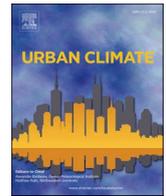




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Urban Climate

journal homepage: www.elsevier.com/locate/uclim

Utilizing world urban database and access portal tools (WUDAPT) and machine learning to facilitate spatial estimation of heatwave patterns

Yuan Shi ^a, Chao Ren ^b, Ming Luo ^c, Jason Ching ^d, Xinwei Li ^b, Muhammad Bilal ^e, Xiaoyi Fang ^{f,*}, Zhihua Ren ^g

^a Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Hong Kong, SAR, China

^b Faculty of Architecture, The University of Hong Kong, Hong Kong, SAR, China

^c School of Geography and Planning, Sun Yat-Sen University, Guangzhou, China

^d Institute for the Environment, University of North Carolina, Chapel Hill, NC, USA

^e Lab of Environmental Remote Sensing (LERS), School of Marine Science, Nanjing University of Information Science and Technology, Nanjing 210044, China

^f Chinese Academy of Meteorological Sciences, China

^g National Meteorological Information Center, China

ARTICLE INFO

Keywords:

Heatwave

Random forest

WUDAPT

Machine learning

Spatial estimation

ABSTRACT

Climate change lead to more intense, higher frequent and prolonged heat extremes. Understanding the spatial pattern of heatwave is vital for providing the corresponding weather services, making climate change adaptation strategies and heat-health actions. In this study, we present an approach to estimate the heatwave spatial patterns by utilizing the WUDAPT Level 0 data and machine learning. The analysis is based on two years (2009 and 2016) of air temperature data from 86 meteorological monitoring stations in Guangdong province of China, a subtropical region with frequent hot and sultry weather in summer. First, heatwave conditions were quantified by calculating the number of hot days and frequency of heatwave events in each year and used as the response variables. Then, random forest models were built by using a geospatial dataset consisting of WUDAPT and urban canopy parameters (UCP) as predictor variables. Based on the resultant models, spatial patterns of heatwave were estimated and mapped at 100 m spatial-resolution. The results show that this approach is able to estimate heatwave spatial patterns using open data and inform urban policy and decision-making. The study is also a new perspective and a feasible pathway of utilizing WUDAPT Level 0 product to facilitate urban environment applications.

1. Introduction

Climate change has been identified as a major challenge to environmental sustainability, human health and well-being (IPCC, 2014; WMO and WHO, 2015). In the context of climate change and the trend of global warming, heatwave events during summertime have become one of the most severe meteorological disasters in cities and societies. Heatwave generally refers to the extreme events of a

* Corresponding author.

E-mail address: fangxy@cma.gov.cn (X. Fang).

<https://doi.org/10.1016/j.uclim.2021.100797>

Received 24 August 2020; Received in revised form 10 December 2020; Accepted 3 February 2021

Available online 9 February 2021

2212-0955/© 2021 Elsevier B.V. All rights reserved.

period of consecutive hot weather (Meehl and Tebaldi, 2004; Nairn and Fawcett, 2011). In the past two decades, heatwave events during summertime become more intense, more frequently-happened, and longer-lasting (Field, 2012; Meehl and Tebaldi, 2004; Stocker, 2014). Such extremely hot weather conditions have brought serious negative impacts on environmental health (Haines et al., 2006) and also considerable economic loss (Epstein and Mills, 2005).

1.1. The associations of negative health outcomes and vulnerabilities with heatwave

There have been many studies emphasizing the associations of negative heat-health outcomes (Campbell et al., 2018; Mayrhuber et al., 2018). It is generally known that the elderly and children were more susceptible and vulnerable to heatwave events (Benmarhnia et al., 2015; Oudin Åström et al., 2011). However, extremely hot weather causes a series of heat-related health impact ranging from sleeping disorder, to heat morbidity and event to the death, not only for the elderly, children, and vulnerable peoples who do not have strong resistance (Bunyavanich et al., 2003; Kenny et al., 2010; Maughan, 2012; Xu et al., 2012) but also increases health risks for youth and working populations (Xiang et al., 2013). Long-term exposure to weather with high ambient temperature could even endanger human's life (Rey et al., 2009) and cost extra economic loss (García-Herrera et al., 2010; WMO, 2011). A national study in China indicates that a total of 4.5% (95% confidence intervals (CI): 1.4%–7.6%) excess deaths were associated with heatwaves in south China (Ma et al., 2015).

1.2. The effect of urbanization on heatwave and its spatial heterogeneity

Globally, the process of urbanization is continuing (UN, 2018; UN, 2019). The influence of urbanization on the spatiotemporal distribution of hot weather and the duration and intensity of heatwave is noticeable, especially for those countries that are experiencing rapid urbanization process (Guo et al., 2020; Oleson et al., 2015). The extremely hot weather condition is exacerbated by urbanization (Chapman et al., 2017; Luo and Lau, 2016; Sun et al., 2016). Generally speaking, urbanization increases the intensity of the urban heat island (UHI) effect, which further increases the intensity, frequency, and duration of the heatwave (Li and Bou-Zeid, 2013; Oke, 1973; Oke, 1997; Tan et al., 2010). In detail, the spatial variability in land-use patterns and inhomogeneous land surface thermal and aerodynamic properties lead to spatial heterogeneity in the near-ground wind field (Comrie, 2000), radiation and energy balance (Arnfield, 2003), and anthropogenic heat (Taha, 1997). All these spatial heterogeneities make the extremely hot weather condition varies among locations (Hart and Sailor, 2009). Such spatial heterogeneity makes city dwellers who live in urbanized areas, especially those who live in compact built-up areas in large cities are more vulnerable to heatwave (Uejio et al., 2011a; WMO and WHO, 2015).

1.3. The necessity of incorporating fine-scale spatial heterogeneity into the estimation of heatwave condition

Most of the heat-health related studies are based on time-series analysis; therefore, more focuses on the temporal characteristics (frequency, duration) of heatwave events (Kovats and Hajat, 2008). These studies usually take a city or an area of interest as a whole to associate health burdens with temperature-related variables (the spatial extent ranges from several kilometers to dozens of kilometers). Therefore, the effect of urbanization and the spatial variability of heatwave weather have become missing elements in most of the above studies and have not been comprehensively investigated (Kaiser et al., 2007; Kyselý, 2002; Le Tertre et al., 2006). Many studies have demonstrated that the vulnerability of citizens to heatwave is associated with demographic variables and socioeconomic factors (Bao et al., 2015; Chan et al., 2012; Gronlund et al., 2015; Uejio et al., 2011b). Most of these studies overlay heat exposure map, surface/air temperature map on the spatial information on vulnerable population, in which spatial scale mismatch issues between the heatwave map and spatial information on the corresponding vulnerability have been a concern (i.e., the Modifiable Areal Unit Problem) (Fraser et al., 2018; Ho et al., 2015; Wong, 2004). Commonly-used spatial units of the city administrative boundaries or urban planning zoning correspond to a spatial resolution approximately a couple of kilometers. In such cases, the assessment of urban vulnerability to heatwave events and the corresponding prevention measures would have higher spatial uncertainties. The uncertainties could be even more considerable in those regions with a complex geographic context.

The importance of locating the groups of people with a high-vulnerability to the impacts of heatwave has been emphasized (Johnson et al., 2012), which means that it is vital to take spatial elements into consideration. In that case, acquiring a detailed fine-scale spatial understanding of the heatwave is essential to heat risk prevention and public health actions (Buscaill et al., 2012). In recent years, relevant studies have been conducted for the spatial mapping of heat-related risks in many large or megacities worldwide (Dugord et al., 2014; El-Zein and Tonmoy, 2015; Klein Rosenthal et al., 2014; Lemonsu et al., 2015; Wolf and McGregor, 2013). Significant spatial variabilities of heat-related health impact were found in all the above cases. Undoubtedly, reliable fine-scale spatial information on heatwave is a fundamental part of all strategic actions in relation to the reduction of heat-related risk and vulnerability. Moreover, it is also a vital part of the information for urban planning for a more resilient built environment under heatwaves (Maragno et al., 2020).

