Building and Environment 158 (2019) 39–50

Contents lists available at ScienceDirect

Building and Environment

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

A diagnostic equation for the maximum urban heat island effect of a typical Chinese city: A case study for Xi'an

Xi Zhang[a,b,*], Gert-Jan Steeneveld[b], Dian Zhou[a], Chengjiang Duan[a], Albert A.M. Holtslag[b]

[a] School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, No.99 Yanxiang Road, Xi'an, Shaanxi Province, China
[b] Meteorology and Air Quality Section, Wageningen University, P.O. Box 47, 6700 AA, Wageningen, The Netherlands

ARTICLE INFO

Keywords:
Urban heat island
Diagnostic equation
Urban morphology
Neighborhood-scale
Xi'an
China

ABSTRACT

To reduce the vulnerability of urban areas facing high temperatures, it is useful to develop methods to obtain the urban heat island (UHI) intensity. However, it is hard to equip all cities with extensive measurement networks and alternative UHI diagnostic methods are needed. Accordingly, in this paper we evaluate and revise the diagnostic equation designed by Theeuwes et al. (2017) and analyze its application for Xi'an (China), based on long-term summer meteorology data. The evaluation of the default diagnostic equation shows that limited accuracy for Xi'an is caused by the sum of morphological parameters being used outside the original range of calibration. Subsequently, we propose an extended equation, which adds the building fraction to express the morphology of additional spatial categories. This new equation is calibrated against a 3-year dataset and independently validated with data from another year. In addition, a class prediction with three spatial categories is proposed, and verified by independent data of 20 stations in Xi'an in 2018, which enables this formula to be applied in more cities. Altogether, the extended diagnostic equation is an effective method to evaluate the daily maximum UHI intensity (UHI_{max}) on neighborhood-scale, which can be generalized for a whole city area and presents internal intensity differences with multi-points.

1. Introduction

Due to global warming, the near surface air temperature is rising sensibly, especially in summer. Also, extreme heat waves occur more frequently and widely [1]. Urban areas are more vulnerable than rural areas facing a high temperature environment, due to the urban heat island effect (here we focus on the 2-m air temperature difference between rural and urban environment), which is caused by complex factors such as canyon structure, enhanced impervious surface, population agglomeration and so on [2–4]. Intense heat in cities may reduce thermal comfort and affect residents' daily commuting, outdoor activities and labor productivity [5.6]. In addition, frequent heatwaves (at least three consecutive days with daily maximum temperature exceeding 35 °C in China) [7], may increase the risk of diseases and excess mortality [8]. Therefore, an accurate, detailed and real-time estimation of the urban heat island intensity seems eminent, to improve the city's responsive capacity and ensure residents' safety and health. Unfortunately, equipping all cities with an extensive measurement network is hard to achieve [9–11], which indicates the need for an indirect but relatively accurate method to diagnose the real-time UHI intensity.

Firstly, the UHI is often estimated using thermal remote sensing observations, through the use of satellite, airborne and aircraft platforms [12]. The main concern in this approach is the comparatively low accuracy, due to limited horizontal resolution (30 m) [13]. Moreover, there are significant differences between the value of air temperature and surface temperatures measured by remote sensing [12].

Apart from observations, simulation approaches have been developed in UHI studies. Sophisticated atmospheric models coupled with an urban canopy model (UCM) are used to investigate the energy budget of the urban canopy layer (Ronda et al., 2017) [14]. Unfortunately, the complete 3D wind flow is not represented by these models [9]. Another simulation method is the computational fluid dynamics technique, which deals with the governing equations of the flow (Navier–Stokes equations) [14]. The modelling approaches has acceptable accuracy but for an entire city area are very computational expensive and require a large amount of information.

Alternatively, statistical models have been utilized because of their simplicity and ease to use. Most of these models are multiple linear regressions and are not purely physically based [15]. Furthermore, they are based on variables that are relatively difficult to measure and obtain, such as the boundary-layer height and the surface sensible heat flux in urban and rural stations [15].

