were all included, while

471   PRE was also included for 11 hours. In the importance assessment, Ta, PRS, VV2 and PRE

472   contributed 70.8%, 9.7%, 6.4% and 2.4%, respectively. Simultaneously, the landscape drivers

473   and current time contributed 7.7% and 3.0%, respectively.

474   We can find that Ta and RH are the most critical drivers of each other, which reflects their

475   clear correlation. However, Ta has a greater influence on RH than RH has on Ta. In addition,

476   the landscape has different influence degrees on Ta and RH. Ta is more influenced by the

477   landscape, including a non-negligible influence of Ta by altitude. This is reflected in the higher

478   correlation of the hourly Ta distribution with the LCZ-based landscape and topography.

479   **4.2. The performance difference between daytime and nighttime of the models**

480   Overall, the RF models for estimating Ta and RH both perform better during nighttime,

481   especially at late night (23:00 to 5:00). Then, they both show a decrease in accuracy during the

482   sunrise period (6:00 to 9:00). Moreover, their estimation errors remain at a high level for most

483   of the subsequent daytime period (10:00 to 15:00). However, the R^2 of the model for estimating

484   RH gradually increases during this period as RH differences between regions become larger

485   during daytime. The models' errors decrease progressively until after the sun gradually sets

486   (after 16:00). Similar better model performance during nighttime has been seen in previous

487   studies of hourly Ta mapping [61]. Thus, we infer that solar radiation is one of the main factors

488   affecting Ta and RH mapping accuracy, as it is the most important source of energy driving

489   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 [36,61], while on a daily scale, these two metrics were reported to be 2.0 °C and 1.5 °C [35]. 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% [62]; an ML-based daily RH mapping could achieve an R^2 of 0.71 [41], while a summer RH mapping reported RMSE and MAE of 7.4% and 2.4%, respectively [42]. In contrast, our hourly RH mapping has

514 R^2, RMSE and MAE of 0.7970, 5.3816% and 3.8641%, respectively. Overall, both our Ta and

515 RH mapping achieved comparable and satisfactory accuracy.

4.5. Potential application

517 The multi-year summer season hourly thermal comfort dataset proposed in this study

518 contains Ta, RH and NET maps, which allows it to be applied to many relevant areas. For

519 example, it can be used to explore the long-term relationship between urban morphology and

520 the thermal environment and to capture the trend of urban heating [19]. Moreover, this dataset

521 has fine spatiotemporal resolution and focuses on high-density urban areas. Thus, it can be

522 combined with big data on human activity trajectories to explore real-time heat exposure [63,64]

523 and heat-related health risk assessment [14,15] at a fine spatiotemporal scale. In addition, this

524 dataset can be used for thermal comfort assessment, which can further guide urban planning,

525 building design, and the assessment of cooling energy consumption [12,13].

4.6. Limitations and future work

527 Despite all our efforts, there are still some limitations in this study. First, the spatial pattern

528 of meteorological drivers was obtained by Kriging interpolation. In the future, finer

529 meteorological spatial data could reduce the estimation error due to interpolation. Second, the

530 LCZ maps that calculated the landscape drivers introduced their errors. The increasing

531 accuracy of LCZ classification in the future will reduce this part of the error. Third, the previous

532 inferred that solar radiation might be an essential driver of hourly Ta and RH mapping, but this

533 study did not consider the drivers directly related to solar radiation. Because Hong Kong is

534 located in a monsoonal climate zone with clouds during most of the summer season,

535 multispectral remote sensing images that observe the land surface at a near-hourly scale are

536 difficult to obtain. In addition, the machine learning model for thermal comfort mapping could

537 be applied to the whole year and other regions in the future to explore the robustness and

538 generalisation of the model.