1.4. Study objective

Understanding the spatial pattern of heatwave is important for public health management and urban development. Despite the necessity of incorporating spatial heterogeneity into the estimation of heatwave has been recognized, yet few studies focus on direct mapping of the spatial distribution of heatwave conditions at a spatial resolution that finer than the commonly-used spatial units of the

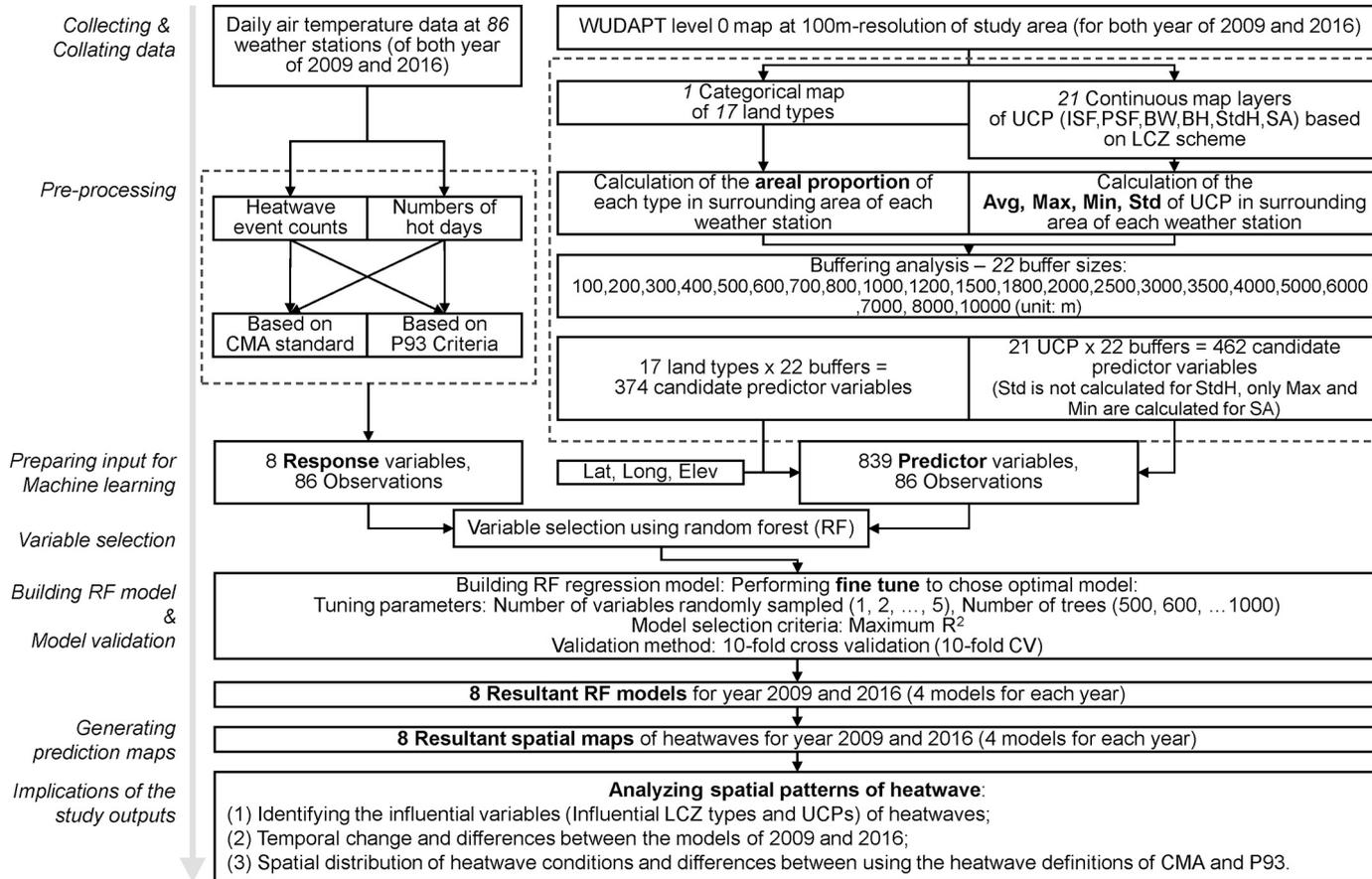


Fig. 1. The workflow of the present study.

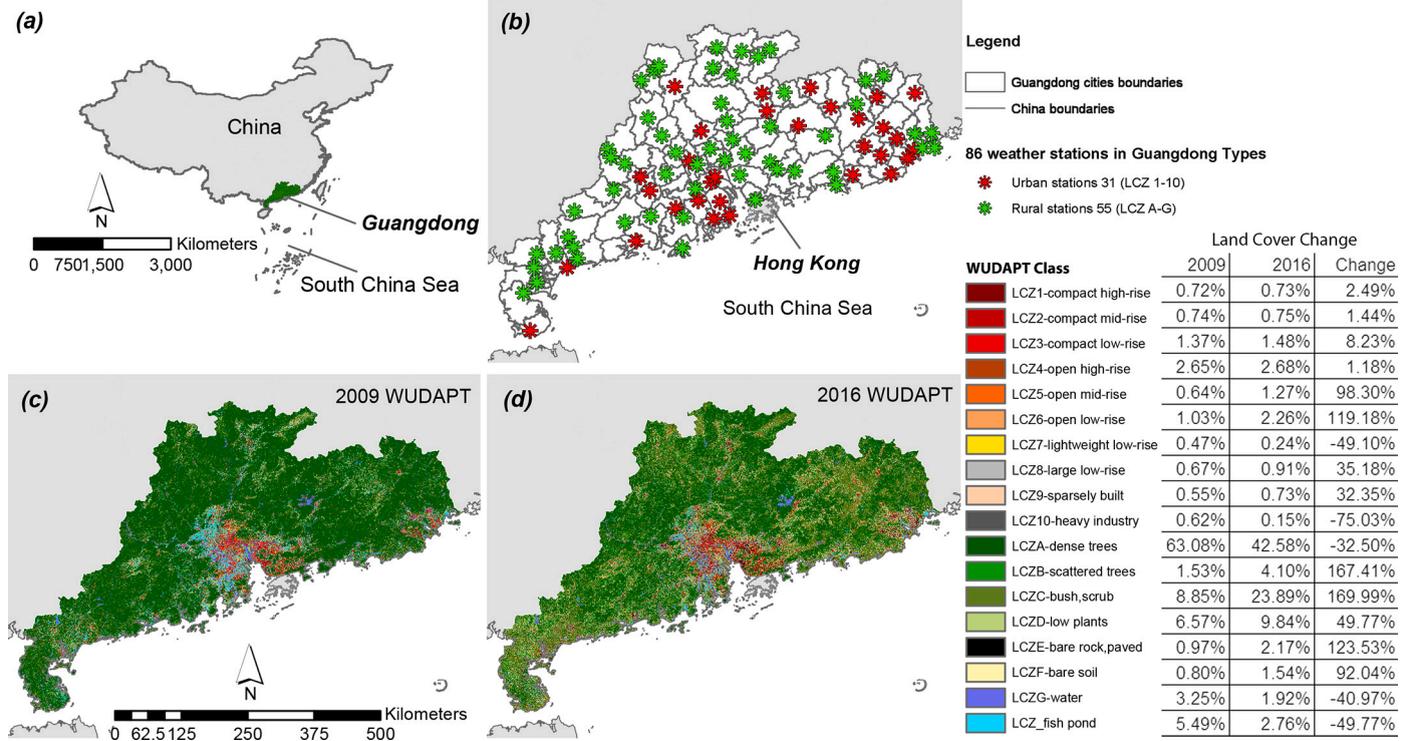


Fig. 2. The location of the study area - Guangdong province in China (a). The location of 86 national meteorological monitoring stations in Guangdong province (b). The WUDAPT map of the year 2009 (c) and 2016 (d). A table (at the right side of the legend) to illustrate the LCZ difference between 2009 and 2016 and the rate of change.

city administrative boundaries or urban planning zoning. To address the above research gap, in this study, we present an approach for the investigation of the spatial heterogeneity of the heatwave condition via machine learning and geospatial mapping techniques. The high-resolution spatial maps of heatwave generated by this study will enhance the robustness of spatial assessment and mapping of heat-health risk. Besides the resultant spatial mapping of the heatwave condition, the important influential variables of heatwave will be identified by the study, which will also provide valuable clues for how the cities should be properly planned to enhance the resilience to heatwave events and heat-related disasters.

2. Materials and methods

In this study, spatial buffering analysis and machine learning technique - random forest were adopted. Guangdong province in China, a subtropical region with highly diversified landscape and spatially heterogeneous land surface coverage, was selected as the study area, as its geographical complexity makes it an ideal testbed for the research. Land cover types were analyzed by utilizing the Level 0 data product of World Urban Database and Access Portal Tools (WUDAPT) (Bechtel et al., 2019; Bechtel et al., 2016). Land surface morphology was also quantified by introducing urban canopy parameters (UCP). First, a buffering method was adopted to analyze WUDAPT Level 0 data and UCP geospatial data and generate predictor variables. Then, the influential predictors of the heatwave were identified via random forest variable selection. Finally, the spatial pattern of heatwave was estimated by performing random forest regression modelling. Fig. 1 shows the workflow of the present study. The year 2016 and 2009 that we selected are two of the representative warmest years on the local records. Local records indicate frequent occurrence of meteorological disasters related to heatwaves in these two years. In this study, the workflow was performed twice to generate the heatwave maps for both of the years 2009 and 2016 in which the WUDAPT land type data were also generated. Thus, it is feasible to explore the temporal change of the effect of urbanization in this period of eight years. All methodological details are introduced in the following paragraphs of this section.