* Corresponding author. School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, No.99 Yanxiang Road, Xi'an, Shaanxi Province, China.
E-mail addresses: celiazhangxi@foxmail.com, xizhang@stu.xjtu.edu.cn (X. Zhang).

https://doi.org/10.1016/j.buildenv.2019.05.004
Received 20 March 2019; Received in revised form 1 May 2019; Accepted 2 May 2019
Available online 03 May 2019

Therefore, in this study, we utilize a diagnostic equation for the UHI designed by Theeuwes et al. [15], which gives a simple and accurate way to diagnose the neighborhood -scale maximum UHI (the max difference between rural and urban temperature in a 24-h circle). This equation only requires routine meteorology data of a rural station and the basic morphology data of the urban area (as representative for an area with radius of 500 m). Data for two aspects estimated in separate parts are multiplied to obtain the UHI_{max} (maximum UHI, see section 2.2). Designed in this way, the equation easily represents the UHI intensity for the whole city area. Moreover, the diagnostic equation estimates to be in good agreement with field observations, according to the independent validation for 14 European cities varied in size [15]. Beyond Europe, it was also tested for 7 Local Climate Zones (LCZ) in Nanjing, China by Lingye Yao et al. [16]. Note that Local Climate Zone according to Stewart and Oke [21], comprises a new and systematic classification of field sites for heat island studies. Generally, the results in Nanjing show an acceptable level of accuracy but the UHI for some LCZ types is overestimated. This indicates that the application range of this equation may need to be further expanded, especially for the new spatial categories different from Europe.

Xi'an is chosen as the study area. As the capital of Shaanxi province, it is a typical Chinese city according to complex land use and diverse urban texture. Moreover, the UHI is prominent in Xi'an (the maximum UHI is about 9 K in July [17]) as contributed by the special topography of subtropical high basin and high-speed urban development and constructions. Earlier UHI studies in Xi'an, partly focused on the multiple timescale analysis of the UHI intensity and discussed the seasonal, monthly and daily variations of UHI tendency [17]. On the other hand, some studies aimed to divide the urban area of Xi'an into a series of zones with graded temperature differences by surface temperature. Being more inclined to a historical perspective, they tried to figure out the correlation between the temporal and spatial distribution pattern of UHI and urbanization, through comparing the UHI intensity with the built-up area, land-use type, vegetation cover and other indicators [18.19]. Taken together, most of these studies applied the remote sensing method, which limited the spatial accuracy and put more focus on historical development, and therefore lack the estimation of current UHI intensity.

In this study, long-term summer climate data of Xi'an are used to test the performance of the UHI_{max} diagnostic equation (see section 3). Then considering the results of bias analysis, the extended equation for different spatial categories is proposed and verified (see section 4.1). Finally, the class prediction with spatial categories based on the formula are confirmed (see section 4.2), which may enable this UHI calculation method being more widely used in more diverse space and cities.

2. Materials and methods

2.1. Research area

Xi'an is located in the central part of the mainland of China (107.4°-109.5°E, 33.4–34.5°N), and the hinterland of Guanzhong basin (Fig. 1). The topography is generally high in the southeast (690 m at highest area on the hill) and low in the northwest and southwest (307 m at lowest area near riverside), with an average altitude of 410 m above sea level [20]. The Beishan Mountains in the north forms a natural barrier around the city area, which also leads to restriction of heat diffusion and aggravating the urban heat island effect.

This paper takes the central urban area of Xi'an as the research object, which covered an urbanized area of 832.17 km^2 and supported a population of 4.67 million in 2016 [20]. Xi'an experiences four distinct seasons. The summers (June, July and August) are hot and the summer average air temperature of 2014–2017 amounts to 27.1 °C [20]. The dominant wind direction in central urban area is the northeast wind, and the recorded summer wind speed is around 2.5 m/s in the city [20].

2.2. Diagnostic equation for the urban heat island

In this paper, we use the diagnostic equation proposed by Theeuwes et al. [15]:

$$UHI_{max} = Ai^{*}\sqrt[4]{\frac{\mathbf{S}\downarrow \boldsymbol{DTR}^{3}}{\boldsymbol{U}}} \tag{1a}$$

Here UHI_{max} is the daily maximum value of UHI, $\mathbf{S}\downarrow$ is the mean downward shortwave radiation in kinematic units during the preceding day ($K{\cdot}ms^{-1}$). Moreover, DTR denotes the diurnal temperature range (K), to introduce a measure for the rural nocturnal atmospheric stability [15]; U denotes the 24 h average 10-m wind speed ($m{\cdot}s^{-1}$) at the rural weather station. Taken together, $\sqrt[4]{\mathbf{S}\downarrow \boldsymbol{DTR}^{3}/\boldsymbol{U}}$ represents the meteorology of the rural station.

$$Ai = 2 - SVF - F_{veg} \tag{1b}$$

Where SVF is the sky-view factor at the urban station and F_{veg} is the surface vegetation fraction around the station (in a radius of 500 m). They are used to express the urban characteristic jointly. Note that the equation was calibrated for $0 < F_{veg} < 0.4$, $0.2 < SVF < 0.9$ only [15].