5. Conclusions

540　In this study, we present an hourly thermal comfort dataset for Hong Kong, including Ta

541 and RH using a machine learning algorithm (RF regression model), which covers the summer

542 season (May to September) from 2008-2018. Since considering the meteorological driver and

543 the LCZ-based landscape driver, the thermal comfort dataset provides fine spatial details at 100

544 m resolution, especially the spatial pattern of hourly Ta presents a closer relationship with the

545 landscape. Moreover, the validation results show excellent accuracy of the thermal comfort

546 dataset in both hourly Ta and RH estimations. The mean values of R^2, RMSE and MAE for Ta

547 estimation achieved 0.8723, 1.1160°C and 0.8227°C, respectively, while those for RH

548 estimation reached 0.7970, 5.3816% and 3.8641%. Further, the NET calculated by combining

549 Ta and RH gives a more relevant picture of human thermal comfort. For example, it can reveal

550 that people in built-up areas feel hotter than the actual temperature at night, thus presenting a

551 more serious challenge to human thermal comfort. In the future, the hourly thermal comfort

552 dataset can be combined in depth with studies on big data, health, energy consumption, urban

553 planning and more to explore the socio-economic and human impact of the thermal

554 environment.

562 of Hong Kong, The Chinese University of Hong Kong, and the Hong Kong Observatory.

565

Reference

567
568
1. Wu, X. et al. Quantitatively evaluating the effect of urbanization on heat waves in China. *Science of the Total Environment*. **731**, 138857 (2020).
2. Stott, P. How climate change affects extreme weather events. *Science*. **352**, 1517-1518 (2016).
3. Ma, F. & Yuan, X. More persistent summer compound hot extremes caused by global urbanization. *Geophysical Research Letters*. **48**, e2021G-e93721G (2021).
4. Zhang, J. & Wu, L. Land-atmosphere coupling amplifies hot extremes over China. *Chinese Science Bulletin*. **56**, 3328-3332 (2011).
5. 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).
6. Matzarakis, A. & Amelung, B. Physiological equivalent temperature as indicator for impacts of climate change on thermal comfort of humans. *Seasonal forecasts, climatic change and human health*: Springer; 2008. pp. 161-172.
7. de Souza, D. O. & Dos Santos Alval A, R. C. E. L. Observational evidence of the urban heat island of Manaus City, Brazil. *Meteorological Applications*. **21**, 186-193 (2014).
8. Khan, A. & Chatterjee, S. Numerical simulation of urban heat island intensity under urban--suburban surface and reference site in Kolkata, India. *Modeling Earth Systems and Environment*. **2**, 1-11 (2016).
9. Liu, W., Ji, C., Zhong, J., Jiang, X. & Zheng, Z. Temporal characteristics of the Beijing urban heat island. *Theoretical and Applied Climatology*. **87**, 213-221 (2007).
10. Dzwonkowski, B. et al. Compounding impact of severe weather events fuels marine heatwave in the coastal ocean. *Nature Communications*. **11**, 1-10 (2020).
11. Li, D., Yuan, J. & Kopp, R. E. Escalating global exposure to compound heat-humidity extremes with warming. *Environmental Research Letters*. **15**, 64003 (2020).
12. Braun, M. R., Altan, H. & Beck, S. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy*. **130**, 305-313 (2014).
13. Ihara, T., Genchi, Y., Sato, T., Yamaguchi, K. & Endo, Y. City-block-scale sensitivity of electricity consumption to air temperature and air humidity in business districts of Tokyo, Japan. *Energy*. **33**, 1634-1645 (2008).
14. Bhaskaran, K. et al. Heat and risk of myocardial infarction: hourly level case-crossover analysis of MINAP database. *Bmj*. **345**, (2012).
15. Xu, D. et al. Acute effects of temperature exposure on blood pressure: An hourly level panel study. *Environment International*. **124**, 493-500 (2019).
16. 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).
17. 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).
18. Lin, X. et al. Empirical estimation of near-surface air temperature in China from MODIS LST data by considering physiographic features. *Remote Sensing*. **8**, 629 (2016).
19. Zhao, J. et al. Exploring difference in land surface temperature between the city centres and urban expansion areas of China's major cities. *International Journal of Remote Sensing*. **41**, 8965-8985 (2020).
20. Emmanuel, R. *Urban Heat Island Mitigation Technologies*. (MDPI, 2021).
21. 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).

22. 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).

23. Tobler, W. On the first law of geography: A reply. *Annals of the Association of American Geographers*. **94**, 304-310 (2004).