2.1. Response variables

2.1.1. Heatwave definition

There is no standardized definition for heatwave worldwide. Currently, China adopts the standard of heatwave developed by the China Meteorological Administration (CMA). The CMA defines a day as a high-temperature hot day when it has a maximum daily temperature $\geq 35^{\circ}\text{C}$, and a heat weather event consists of a consecutive three hot days or more than three hot days is defined as a heatwave event. However, using a single and absolute definition for heatwave investigations for China could possibly introduce bias, as China's vast territorial area spans a wide range of latitudes and contains many different climatic zones. It has been found that the peoples' group-specific mortality risks in a city in humid subtropical climate zone (Köppen Cfa) in China using the standardized heatwave definition of CMA was significantly underestimated than that using the heatwave definition of daily mean temperature ≥ 99.0 th percentile (P99) with a duration ≥ 3 days (Zhang et al., 2017). In that case, a national-scale study on the adjustment of heatwave definition has emphasized the importance of area-specific definitions of heatwave in heat-health risk assessments and developed a series of area-specific heatwave definitions for different regions in China (Lin et al., 2019). In the study of China, the regional heatwave of Northeast, North, Northwest, East, Central and Southwest China were defined separately as being two or more consecutive days with a daily mean temperature higher than or equal to the P64, P71, P85, P67, P75 and P77 of the warm season (May to October) temperature. The heatwave in South China (for example, Guangdong province) was defined as five or more consecutive days with a daily mean temperature higher than or equal to the P93 of the warm season temperature. In the present study, both the CMA heatwave definition and the P93 definition were adopted to define heatwave. As a result, two groups of models will be generated based on two different criteria.

2.1.2. Heatwave event counts and the numbers of hot days

There is a total of 86 national meteorological monitoring stations currently located in the study area – Guangdong province (Fig. 2). All the stations are operated by CMA and conform to the WMO guide (WMO, 2008). Hourly air temperature data of the year 2009 and 2016 were collected and used for the analysis. Using CMA and P93 definition, the numbers of hot days and heatwave event counts of the entire year were calculated for both of the two years, respectively. Comprehensively considering the event counts and hot day numbers leads to a holistic understanding of the heatwave characteristics about duration and frequency. As a result, a total of eight parameters were calculated to represent the heatwave conditions and used as the model response variables (2 definitions \times 2 parameters \times 2 years). Consequently, eight models and corresponding prediction maps will be generated.

2.2. Predictor variables

The physical basis behind the spatial variability in air temperature has been comprehensively understood from the viewpoint of urban climatology (Landsberg, 1981; Oke, 1982; Oke, 1987; Oke, 1988). In the present study, to maximize the reproducibility and worldwide applicability of the workflow, all data used for generating predictor variables are open data based on worldwide unifiable standards. Two major parts of data were selected and used as the input for generating predictor variable datasets, which are: (a) Land surface cover – WUDAPT, (b) Urban canopy parameters (UCP), as they have been proved to be ideal proxies of land surface form as well as good indicators of the spatial variability in near-ground ambient air temperature (Bechtel et al., 2015; Salamanca et al., 2011).

2.2.1. Land surface cover – WUDAPT level 0 map

WUDAPT Level 0 data has been popularly used for the investigation of spatial variability in air temperature (Leconte et al., 2015; Shi et al., 2018b). It is a well-established urban data portal which aims to provide a globally standardized and detailed urban morphological database of Local Climate Zone (LCZ) (Mills et al., 2015). LCZ is a standardized urban morphology scheme which provides a 17-LCZ type land surface classification for urban climate research (Stewart and Oke, 2012). Different LCZ types represent various combinations of surface structure (sky view factor, aspect ratio, surface roughness), surface cover (ground coverage ratio of buildings, vegetation, and impervious/paved surface), surface thermal properties, surface albedo, and human activity (building functions, anthropogenic heat). Moreover, different from other existing land use/land cover (LU/LC) classification products (e.g., USGS Global Land Cover Characterization (GLCC), Climate Change Initiative (CCI) Land Cover, and GlobeLand30), the WUDAPT introduces 3-level product to meet different needs in urban climate study and as well as provides a detailed LU/LC classification, especially for the built-environment. The above features make it a better proxy to depict the aerodynamic and thermal properties of the land surface. There are mainly two types of methods of generating LCZ map - GIS-based method and satellite image-based method (Gál et al., 2015; Wang et al., 2018). The GIS-based method is more city-specific as it uses local datasets. The robustness of GIS-based results depends on the quality of local urban datasets, thus varying from place to place. Oppositely, the satellite image-based method, as the most popular one of the WUDAPT Level 0 product methods, was designed to be universal to be part of a global dataset of urban form (Ching et al., 2018). Specifically, LCZ map at a high spatial-resolution of 100 m (Level 0 data) can be generated by using open-source satellite images (Bechtel et al., 2015). As the input data, the highly standardized database of WUDAPT also enables cross-comparison between urban climate related studies in different regions of the world. The above advantages make WUDAPT a superior choice of input data to facilitate urban climate and environmental modelling (Ching et al., 2014). The WUDAPT Level 0 map of Guangdong province (Fig. 2) was generated for both of the years 2009 and 2016 in previously published peer-reviewed research. The information on accuracy assessment can be found in two papers. WUDAPT Map of 2009 is described in the study by Wang et al. (2019b). WUDAPT Map of 2016 is described in the study Ren et al. (2019).

2.2.2. Urban canopy parameters (UCP)

There have been studies focus on the relationship between parameters of urban surface parameterization such as the urban canopy parameters (UCP) and the ambient air temperature or urban heat island effects (Chen et al., 2011; Garuma, 2018; Salamanca et al., 2011; Sharma et al., 2017). In the present study, besides WUDAPT, six commonly used parameters were also used for generating continuous data layers of predictor variables, which are impervious surface fraction (ISF, also known as urban fraction), pervious surface fraction (PSF), building width (BW), building height (BH), standard deviation of building heights (StdH), surface albedo (SA). All building-related parameters were estimated from 30 m resolution DSM and DEM datasets by using the method developed by Ren et al. (2020). Specifically, the building footprints are extracted by utilizing the Google Maps Static API. The height information is estimated from the Advanced Land Observing Satellite (ALOS) World 3D Digital Surface Model dataset. The 3D building morphology can be achieved by combing the above two parts of information and be used to map UCP at a spatial resolution of 100 m. Instead of exhaustively processing data for the large spatial extent (approximately 179,800 km²) of the entire study area, an inexhaustive sampling strategy was used for generating the spatial maps of the six parameters. Specifically, for each LCZ type in the study area, we randomly select 30–100 typical LCZ sample sites that locate separately from each other. The average (Avg), maximum (Max), minimum (Min), and standard deviation (Std) values of building-related parameters were calculated for all samples of each LCZ type and then assigned to each 100 × 100 m pixel based on their LCZ type. Similarly, minimum and maximum SA values were assigned to each pixel based on the representative values of the LCZ scheme (Stewart and Oke, 2012). ISF and PSF were based on the High-resolution Multi-temporal Mapping of Global Urban Land product (Liu et al., 2018). The 2010 and 2015 data layers were used in this study, as the product only provides data layer at a 5-year interval.

2.2.3. Buffering analysis

The measured air temperature depends on not only the physical environment at the location of meteorological stations but also its surroundings at a longer spatial range (KJjun et al., 2004; Konarska et al., 2016). Therefore, instead of training the model by directly using the predictor data extract at those pixels in which the meteorological stations located, we perform buffering analysis to generate predictor variable sets. Similar buffering analysis has been adopted and found to be a useful way of investigating fine-scale spatial variability of air temperature in several studies (Brandsma and Wolters, 2012; Johnson et al., 2020; Schatz and Kucharik, 2014; Shi et al., 2018a). In this present study, a total of 22 buffer radius range from 100 m to 10 km was used for buffering analysis (For example, a buffer zone with a radius of 500 m contains a total of 80 pixels on a data layer of 100 m resolution). For the categorical data of each of the 17 LCZ types, the areal proportion within the range of each buffer radius was calculated. For all UCP parameters, the average value within the range of each buffer radius was calculated based on the continuous data layers of UCP (mentioned in Section 2.2.2). All predictor variables were calculated for both the year of 2009 and 2016. The above process, along with the latitude, longitude, and elevation of meteorological stations, resulted in a total of 839 candidate predictors, which correspond to 839 spatial data layers for each year. To be consistent with the WUDAPT maps, all spatial data layers were generated using the resolution of 100 × 100 m; thus all have a data amount of 46.1 Megapixel. The prediction map will be using the same spatial resolution which is a relatively fine resolution with regards to the spatial extent of the study area and allows to spatially continuous data layer of heatwave patterns. All geospatial data processing was completed in QGIS Desktop software (v3.10.5 LTR).