Here, Ai denotes a morphological factor specific to each urban station and was originally proposed for typical cities in Northwest Europe. The design of this equation considers the fact that the background weather condition of urban stations are the same and their temperature differences are mainly caused by various urban characteristics. Accordingly, Ai could be considered as a comprehensive index that presents urban characteristics, such as street geometry and urban vegetation fraction, which influence the UHI generally.

2.3. Observations

(1) Dataset 1: urban station dataset from 2014 to 2017

This study uses long-term summer meteorological data, from 2014 to 2017 (Jun, July, August), of 11 meteorological stations in central urban area in Xi'an, provided by Xi'an Meteorological Bureau. Considering the equation's requirement of SVF and F_{veg}, 9 suitable stations are retained (see Table 1). Data of 2014–2016 are used for testing and analysis of Equation (1) (see section 3), data of 2017 are reserved for validating the modified equation in section 4.1.

(2) Dataset 2: urban station dataset in 2018

Except the above 8 stations in datasetc1 (ZL station is out of use in 2018 and not included in dataset 2), we obtained meteorological data for 12 additional urban stations for 16–25 July 2018, which was a heatwave period (Fig. 1). These stations were built at the same time as the ones in dataset 1, but provided by the Xi'an Meteorological Bureau this year. These data could be regarded as independent data to validate the class prediction with three spatial categories in 4.2.

The above 20 meteorological stations are evenly distributed in the central area and containing various morphology and land use, such as public service, commercial facility, middle schools, universities, and low-rise and mid-rise communities. Generally, the main spaces are characterized as LCZ5 and LCZ6, and some LCZ2 spaces located in newly high-intensity urban center as the XA and XZ stations, and LCZ 8 spaces in open landmark as the HZZX and FRY stations. Note that a number of mixed and inhomogeneous spaces existing in Xi'an, which is quite different from European cities. This is partly due to the fact that Xi'an as an ancient capital remained the old spatial pattern and preserved some node spaces with large-scale, as for the NM and ZL stations. On the other side, during the development process, some unplanned villages are still left inside the city, which are consisting of densely low-rise housing and bare soil or farmland, such as the LCB, SLP and MTK