24. Oke, T. R., Mills, G., Christen, A. & Voogt, J. A. *Urban climates*. (Cambridge University Press, 2017).

25. 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.

26. Yan, D. et al. Integrating remote sensing data with WRF model for improved 2-m temperature and humidity simulations in China. *Dynamics of Atmospheres and Oceans*. **89**, 101127 (2020).

27. Lamarque, J. et al. The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geoscientific Model Development*. **6**, 179-206 (2013).

28. Acero, J. A. & Arrizabalaga, J. Evaluating the performance of ENVI-met model in diurnal cycles for different meteorological conditions. *Theoretical and Applied Climatology*. **131**, 455-469 (2018).

29. 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).

30. 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).

31. Chevalier, R. F., Hoogenboom, G., McClendon, R. W. & Paz, J. A. Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. *Neural Computing and Applications*. **20**, 151-159 (2011).

32. Chronopoulos, K. I., Tsiros, I. X., Dimopoulos, I. F. & Alvertos, N. An application of artificial neural network models to estimate air temperature data in areas with sparse network of meteorological stations. *Journal of Environmental Science and Health Part A*. **43**, 1752-1757 (2008).

33. Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F. & Shiri, J. Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. *Ish Journal of Hydraulic Engineering*. **26**, 376-386 (2020).

34. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232 (2001).

35. 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).

36. 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).

37. Hodam, S., Sarkar, S., Marak, A. G., Bandyopadhyay, A. & Bhadra, A. Spatial interpolation of reference evapotranspiration in India: comparison of IDW and Kriging Methods. *Journal of the Institution of Engineers (India): Series a*. **98**, 511-524 (2017).

38. Li, T. et al. Mapping near-surface air temperature, pressure, relative humidity and wind speed over Mainland China with high spatiotemporal resolution. *Advances in Atmospheric Sciences*. **31**, 1127-1135 (2014).

39. Maharjan, S. & Regmi, R. P. Grid Based Temperature and Relative Humidity Distribution Map of the Kathmandu Valley. *Journal of Institute of Science and Technology*. **19**, 7-13 (2015).

40. Fries, A., Rollenbeck, R. U. T., Nau Ss, T., Peters, T. & Bendix, J. O. R. Near surface air humidity in a megadiverse Andean mountain ecosystem of southern Ecuador and its regionalization. *Agricultural and Forest Meteorology*. **152**, 17-30 (2012).

41. Hanoon, M. S. et al. Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia. *Scientific Reports*. **11**, (2021).

42. Li, L. & Zha, Y. Mapping relative humidity, average and extreme temperature in hot summer over China. *Science of the Total Environment*. **615**, 875-881 (2018).

43. Di Napoli, C., Barnard, C., Prudhomme, C., Cloke, H. L. & Pappenberger, F. ERA5‐HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. *Geoscience Data Journal*. **8**, 2-10 (2021).

44. Fricke, C., Pongr A Cz, R. & Unger, J. A. N. Comparison of daily and monthly intra-urban thermal reactions based on LCZ classification using surface and air temperature data. *Geographica Pannonica*. **26**, (2022).

45. 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).

46. Ren, J. et al. Exploring thermal comfort of urban buildings based on local climate zones. *Journal of*

33

Cleaner Production. **340**, 130744 (2022).

47. 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).

48. 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).

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

50. Chen, G. et al. Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning. *Science of the Total Environment*. **841**, 156737 (2022).

51. 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).

52. 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).

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

54. Breiman, L. Random forests. *Machine Learning*. **45**, 5-32 (2001).

55. 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).

56. 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).

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

59. Yip, K. M., Leung, Y. K. & Yeung, K. H. Long-term trend in thermal index and its impact on mortality in Hong Kong. *Hong Kong Observatory Reprint*, (2007).

60. Li, P. W. & Chan, S. T. Application of a weather stress index for alerting the public to stressful weather in Hong Kong. *Meteorological Applications*. **7**, 369-375 (2000).

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. Bregaglio, S., Donatelli, M., Confalonieri, R., Acutis, M. & Orlandini, S. An integrated evaluation of thirteen modelling solutions for the generation of hourly values of air relative humidity. *Theoretical and Applied Climatology*. **102**, 429-438 (2010).