2.3. Variable selection and regression modelling using random forests

It has been found that non-linear modelling techniques are necessary for the estimation of spatial variability in air temperature, as they usually result in a better prediction performance than linear modelling approaches (Brandtma and Wolters, 2012; Voelkel and Shandas, 2017). Ensemble approaches such as random forest allow building complex non-linear models while still provide reasonable interpretability by ranking variable importance. In this study, the random forest algorithm (Breiman, 2001) was used for both the variable selection and the regression modelling.

2.3.1. Predictor variable selection

As the predictor variable dataset contains a massive amount of candidate predictor variables that have to be examined, a random forest-based strategy of variable selection developed by Genuer et al. (2010) was adopted by the present study for variable selection. Simply speaking, the method ranks all the candidate predictor variables via the random forest permutation-based score of importance, and during the process a forward stepwise strategy is employed for adding predictor variables (Genuer et al., 2015). This variable selection method identifies two subsets of important predictor variables: a larger subset of variables aims to all possible interpretation but with redundancy in explaining the variability in the response variable and a smaller subset of predictors aims to a more robust prediction without redundancy. In this study, we use the later subset, as the study aims to prediction mapping. The above variable selection process was performed in R (v3.6.3) using the VSURF package (v1.1.0) (Genuer et al., 2015) and finished using data-driven default values.

2.3.2. Random forest regression modelling and model fine-tuning

Random forest is a supervised machine learning algorithm that uses the ensemble learning method and uses the Out-of-bag (OOB) error to measure the prediction error (Breiman, 1996; Breiman, 2001; Liaw and Wiener, 2002). The prediction performance of the random forest regression model is sensitive to parameter tuning (Probst et al., 2019). The number of trees (ntree) and the number of variables considered for splitting at each node (mtry) are two commonly considered tuning parameters. In this study, a fine-tuning process is employed to let the random forest algorithm to automatically choose the optimal prediction model. Specifically, an extended tune grid was set to automatically repeat the experiment with all possible combinations between mtry ranging from 1 to 5 and ntree ranging from 500 to 1000 (using an increment of 100). In order to avoid the overfitting issue and evaluate the regression model, the resampling process was done by using repeated 10-fold cross-validation (Burman, 1989). All response data were randomly divided into ten subsets, with nine subsets used as the training dataset and the other one subset used as validation datasets. This process was repeated ten times until all data have been used as validation data once. The coefficient of determination (R^2) is selected as the metric to determine the optimal model. The above random forest regression and fine-tuning process were performed in R using the caret package (v6.0–86) (Kuhn, 2008).

3. Results

3.1. Predictor variable selection results

Following the method mentioned in Section 2.3.1, for each of the eight response variables, a subset of important predictor variables was identified. The identified variables for each response are summarized and shown in Table 1.

3.2. Resultant models and evaluation

Following the method mentioned in Section 2.3.2, the R^2 , mean absolute error (MAE) and root mean square error (RMSE) were calculated for the model evaluation. Fig. 3 shows the results and the comparison between the predicted and monitored values of the number of hot days and heatwave event counts based on CMA and P93 heatwave definitions, in the year of 2009 and 2016,

Table 1

Summary of important predictor variables identified by the variable selection process and the formula of eight prediction models. For the nomenclature of predictor variables: LCZ8_1800 means the areal fraction of LCZ8 within the buffer radius of 1800 m; SA_Max_2000 means the averaged value of the maximum surface albedo in all pixels within the buffer radius of 2000 m, so on and so forth.

Year	Heatwave definition	Model formula
2009	CMA	Number of hot days ~ Lat + Elev + LCZE_8000 + LCZE_7000 + Lon + LCZG_10000 + LCZD_6000 + LCZ8_1800 + LCZ8_2500 HW events count ~ Elev + Lat + LCZE_8000 + LCZE_7000 + Lon + LCZ8_1800 + LCZA_8000
	P93	Number of hot days ~ LCZ8_8000 + LCZ8_7000 + LCZ6_10000 + LCZ8_5000 + LCZA_8000 HW events count ~ LCZ8_10000 + LCZ8_8000 + LCZ8_7000 + LCZG_200 + LCZG_10000 + LCZ6_10000 + SA_Min_10000
2016	CMA	Number of hot days ~ Lat + LCZE_8000 + LCZ2_600 + Elev + LCZ3_300 + LCZF_8000 + ISF_Std_200 HW events count ~ Lat + Lon + LCZE_8000
	P93	Number of hot days ~ LCZ1_5000 + LCZC_1500 + SA_Min_2000 + LCZ2_2000 + SA_Max_1500 HW events count ~ LCZ1_5000 + LCZC_1500 + LCZ1_400 + SA_Max_2000 + LCZ1_6000 + SA_Min_2000 + LCZG_6000 + LCZG_2500 + LCZG_3000 + LCZ2_2000 + LCZ5_200

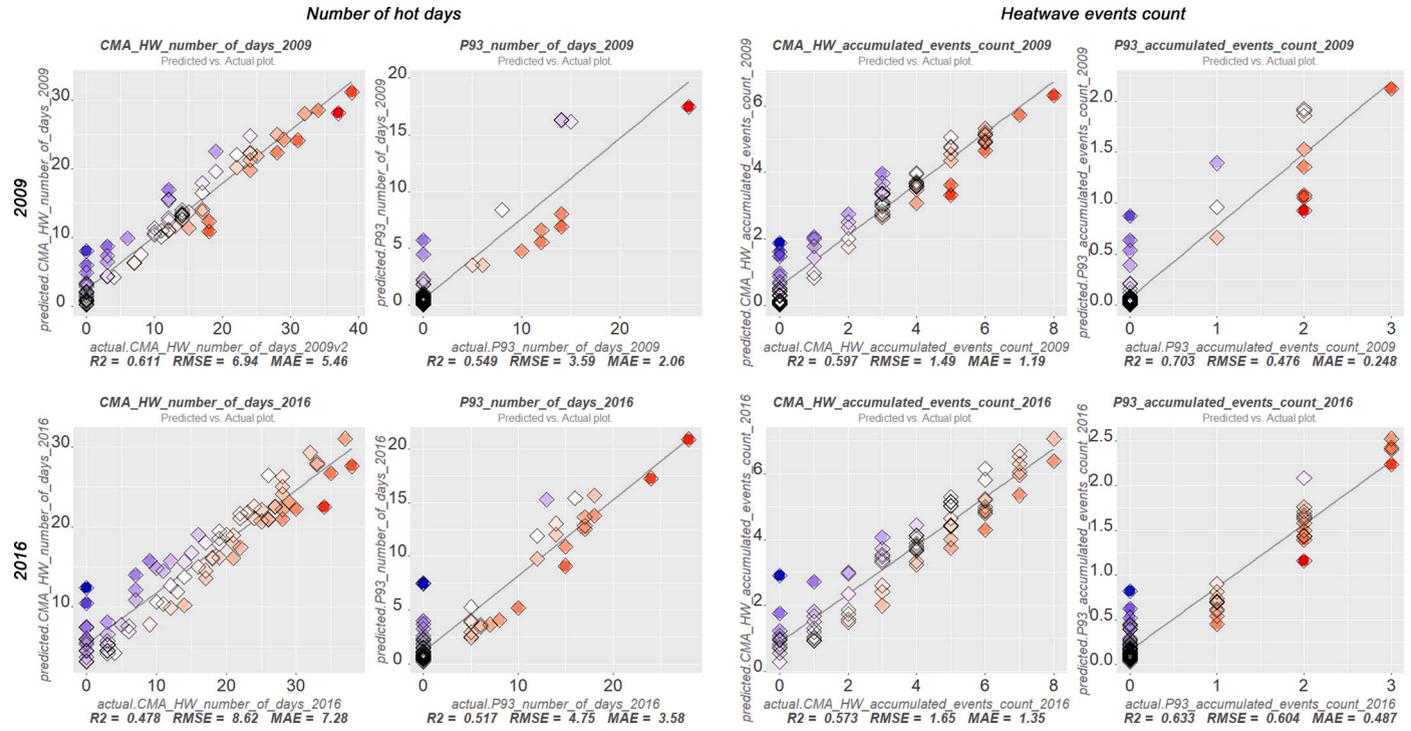


Fig. 3. The actual-by-predicted data plot of the number of hot days and heatwave event counts based on CMA and P93 heatwave definitions, in the year of 2009 and 2016, respectively. The R^2 , MAE, RMSE of corresponding models are shown under each plot.