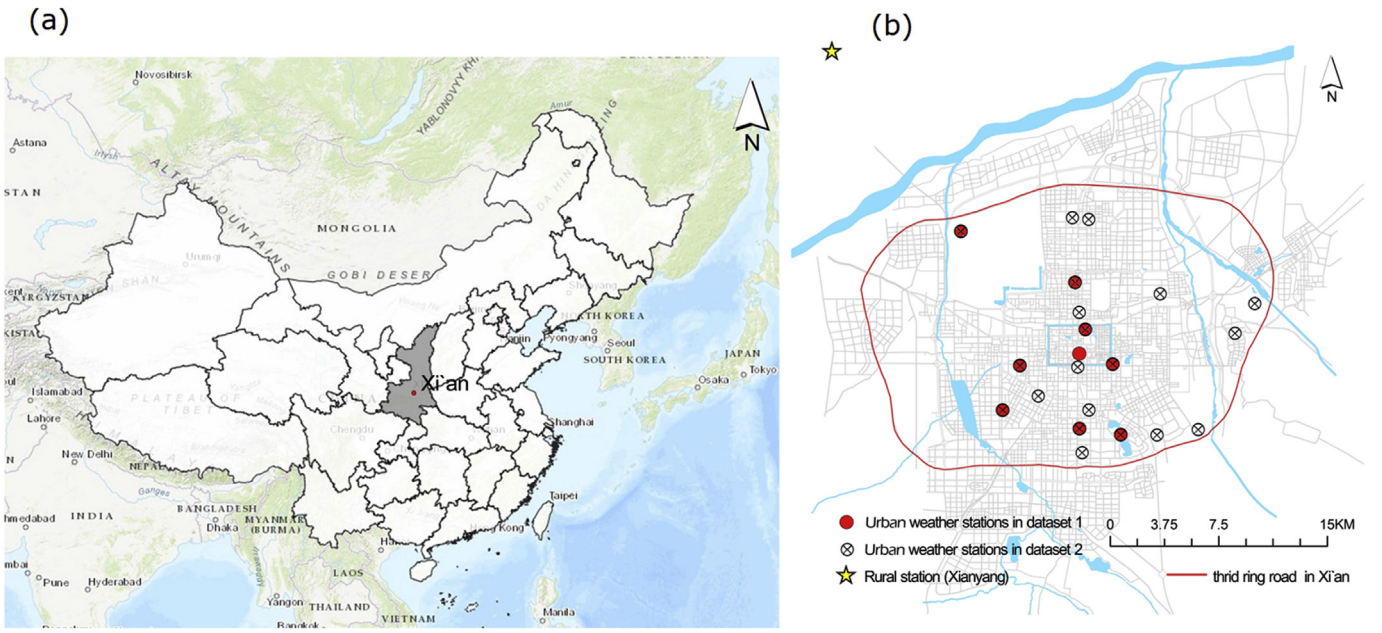

Fig. 1. (a) Map of China with the location of Xi'an. (b) Locations of meteorological stations in Xi'an, containing 2 data sets. The red dots indicate the stations in dataset 1 and the cross dots indicate dataset 2. The yellow star is the Xianyang airport in dataset 3, as the rural station. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

stations.

All meteorological data are collected by automatic weather stations (Fig. 4), and the specific models of each stations are listed (see Tables 1 and 2). These stations are located on open ground without shading of buildings or trees and record hourly data at 1.5 m, including air temperature, relative humidity and precipitation. The accuracy of the thermometer is ± 0.1 ˚C and of the hygrometer is ± 1%. Meanwhile, the SVF of each urban station was calculated using the Rayman model [29] and the F_{veg} was manually outlined in the satellite imagery and calculated using ArcGIS.

(3) Dataset 3: rural reference dataset

To compute the UHI_{max}, the observations from a rural station are also needed and Xianyang Station (34.45˚N, 108.75˚E) is used in this paper (Fig. 1b). It is located near Xianyang Airport over a grass surface (LCZ D, low plants). The meteorological data have been obtained from NOAA website (www.ncdc.noaa.gov), concluding the observations of hourly air temperature and wind speed at 10 m. The **S↓** is calculated by the Rayman model [29] with input information of time, location, meteorological data and obstacle file of Xianyang station. Also, in order to exclude frontal systems, rain events (daily sum > 0.3 mm) and fog events (daily maximum relative humidity > 80%) have been excluded, which is consistent with the requirements in Ref. [15]. Finally, 142 days are used for further analysis.

3. Results

3.1. Default evaluation of the diagnostic UHI_{max} equation

The performance of the diagnostic equation is displayed in Fig. 5, where the modelled UHI_{max} (as calculated by Equation (1a)) on the Y axis is compared with the observed UHI_{max} (e.g. the daily maximum of urban-rural temperature difference in observations) on the X axis. The Xi'an result (RMSE = 1.68 K, MEAE = 1.14 K) exhibits much more scatter and systematic differences than that of Europe (RMSE = 0.91 K, MEAE = 0.58 K) [15]. From the histogram of daily UHI_{max} bias (Fig. 6), we identify a number of large biases (bias > 1.5 K or bias < -1.5 K), which indicates that the equation needs further improvement. Obviously, large biases occur for stations which have morphological properties outside the original calibrated range.

Table 1
Summary of urban measurement stations in dataset 1.

Station Abr.	Main Land Use	Coordinates	Instrumentation	Distance from Rural Station,km	F_{veg}	SVF	F_{build}	Geomean Height,m	Population Density, person/km^2	LCZ
XA	Public Service	34.31˚N,108.94˚E	DZZ5	23.2	0.13	0.47	0.25	12.7	11304.6	2
XZ	Commercial Facility	34.21˚N,108.94˚E	CAWS600RE	31.3	0.14	0.44	0.29	12.4	30693.4	2
LCB	Low-rise Community	34.34˚N,108.85˚E	CAWS600RT	15.4	0.22	0.81	0.22	4.4	2785.2	6
FRY	Park	34.21˚N,108.98˚E	DZZ4	33.3	0.30	0.86	0.24	8.1	2667.8	6
ZL	Commercial Facility	34.26˚N,108.94˚E	CAWS600RE	27.1	0.13	0.88	0.45	12	7204.8	5
XQ	Mid-rise Community	34.26˚N,108.97˚E	CAWS600RT	29.1	0.14	0.79	0.30	13.2	33930.9	5
FQL	Public Service	34.25˚N,108.90˚E	WUSH-RG	25.4	0.25	0.48	0.23	11	15553.1	5
AZZX	Middle school	34.28˚N,108.95˚E	CAWS600RT	26.1	0.14	0.59	0.31	10.6	23526.1	5
XGX	Middle school	34.23˚N,108.89˚E	CAWS600RE	27.5	0.23	0.52	0.16	18.8	3148.8	5

The morphological indicators F_{veg}, SVF, F_{build} and geomean height are calculated around the meteorology stations within a radius of 500 m. LCZ is the local climate zone according to Stewart, I. D., & Oke, T. R. (2012) [21].