63. Epstein, Y. & Moran, D. S. Thermal comfort and the heat stress indices. *Industrial Health*. **44**, 388-398 (2006).

64. Varghese, B. M. et al. Characterising the impact of heatwaves on work-related injuries and illnesses in three Australian cities using a standard heatwave definition-Excess Heat Factor (EHF). *Journal of exposure science \& environmental epidemiology*. **29**, 821-830 (2019).

65. Demuzere, M., Arg U Eso, D., Zonato, A. & Kittner, J. W2W: A Python package that injects WUDAPT's Local Climate Zone information in WRF. *Journal of Open Source Software*. **7**, 4432 (2022).

1. Wu, X. et al. Quantitatively evaluating the effect of urbanization on heat waves in China. *Science of the Total Environment*. **731**, 138857 (2020).

2. Stott, P. How climate change affects extreme weather events. *Science*. **352**, 1517-1518 (2016).

3. Ma, F. & Yuan, X. More persistent summer compound hot extremes caused by global urbanization. *Geophysical Research Letters*. **48**, e2021G-e93721G (2021).

4. Zhang, J. & Wu, L. Land-atmosphere coupling amplifies hot extremes over China. *Chinese Science Bulletin*. **56**, 3328-3332 (2011).

5. 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).

6. Matzarakis, A. & Amelung, B. Physiological equivalent temperature as indicator for impacts of climate change on thermal comfort of humans. *Seasonal forecasts, climatic change and human health*: Springer; 2008. pp. 161-172.

7. de Souza, D. O. & Dos Santos Alval A, R. C. E. L. Observational evidence of the urban heat island of Manaus City, Brazil. *Meteorological Applications*. **21**, 186-193 (2014).

8. Khan, A. & Chatterjee, S. Numerical simulation of urban heat island intensity under urban--suburban surface and reference site in Kolkata, India. *Modeling Earth Systems and Environment*. **2**, 1-11 (2016).

9. Liu, W., Ji, C., Zhong, J., Jiang, X. & Zheng, Z. Temporal characteristics of the Beijing urban heat island. *Theoretical and Applied Climatology*. **87**, 213-221 (2007).

10. Dzwonkowski, B. et al. Compounding impact of severe weather events fuels marine heatwave in the coastal ocean. *Nature Communications*. **11**, 1-10 (2020).

11. Li, D., Yuan, J. & Kopp, R. E. Escalating global exposure to compound heat-humidity extremes with warming. *Environmental Research Letters*. **15**, 64003 (2020).

12. Braun, M. R., Altan, H. & Beck, S. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy*. **130**, 305-313 (2014).

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

14. Bhaskaran, K. et al. Heat and risk of myocardial infarction: hourly level case-crossover analysis of MINAP database. *Bmj*. **345**, (2012).

15. Xu, D. et al. Acute effects of temperature exposure on blood pressure: An hourly level panel study. *Environment International*. **124**, 493-500 (2019).

16. 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).

17. 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).

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

19. Zhao, J. et al. Exploring difference in land surface temperature between the city centres and urban expansion areas of China's major cities. *International Journal of Remote Sensing*. **41**, 8965-8985 (2020).

20. Emmanuel, R. *Urban Heat Island Mitigation Technologies*. (MDPI, 2021).

21. 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).

22. 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).

23. Tobler, W. On the first law of geography: A reply. *Annals of the Association of American Geographers*. **94**, 304-310 (2004).

24. Oke, T. R., Mills, G., Christen, A. & Voogt, J. A. *Urban climates*. (Cambridge University Press, 2017).

25. 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.

26. Yan, D. et al. Integrating remote sensing data with WRF model for improved 2-m temperature and humidity simulations in China. *Dynamics of Atmospheres and Oceans*. **89**, 101127 (2020).

27. Lamarque, J. et al. The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geoscientific Model Development*. **6**, 179-206 (2013).

28. Acero, J. A. & Arrizabalaga, J. Evaluating the performance of ENVI-met model in diurnal cycles for different meteorological conditions. *Theoretical and Applied Climatology*. **131**, 455-469 (2018).