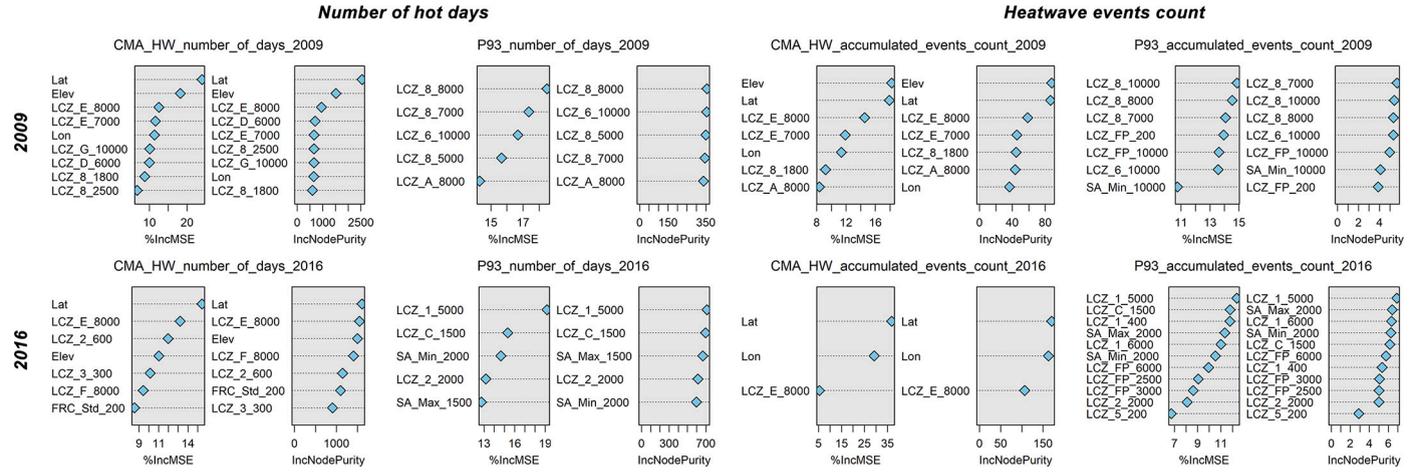


Fig. 4. The variable importance plot (%IncMSE and IncNodePurity sorted decreasingly from top to bottom) of all resultant models.

c

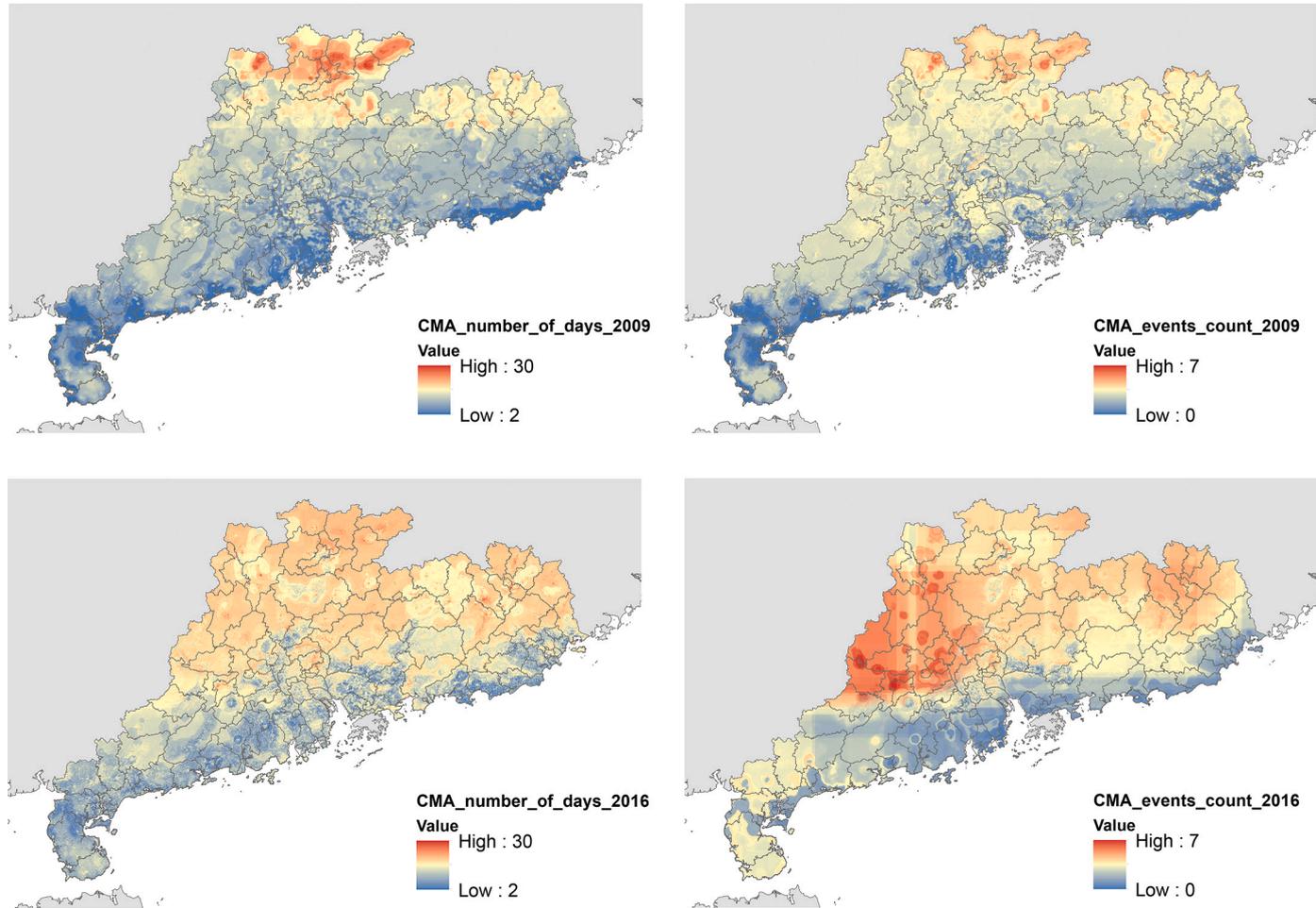


Fig. 5. Spatial prediction maps of the number of hot days and heatwave event counts based on CMA heatwave definitions, in the year of 2009 and 2016, respectively. For the number of hot days, only those days belongs to CMA heatwave events are counted. For example, if a day has a maximum temperature ≥ 35 °C, but not belongs to a ≥ 3 -day consecutive series, then it is not counted.