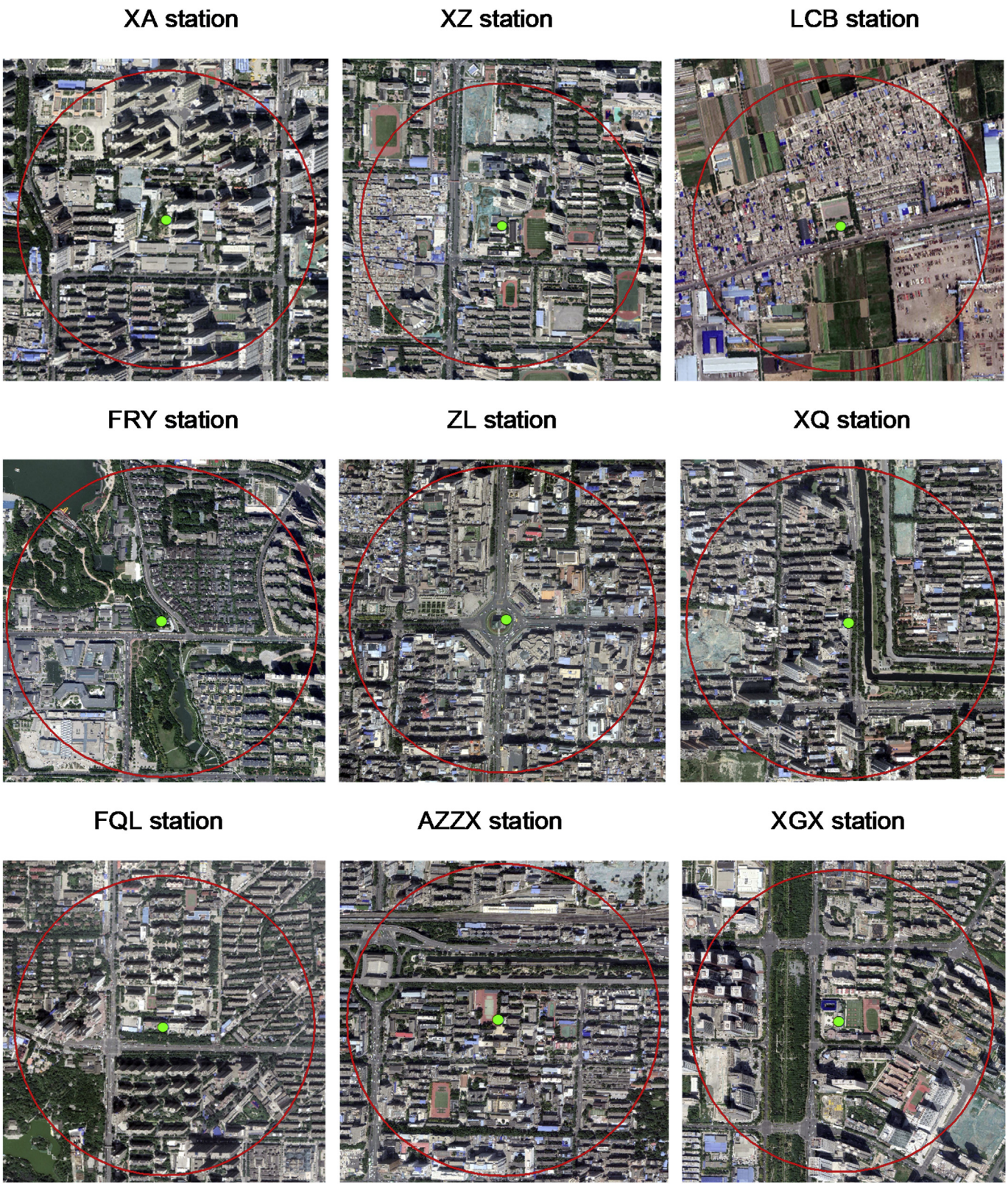

Fig. 2. Peripheral photographs of stations in database1. The green dot shows the location of station and the red circle represents a radius of 500 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.2. Bias analysis

More specifically, the results of the bias distribution with different stations (see Table 3) also indicate that the large biases mainly concentrate on the XA, XZ, FRY, LCB and ZL stations (Fig. 3). The XA and XZ stations have 39.5% and 43.5% overestimated UHI_{max} respectively. On the contrary, the UHI_{max} of the FRY (45.3%), LCB (31.6%) and ZL (81.7%) stations are underestimated.