29. 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).

30. 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).

31. Chevalier, R. F., Hoogenboom, G., McClendon, R. W. & Paz, J. A. Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. *Neural Computing and Applications*. **20**, 151-159 (2011).

32. Chronopoulos, K. I., Tsiros, I. X., Dimopoulos, I. F. & Alvertos, N. An application of artificial neural network models to estimate air temperature data in areas with sparse network of meteorological stations. *Journal of Environmental Science and Health Part A*. **43**, 1752-1757 (2008).

33. Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F. & Shiri, J. Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. *Ish Journal of Hydraulic Engineering*. **26**, 376-386 (2020).

34. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232 (2001).

35. 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).

36. 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).

37. Hodam, S., Sarkar, S., Marak, A. G., Bandyopadhyay, A. & Bhadra, A. Spatial interpolation of reference evapotranspiration in India: comparison of IDW and Kriging Methods. *Journal of the Institution of Engineers (India): Series a*. **98**, 511-524 (2017).

38. Li, T. et al. Mapping near-surface air temperature, pressure, relative humidity and wind speed over Mainland China with high spatiotemporal resolution. *Advances in Atmospheric Sciences*. **31**, 1127-1135 (2014).

39. Maharjan, S. & Regmi, R. P. Grid Based Temperature and Relative Humidity Distribution Map of the Kathmandu Valley. *Journal of Institute of Science and Technology*. **19**, 7-13 (2015).

40. Fries, A., Rollenbeck, R. U. T., Nau Ss, T., Peters, T. & Bendix, J. O. R. Near surface air humidity in a megadiverse Andean mountain ecosystem of southern Ecuador and its regionalization. *Agricultural and Forest Meteorology*. **152**, 17-30 (2012).

41. Hanoon, M. S. et al. Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia. *Scientific Reports*. **11**, (2021).

42. Li, L. & Zha, Y. Mapping relative humidity, average and extreme temperature in hot summer over China. *Science of the Total Environment*. **615**, 875-881 (2018).

43. Di Napoli, C., Barnard, C., Prudhomme, C., Cloke, H. L. & Pappenberger, F. ERA5‐HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. *Geoscience Data Journal*. **8**, 2-10 (2021).

44. Fricke, C., Pongr A Cz, R. & Unger, J. A. N. Comparison of daily and monthly intra-urban thermal reactions based on LCZ classification using surface and air temperature data. *Geographica Pannonica*. **26**, (2022).

45. 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).

46. Ren, J. et al. Exploring thermal comfort of urban buildings based on local climate zones. *Journal of Cleaner Production*. **340**, 130744 (2022).

47. 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).

48. 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).

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

50. Chen, G. et al. Integrating weather observations and local-climate-zone-based landscape patterns for regional hourly air temperature mapping using machine learning. *Science of the Total Environment*. **841**, 156737 (2022).

51. 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).

52. 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).

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

54. Breiman, L. Random forests. *Machine Learning*. **45**, 5-32 (2001).

55. 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).

56. 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).

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

59. Yip, K. M., Leung, Y. K. & Yeung, K. H. Long-term trend in thermal index and its impact on mortality in Hong Kong. *Hong Kong Observatory Reprint*, (2007).

60. Li, P. W. & Chan, S. T. Application of a weather stress index for alerting the public to stressful weather in Hong Kong. *Meteorological Applications*. **7**, 369-375 (2000).

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. Bregaglio, S., Donatelli, M., Confalonieri, R., Acutis, M. & Orlandini, S. An integrated evaluation of thirteen modelling solutions for the generation of hourly values of air relative humidity. *Theoretical and Applied Climatology*. **102**, 429-438 (2010).

63. Epstein, Y. & Moran, D. S. Thermal comfort and the heat stress indices. *Industrial Health*. **44**, 388-398 (2006).

64. Varghese, B. M. et al. Characterising the impact of heatwaves on work-related injuries and illnesses in three Australian cities using a standard heatwave definition-Excess Heat Factor (EHF). *Journal of exposure science \& environmental epidemiology*. **29**, 821-830 (2019).