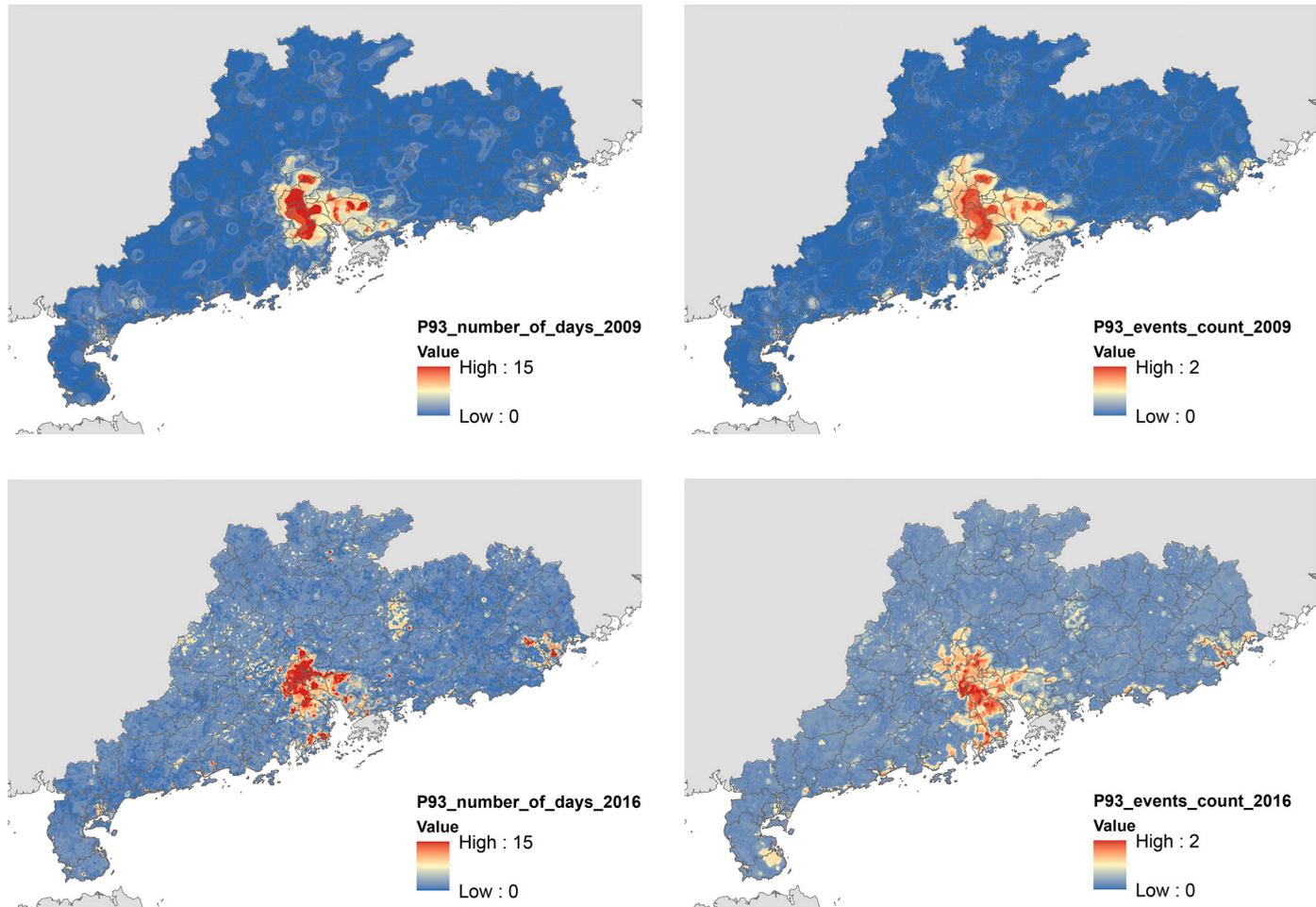


Fig. 6. Spatial prediction maps of the number of hot days and heatwave event counts based on P93 heatwave definitions, in the year of 2009 and 2016, respectively.

respectively. The eight resultant models explain 47.8% to 70.3% of the spatial variability of heatwave conditions. Most of the resultant models can explain approximately 60% of the spatial variability. Fig. 4 shows the variable importance measured and ordered based on Mean Decrease Accuracy (%IncMSE) and Mean Decrease Gini (IncNodePurity).

3.3. Spatial mapping of heatwave conditions

Using the model formulas shown in Table 1 and the predictor data layers described in Section 2.2.3, we generate the spatial prediction maps of the number of hot days and heatwave event counts. As a result, eight maps were generated using the resolution of 100×100 m for the entire spatial extent of the study area (Figs. 5 and 6).

4. Discussion

4.1. Influential LCZ types and UCPs

In the eight resultant models, some LCZ types have been found that are commonly included, which are LCZ 8 (large low-rise buildings), LCZ E (bare rock, paved surface), LCZ G (water body). Besides the above LCZ types, LCZ 6 (open low-rise buildings) and LCZ A (dense trees) were also included by some of the resultant models. Two UCPs were found to be influential factors, i.e., impervious surface fraction, surface albedo. Impervious ground surfaces in LCZ E, especially those artificial surfaces such as the pavements of concrete, asphalt store more heat compared to natural pervious land surfaces (Arnold and Gibbons, 1996). Surface albedo has been proven to be a crucial factor in urban air temperature (Susca et al., 2011; Taha et al., 1988). In the urban context of China, LCZ 8 mostly contains those buildings with a very large area of flat dark-colored rooftop of low surface albedo, or the color-coated steel sheet roofing. The thermal properties of roofing materials determine the surface energy balance, consequently alters the atmospheric heating, thus affect ambient air temperature (Coutts et al., 2013). LCZ A and LCZ G represent dense tree-dominated green space and blue space. The blue-green space has strong and complex relationships with the urban canopy and boundary-layer temperatures and has a significant impact on it (Garuma, 2018; Gunawardena et al., 2017; Morris et al., 2016). Therefore, it is reasonable that the above candidate predictor variables are identified by the data-driven variable selection process and included in the resultant models.

4.2. Temporal change of the models between 2009 and 2016

Comparing the 2009 and 2016 models, we found that LCZ1 (compact high-rise buildings), LCZ2 (compact mid-rise buildings), and LCZ3 (compact low-rise buildings) which were not included in the 2009 resultant models are included in the 2016 resultant models. These LCZ types represent highly-urbanized compact built-up areas. During 2009–2016, Guangdong province continues a rapid development and urbanization process, as planned in the “PRD Region Reform and Development Planning Guidelines (2008–2020)” which was released by the local authorities (Shen and Kee, 2016; Yu-shek, 2018). Affected by the regional planning and development policies, the land area for urban development and construction is gradually expanding, consequently causes an increase in local temperature. In cities, the amount of high-rise buildings is increasing rapidly, and the building density is also increased. Such changes in the urban morphology are conducive to the intensification of urban heat islands. The inclusion of predictors about LCZ 1, LCZ2, and LCZ3 in the 2016 models implies that the effect of urbanization on heatwave conditions becomes stronger during the past several years. Such impacts of urbanization on local climate is also found in regional weather simulation for Guangdong (Tse et al., 2018).

4.3. Spatial distribution of heatwave conditions and differences between using the heatwave definitions of CMA and P93

As shown in Figs. 5 and 6, the resultant maps based on CMA and P93 heatwave definitions have noticeable differences in the spatial pattern. In the spatial prediction maps based on the CMA definition, it is observed that the total number of hot days and heatwave events counted in coastal areas of Guangdong province are less than those in the inland areas. This finding breaks through the common mindset that the higher the average temperature, the more heatwave events. Using the absolute definition like CMA definition, the spatial distribution of hot days/heatwave events does not have to be consistent with the climatological mean temperature distribution. In South China, it has been found that a consecutive series of several hot days have higher health impacts than a single hot day with an extremely high temperature, as people cannot be relieved and physically recover from the hot weather (Wang et al., 2019a). Therefore, the heatwave investigation consider not only the air temperature but also the duration of the events. In this study, for the number of hot days, only those days belongs to CMA heatwave events are counted. For example, if a day has a maximum temperature $\geq 35^\circ\text{C}$, but not belongs to a ≥ 3 -day consecutive series, then it is not counted. This has brought a counter-intuitive result that there are more heat waves in the northern part of the study area. It is also found that the number of hot days and heatwave events have a significant correlation with latitude and longitude (particularly in the map of heatwave events count in 2016). It is commonly known that the air temperature in a specific geographic location highly depends on the latitude and whether the location is near to the coastal area. The CMA definition uses a fixed threshold of air temperature (a maximum daily temperature $\geq 35^\circ\text{C}$) to define hot days. Therefore, it is reasonable that the spatial distribution of heatwave conditions based on CMA definition is strongly correlated with geolocation predictors (latitude and longitude). Using a fixed threshold of air temperature to investigate the spatial pattern of heatwave conditions in a relatively large spatial extent could introduce bias. Specifically, it might overestimate the heatwave condition and relevant environmental and health risks in low latitudes and underestimate the situation in high latitudes, as the surface air temperature is

already a function of latitude, and those people live in different latitudes has varying tolerance and adaptations to heat (McCarthy et al., 2001). Unlike CMA based results, the effect of urbanization can be clearly observed in the resultant prediction maps produced using P93 definition. The hotspot shows in Fig. 6 is the location of Guangzhou which is the capital city and also the most urbanized and the most populated city of Guangdong province. It has been found that city areas do have more very hot days than rural areas due to the high-density building clusters trap heat within the city during nighttime (Shi et al., 2019). Therefore, it is reasonable that Fig. 6 shows a distribution pattern that similar to the urbanized areas and urban heat island effect.

4.4. Limitations and future works

There are still certain limitations in the present study, which could be further overcome in future works. First, in this study, only moderate prediction performance is achieved in the estimation of heatwave spatial patterns. The resultant model explains approximately 50–60% of the spatial variability in heatwave conditions. In this study, the automatic variable selection process identifies five to eleven important variable predictors in most cases. However, the variable selection process only identifies three predictors in the analysis of CMA heatwave events count in 2016, which are Latitude, Longitude, and LCZ_E 8000. It can be observed that the inclusion of Latitude, Longitude as spatial predictors leads to the appearance of edge in the prediction map of CMA heatwave events count in 2016. In that case, future work should focus on further fine-tuning the model to improve the prediction performance. Currently, the models are developed based on data from 86 locations (the 86 major national standard meteorological stations) in the study area. The data amount is usable but not sufficient for the development of very high-performance spatial models yet. The model development would be beneficial from a denser spatial distribution of input data. In future work, the model will be fine-tuned by using hundreds of automatic weather stations (AWS) that are densely distributed in the study area. Second, in this study, so far, only the areal proportion of LCZ types was calculated and used as predictors. The spatial pattern of the configuration (i.e., the evenness, fragmentation, clustering, which can be quantified by means of landscape ecology methods) of specific LCZ types not been analyzed yet by the present study. More complex spatial pattern analysis for WUDAPT could be introduced for generating more useful predictor variables in future studies to improve the model performance. Last but not least, this pilot study is currently only conducted in subtropical regions, which means that more tests should be conducted in other areas under different climatic zones in order to verify and increase the worldwide applicability of the proposed approach.