To further explain the causes of biases at these stations and considering their obviously morphological difference with Europe, we

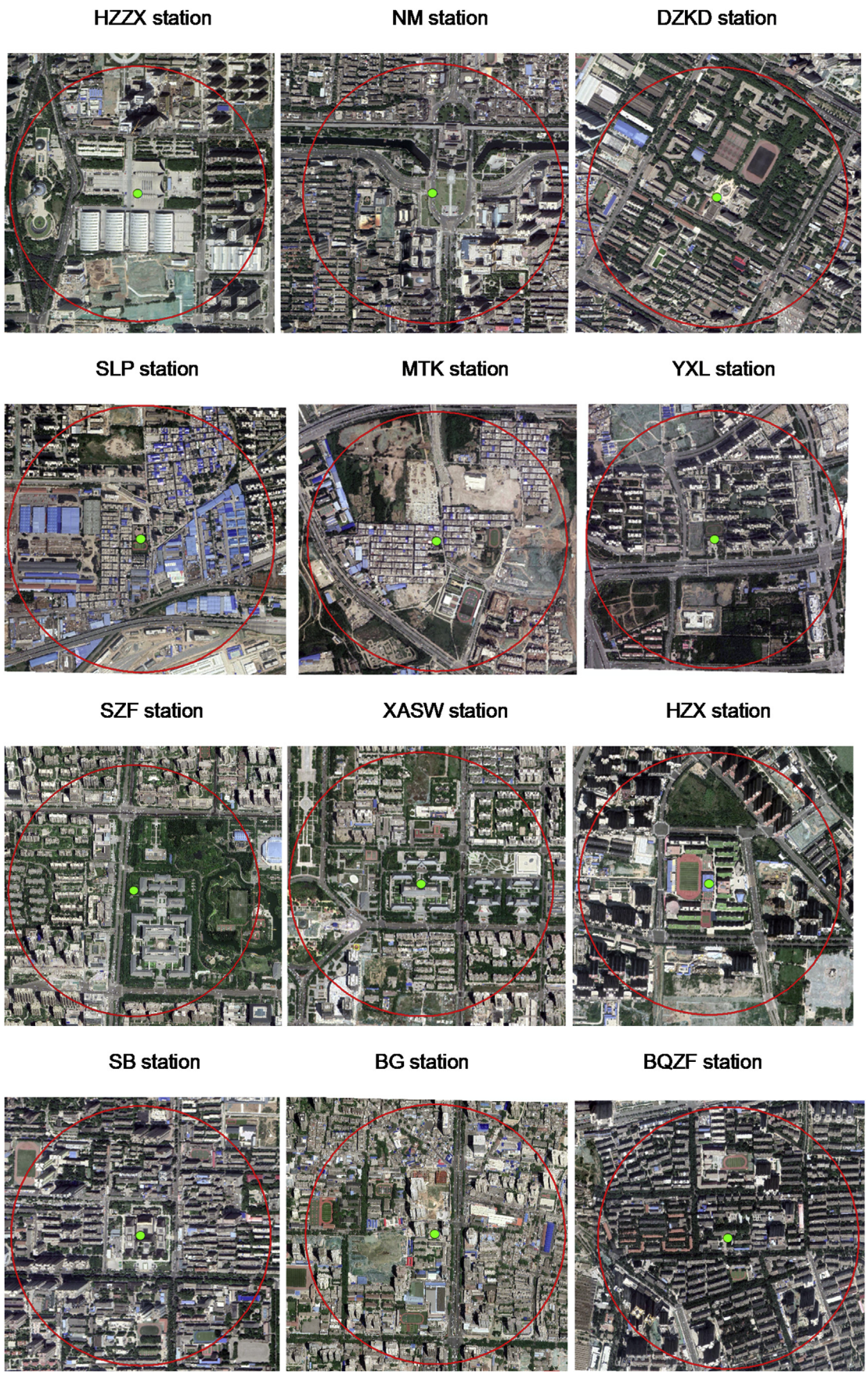

Fig. 3. Peripheral photographs of new added 12 stations in database 2. The green dot shows the location of station and the red circle represents a radius of 500 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4. Photos of three types of automatic stations in above meteorological stations.

calculate the *Observed Ai* according to Equation (1a), using a linear regression. The lines are indicated in light grey in Fig. 7. Meanwhile, the results are compared with the *Modelled Ai* according to Equation (1b), which is represented by linear regression lines in dark grey (Fig. 7). These two Ai should be very close originally. However, here the *Modelled Ai* are larger than the *Observed Ai* for the XA and XZ stations and smaller for the FRY, LCB and ZL stations, which contribute to the majority of large biases.

3.3. Causes of bias

The reason for the large deviations between *Observed Ai and Modelled Ai* is that the sum of morphological parameters (F_{veg} + SVF) is applied outside the original range of calibration in Theeuwes et al. (2017). Considering the definition of Ai, the value of the *Observed Ai* reflects the urban morphology around the station. Thereto *the Modelled Ai as* calculated by Equation (1b) is far from the *Observed Ai*, and as such unable to express these stations' morphology accurately.

From the peripheral photographs, the stations with overestimated UHI_{max} are dense areas with intensive constructions and numerical high-rise buildings (Fig. 2), which also contributes to a lower F_{veg} (0.14, 0.13). Accordingly, the SVF affected by the nearby high-rise buildings shows comparatively smaller (0.44, 0.47) value than the other stations. Conversely, the stations with underestimated UHI_{max}, FRY, LCB and ZL, have the highest SVF (0.86, 0.81, 0.88) according to the highly open morphology, which is different from the surroundings and relatively higher F_{veg}. The peripheral photographs indicate that various urban morphologies occur for these stations and present inhomogeneous spatial features. This could cause the morphological contrast between the open adjacent circle and the dense periphery circle of the ZL and FRY stations, or the non-construction area with extensive grassland and the construction area with densely low-rise buildings of the LCB station.

Thus far, taken the above description of spatial characteristics of these stations, we could name them as “Fully-developed space” (the XA and XZ stations) and “Mixed space” (the LCB, FRY and ZL stations). Also, the stations that perform well in the original equation could be named as “Normal space” below. The “Fully-developed space” and “Mixed space” are obviously different from low-density and homogeneous European cities, which were considered in the design of Equation (1b). So here, we develop two additional spatial categories with various urban morphology to extend the former research of Europe [15].

4. Extended model

4.1. Selection of morphological indices to modify equation

The above analysis reveals that the decline in modelled accuracy is caused by the “Fully-developed space” and “Mixed space”, but for “Normal space” the equation still performs well. It indicated that F_{veg} and SVF are key morphological indicators and should be reserved in Ai preferentially. Meanwhile, to narrow the disparity of *Observed Ai and Modelled Ai* in “Fully-developed space” and “Mixed space”, some other morphology index might be added to quantify their spatial characteristics more comprehensively.

Therefore, the correlation analysis is carried out using the *Observed Ai* of 9 stations in dataset 1 and various urban morphology indexes which proved to correlate with UHI intensity in former research [22–24]. Here (see Table 4), F_{build} shows the best relation with *Observed Ai* and is selected to modify Equation (1b) for additional spatial categories.

Consequently, we propose two modifications to Equation (1b) with $F_{build,}$ as adding it to the morphology part or replacing it with one of the existing indicators (see Table 5). Meanwhile, the modifications need to consider the departure of *Observed Ai* and *Modelled Ai*, which is about 0.3 in “Fully-developed space” and −0.3 in “Mixed space”. Note that the equation should also obey the physical rule that when there are no buildings, the value of Ai should be zero and the UHI_{max} will vanish (see Table 5). In such a situation, the morphology of urban area is the same as the rural area, and no air temperature difference will still exist.

According to the three equations of Table 5, we adjust the value of *Modelled Ai* separately (see Table 6) and bring them into Equation (1a). In Fig. 8, the observed UHI_{max} is compared with the modelled UHI_{max}. Equation (2) has the comparably best performance (Fig. 8a), with less peripheral scatters and enhanced linear characteristics. Equation 3 has a number of underestimated values (Fig. 8b). Especially when the observed UHI_{max} is higher than 6 K, larger deviations appear than for Equation (1). The results of Equation 4 are shown in Fig. 8c, the number of scatters and distance from 1:1 line increase simultaneously.

Overall, Equation (2) best represents the UHI_{max} for all stations. In addition, we compare the 95% confidence interval of the biases for the XA, XZ, LCB, FRY and ZL stations (Fig. 9a) before and after modification. For example, the interval of XA station moved from 0.9 to 1.4 K to −0.1–0.4 K, which means substantially large biases have been corrected. Also, the validation of two equations with independent data in 2017 is showed in Fig. 9b. Equation (2) (RMSE = 1.35 K, MEAE = 1.01 K) has a better performance than Equation (1)

Table 2
Summary of 12 additional urban measurement stations in dataset 2.