5. Conclusion

Investigating the spatial distribution of heatwave conditions is essential to the evaluation of the heat-related vulnerability and relevant potential social and economic impacts, especially for regions that are in rapid urbanization and economic development. In this study, we present an effective approach to spatially estimate the heatwave patterns using machine learning and WUDAPT. The approach enables direct mapping of the spatial distribution of heatwave conditions by taking advantage of open urban data, which incorporates spatial heterogeneity into the estimation of heatwave. The resultant models and high-resolution spatial maps of heatwave generated by this study enhance the spatial assessment of heat-health risk, and also provide valuable clues for how the cities should be properly planned to enhance the resilience to heatwave events and heat-related disasters. The study also put up a new perspective and a feasible pathway of utilizing WUDAPT to facilitate urban environment applications.

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.

Acknowledgement

This research is supported by the General Research Fund (RGC Ref No. 14610717) and Research Impact Fund (RGC Ref No. R4046-18F) from the Research Grants Council (RGC) of Hong Kong. The authors appreciate reviewers for their insightful comments and constructive suggestions on our research work. The authors also want to thank editors for their patient and meticulous work for our manuscript.

References

- Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *Int. J. Climatol.* 23, 1–26.
- Arnold, C.L., Gibbons, C.J., 1996. Impervious surface coverage: the emergence of a key environmental indicator. *J. Am. Plan. Assoc.* 62, 243–258.
- Bao, J., et al., 2015. The construction and validation of the heat vulnerability index, a review. *Int. J. Environ. Res. Public Health* 12, 7220–7234.
- Bechtel, B., et al., 2015. Mapping local climate zones for a worldwide database of the form and function of cities. *ISPRS Int. J. Geo Inf.* 4, 199.
- Bechtel, B., et al., 2016. Towards consistent mapping of urban structure - global human settlement layer and local climate zones. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B8, pp. 1371–1378.
- Bechtel, B., et al., 2019. Generating WUDAPT level 0 data – current status of production and evaluation. *Urban Clim.* 27, 24–45.
- Benmarhnia, T., et al., 2015. Review article: vulnerability to heat-related mortality. *Epidemiology*. 26, 781–793.
- Brandsma, T., Wolters, D., 2012. Measurement and statistical modeling of the urban Heat Island of the City of Utrecht (the Netherlands). *J. Appl. Meteorol. Climatol.* 51, 1046–1060.

- Breiman, L., 1996. Bagging predictors. *Mach. Learn.* 24, 123–140.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Bunyavanich, S., et al., 2003. The impact of climate change on child health. *Ambul. Pediatr.* 3, 44–52.
- Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika*. 76, 503–514.
- Buscaill, C., et al., 2012. Mapping heatwave health risk at the community level for public health action. *Int. J. Health Geogr.* 11, 38.
- Campbell, S., et al., 2018. Heatwave and health impact research: a global review. *Health Place* 53, 210–218.
- Chan, E.Y., et al., 2012. A study of intracity variation of temperature-related mortality and socioeconomic status among the Chinese population in Hong Kong. *J. Epidemiol. Community Health* 66, 322–327.
- Chapman, S., et al., 2017. The impact of urbanization and climate change on urban temperatures: a systematic review. *Landsch. Ecol.* 32, 1921–1935.
- Chen, F., et al., 2011. The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.* 31, 273–288.
- Ching, J., et al., 2014. WUDAPT: Facilitating Advanced Urban Canopy Modeling for Weather, Climate and Air Quality Applications.
- Ching, J., et al., 2018. WUDAPT: an urban weather, climate, and environmental modeling infrastructure for the Anthropocene. *Bull. Am. Meteorol. Soc.* 99, 1907–1924.
- Comrie, A.C., 2000. Mapping a wind-modified urban Heat Island in Tucson, Arizona (with comments on integrating research and undergraduate learning). *Bull. Am. Meteorol. Soc.* 81, 2417–2431.
- Coutts, A.M., et al., 2013. Assessing practical measures to reduce urban heat: green and cool roofs. *Build. Environ.* 70, 266–276.
- Dugord, P.-A., et al., 2014. Land use patterns, temperature distribution, and potential heat stress risk – the case study Berlin, Germany. *Comput. Environ. Urban. Syst.* 48, 86–98.
- El-Zein, A., Tonmoy, F.N., 2015. Assessment of vulnerability to climate change using a multi-criteria outranking approach with application to heat stress in Sydney. *Ecol. Indic.* 48, 207–217.
- Epstein, P.R., Mills, E., 2005. *Climate Change Futures: Health, Ecological and Economic Dimensions*. The Center for Health and the Global Environment, Harvard Medical School.
- Field, C.B., 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Fraser, A.M., et al., 2018. Strategic locating of refuges for extreme heat events (or heat waves). *Urban Clim.* 25, 109–119.
- Gál, T., et al., 2015. Comparison of Two Different Local Climate Zone Mapping Methods.
- García-Herrera, R., et al., 2010. A review of the European summer heat wave of 2003. *Crit. Rev. Environ. Sci. Technol.* 40, 267–306.
- Garuma, G.F., 2018. Review of urban surface parameterizations for numerical climate models. *Urban Clim.* 24, 830–851.
- Genuer, R., et al., 2010. Variable selection using random forests. *Pattern Recogn.* Lett. 31, 2225–2236.
- Genuer, R., et al., 2015. VSURF: An R Package for Variable Selection Using Random Forests.
- Gronlund, C.J., et al., 2015. Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 1990–2007. *Environ. Res.* 136, 449–461.
- Gunawardena, K.R., et al., 2017. Utilising green and bluespace to mitigate urban heat island intensity. *Sci. Total Environ.* 584–585, 1040–1055.
- Guo, A., et al., 2020. Influences of urban spatial form on urban heat island effects at the community level in China. *Sustain. Cities Soc.* 53, 101972.
- Haines, A., et al., 2006. Climate change and human health: impacts, vulnerability and public health. *Public Health* 120.
- Hart, M.A., Sailor, D.J., 2009. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theor. Appl. Climatol.* 95, 397–406.
- Ho, H.C., et al., 2015. A spatial framework to map heat health risks at multiple scales. *Int. J. Environ. Res. Public Health* 12, 16110–16123.
- IPCC, 2014. *Climate Change 2014—Impacts, Adaptation and Vulnerability: Regional Aspects*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Johnson, D.P., et al., 2012. Developing an applied extreme heat vulnerability index utilizing socioeconomic and environmental data. *Appl. Geogr.* 35, 23–31.
- Johnson, S., et al., 2020. Characterization of intra-urban spatial variation in observed summer ambient temperature from the new York City Community air survey. *Urban Clim.* 31, 100583.
- Kaiser, R., et al., 2007. The effect of the 1995 heat wave in Chicago on all-cause and cause-specific mortality. *Am. J. Public Health* 97, S158–S162.
- Kenny, G.P., et al., 2010. Heat stress in older individuals and patients with common chronic diseases. *Can. Med. Assoc. J.* 182, 1053.
- Klein Rosenthal, J., et al., 2014. Intra-urban vulnerability to heat-related mortality in new York City, 1997–2006. *Health Place* 30, 45–60.
- Kljun, N., et al., 2004. A simple parameterisation for flux footprint predictions. *Bound.-Layer Meteorol.* 112, 503–523.
- Konarska, J., et al., 2016. Influence of vegetation and building geometry on the spatial variations of air temperature and cooling rates in a high-latitude city. *Int. J. Climatol.* 36, 2379–2395.
- Kovats, R.S., Hajat, S., 2008. Heat stress and public health: a critical review. *Annu. Rev. Public Health* 29, 41–55.
- Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26.
- Kysely, J., 2002. Temporal fluctuations in heat waves at Prague–Klementinum, the Czech Republic, from 1901–97, and their relationships to atmospheric circulation. *Int. J. Climatol.* 22, 33–50.
- Landsberg, H.E., 1981. *The Urban Climate*. Academic Press, London.
- Le Tertre, A., et al., 2006. Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*. 17.
- Leconte, F., et al., 2015. Using local climate zone scheme for UHI assessment: evaluation of the method using mobile measurements. *Build. Environ.* 83, 39–49.
- Lemonsu, A., et al., 2015. Vulnerability to heat waves: impact of urban expansion scenarios on urban heat island and heat stress in Paris (France). *Urban Clim.* 14, 586–605.
- Li, D., Bou-Zeid, E., 2013. Synergistic interactions between urban Heat Islands and heat waves: the impact in cities is larger than the sum of its parts. *J. Appl. Meteorol. Climatol.* 52, 2051–2064.
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News.* 2, 18–22.
- Lin, Q., et al., 2019. The definition of heat-wave based on mortality risk assessment in different regions of China. *Zhonghua yu Fang yi Xue za Zhi [Chin. J. Prev. Med.]* 53, 97–102.
- Liu, X., et al., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google earth engine platform. *Remote Sens. Environ.* 209, 227–239.
- Luo, M., Lau, N.-C., 2016. Heat waves in southern China: synoptic behavior, long-term change, and urbanization effects. *J. Clim.* 30, 703–720.
- Ma, W., et al., 2015. The short-term effect of heat waves on mortality and its modifiers in China: an analysis from 66 communities. *Environ. Int.* 75, 103–109.
- Maragno, D., et al., 2020. Mapping heat stress vulnerability and risk assessment at the neighborhood scale to drive urban adaptation planning. *Sustainability*. 12, 1056.
- Maughan, R.J., 2012. Hydration, morbidity, and mortality in vulnerable populations. *Nutr. Rev.* 70, S152–S155.
- Mayrhuber, E.A.-S., et al., 2018. Vulnerability to heatwaves and implications for public health interventions – a scoping review. *Environ. Res.* 166, 42–54.
- McCarthy, J.J., et al., 2001. *Climate Change 2001: Impacts, Adaptation, and Vulnerability: Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*. 305, 994–997.
- Mills, G., et al., 2015. An Introduction to the WUDAPT project. In: ICUC9 - 9th International Conference on Urban Climate jointly with 12th Symposium on the Urban Environment. The International Association for Urban Climate (IAUC) and the American Meteorological Society (AMS) Toulouse, France.
- Morris, K.I., et al., 2016. Effect of vegetation and waterbody on the garden city concept: an evaluation study using a newly developed city, Putrajaya, Malaysia. *Comput. Environ. Urban. Syst.* 58, 39–51.

- Nairn, J., Fawcett, R., 2011. Defining heatwaves: heatwave defined as a heat-impact event servicing all. *Europe*. 220, 224.
- Oke, T.R., 1973. City size and the urban heat island. *Atmos. Environ.* 7, 769–779 (1967).
- Oke, T.R., 1982. The energetic basis of the urban heat island. *Quart. J. R. Met. Soc.* 108.
- Oke, T.R., 1987. *Boundary Layer Climates*, 2nd ed. Routledge, London and New York.
- Oke, T.R., 1988. The urban energy balance. *Prog. Phys. Geogr.* 12, 471–508.
- Oke, T.R., 1997. *Urban Environments*. McGill-Queen's University Press.
- Oleson, K.W., et al., 2015. Interactions between urbanization, heat stress, and climate change. *Clim. Chang.* 129, 525–541.
- Oudin Åström, D., et al., 2011. Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. *Maturitas*. 69, 99–105.
- Probst, P., et al., 2019. Hyperparameters and tuning strategies for random forest. *WIREs Data Min. Knowledge Discov.* 9, e1301.
- Ren, C., et al., 2019. Assessment of local climate zone classification maps of cities in China and feasible refinements. *Sci. Rep.* 9, 18848.
- Ren, C., et al., 2020. Developing a rapid method for 3-dimensional urban morphology extraction using open-source data. *Sustain. Cities Soc.* 53, 101962.
- Rey, G., et al., 2009. Heat exposure and socio-economic vulnerability as synergistic factors in heat-wave-related mortality. *Eur. J. Epidemiol.* 24, 495–502.
- Salamanca, F., et al., 2011. A study of the urban boundary layer using different urban parameterizations and high-resolution urban canopy parameters with WRF. *J. Appl. Meteorol. Climatol.* 50, 1107–1128.
- Schatz, J., Kucharik, C.J., 2014. Seasonality of the urban Heat Island effect in Madison, Wisconsin. *J. Appl. Meteorol. Climatol.* 53, 2371–2386.
- Sharma, A., et al., 2017. Urban meteorological modeling using WRF: a sensitivity study. *Int. J. Climatol.* 37, 1885–1900.
- Shen, J., Kee, G., 2016. *Development and Planning in Seven Major Coastal Cities in Southern and Eastern China*. Springer International Publishing.
- Shi, Y., et al., 2018a. Modelling the fine-scale spatiotemporal pattern of urban heat island effect using land use regression approach in a megacity. *Sci. Total Environ.* 618, 891–904.
- Shi, Y., et al., 2018b. Evaluating the local climate zone classification in high-density heterogeneous urban environment using mobile measurement. *Urban Clim.* 25, 167–186.
- Shi, Y., et al., 2019. Assessing spatial variability of extreme hot weather conditions in Hong Kong: a land use regression approach. *Environ. Res.* 171, 403–415.
- Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. *Bull. Am. Meteorol. Soc.* 93, 1879–1900.
- Stocker, T., 2014. *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Sun, Y., et al., 2016. Contribution of urbanization to warming in China. *Nat. Clim. Chang.* 6, 706–709.
- Susca, T., et al., 2011. Positive effects of vegetation: urban heat island and green roofs. *Environ. Pollut.* 159, 2119–2126.
- Taha, H., 1997. Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. *Energy Build.* 25, 99–103.
- Taha, H., et al., 1988. Residential cooling loads and the urban heat island—the effects of albedo. *Build. Environ.* 23, 271–283.
- Tan, J., et al., 2010. The urban heat island and its impact on heat waves and human health in Shanghai. *Int. J. Biometeorol.* 54, 75–84.
- Tse, J.W.P., et al., 2018. Investigation of the meteorological effects of urbanization in recent decades: a case study of major cities in Pearl River Delta. *Urban Clim.* 26, 174–187.
- Uejio, C.K., et al., 2011a. Intra-urban societal vulnerability to extreme heat: the role of heat exposure and the built environment, socioeconomics, and neighborhood stability. *Health Place* 17, 498–507.
- Uejio, C.K., et al., 2011b. Intra-urban societal vulnerability to extreme heat: the role of heat exposure and the built environment, socioeconomics, and neighborhood stability. *Health Place* 17, 498–507.
- UN, 2018. *World Urbanization Prospects: The 2018 Revision*. United Nations, Department of Economic and Social Affairs (DESA), Population Division, Population Estimates and Projections Section. United Nations, Department of Economic and Social Affairs (DESA), Population Division, Population Estimates and Projections Section, New York.
- UN, 2019. *World Population Prospects, The 2019 Revision*. United Nations, Department of Economic and Social Affairs, Population Division.
- Voelkel, J., Shandas, V., 2017. Towards systematic prediction of urban Heat Islands: grounding measurements, assessing modeling techniques. *Climate*. 5, 41.
- Wang, D., et al., 2019a. The impact of extremely hot weather events on all-cause mortality in a highly urbanized and densely populated subtropical city: a 10-year time-series study (2006–2015). *Sci. Total Environ.* 690, 923–931.
- Wang, R., et al., 2018. Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: a case study of Hong Kong. *Urban Clim.* 24, 567–576.
- Wang, R., et al., 2019b. Detecting multi-temporal land cover change and land surface temperature in Pearl River Delta by adopting local climate zone. *Urban Clim.* 28, 100455.
- WMO, 2008. *Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8)*. Secretariat of the World Meteorological Organization, Geneva, Switzerland.
- WMO, 2011. *Weather Extremes in a Changing Climate: Hindsight on Foresight*. World Meteorological Organization Geneva.
- WMO, WHO, 2015. *Heatwaves and Health: Guidance on Warning-System Development*. World Meteorological Organization.
- Wolf, T., McGregor, G., 2013. The development of a heat wave vulnerability index for London, United Kingdom. *Weather Clim. Extrem.* 1, 59–68.
- Wong, D.W.S., 2004. The Modifiable Areal Unit Problem (MAUP). In: Janelle, D.G., et al. (Eds.), *WorldMinds: Geographical Perspectives on 100 Problems: Commemorating the 100th Anniversary of the Association of American Geographers 1904–2004*. Springer Netherlands, Dordrecht, pp. 571–575.
- Xiang, J., et al., 2013. Health Impacts of Workplace Heat Exposure: An Epidemiological Review. *Industrial Health (Advpub)*.
- Xu, Z., et al., 2012. Impact of ambient temperature on children's health: a systematic review. *Environ. Res.* 117, 120–131.
- Yu-shek, C.J., 2018. *Development of Guangdong, The China's Economic Powerhouse*. World Scientific Publishing Company.
- Zhang, Y., et al., 2017. Global climate change: impact of heat waves under different definitions on daily mortality in Wuhan, China. *Glob. Health Res. Pol.* 2, 10.