Station Abr.	Main Land Use	Coordinates	Instrumentation	Distance from Rural Station,km	F_{veg}	SVF	F_{build}	Geomean Height,m	Population Density, person/km2	LCZ
SZF	Public Service	34.34°N,108.94°E	CAWS600RE	20.3	0.28	0.81	0.19	16.1	11304.6	5
HZZX	Public Service	34.20°N,108.95°E	CAWS599RE	32.9	0.15	0.90	0.23	9.5	12199.5	8
NM	Commercial Facility	34.25°N,108.94°E	CAWS600RE	27.9	0.15	0.87	0.30	12.3	19979.1	5
DZKD	University campus	34.23°N,108.91°E	DZZ5	28.0	0.23	0.77	0.25	13.1	13440.4	5
SLP	Low-rise Community	37.30°N,109.00°E	CAWS600RE	28.3	0.15	0.85	0.35	7.2	7204.8	6
YXL	Mid-rise Community	34.21°N,109.00°E	CAWS600RE	35.0	0.29	0.74	0.12	14.0	2667.8	5
MTK	Low-rise Community	34.21°N,109.03°E	CAWS600RE	36.6	0.23	0.77	0.16	5.9	5634.2	6
XASW	Public Service	34.34°N,108.95°E	DZZ4	21.3	0.28	0.46	0.17	12.3	11304.6	5
BHZX	Middle school	34.29°N,109.07°E	DZZ4	34.2	0.20	0.45	0.15	17.7	1883.1	5
SB	Public Service	34.23°N,108.95°E	CAWS600RE	30.6	0.14	0.48	0.29	14.8	30693.4	2
BG	Mid-rise Community	34.29°N,108,94°E	CAWS600RE	25.0	0.12	0.49	0.33	11.5	35148.1	2
BQZF	Mid-rise Community	34.27°N,109.06°E	CAWS600RE	34.2	0.14	0.56	0.29	13.0	18487.0	5

The morphological indicators, F_{veg}, SVF, F_{build} and geomean height are calculated around the meteorology stations within a radius of 500 m. LCZ is the local climate zone according to Stewart, I. D., & Oke, T. R. (2012) [21].

Fig. 5. The observed UHI_{max} against modelled UHI_{max} using Equation (1) for stations in dataset 1, from 2014 to 2016. The solid line is 1:1 line.

(RMSE = 1.69 K, MEAE = 1.22 K).

Therefore, in the remainder of this study, Equation (2) is used to estimate $UHI_{max,}$ e.g.:

$$\begin{aligned}&\text{Fully-developed space: UHImax} = (2 - \text{SVF} - \text{Fveg} - \text{Fbuild})*\sqrt[4]{\frac{S \downarrow DTR^3}{U}}\\ &\text{Mixed space: UHImax}= (2 - \text{SVF} - \text{Fveg} + \text{Fbuild})*\sqrt[4]{\frac{S \downarrow DTR^3}{U}}\\ &\text{Normal space: UHImax}= (2 - \text{SVF} - \text{Fveg})*\sqrt[4]{\frac{S \downarrow DTR^3}{U}}\end{aligned} \quad (2)$$

4.2. Application for three spatial categories

From the above analysis, an extended equation with better modelled accuracy is confirmed. To apply it to more cities, a class prediction with different spatial categories will be discussed below.

The frequency of *Daily Ai*, which is the ratio of daily observed UHI_{max} to $\sqrt[4]{S \downarrow DTR^3/U}$ and represents numerical results for Ai (Fig. 10). According to Fig. 10 and considering the value of *Observe Ai* of the above stations in dataset 1(Table 6), the most reliable range of Ai seems to be 1–1.4, which indicates that the Normal space should meet the requirement of $0.6 < \text{SVF} + F_{veg} < 1$. Furthermore, we can infer that if:

(1) $\text{SVF} + F_{veg} \leq 0.6$, this station could be regarded as Fully−developed space.

(2) $\text{SVF} + F_{veg} \geq 1$, this station could be regarded as Mixed space.

(3) $0.6 < \text{SVF} + F_{veg} < 1$,this station could be regarded as Normal space. (5)

The "Fully-developed space" with massive construction and low vegetation fraction has an SVF less than 0.5 and F_{veg} less than 0.15 generally. Meanwhile, according to the open form around the station in "Mixed space", the value of SVF is usually more than 0.8 and F_{veg} not smaller than 0.1. Finally, the "Normal space" is relatively comfortable with appropriate density and vegetation, and exhibits a uniform and consistent spatial texture. Specifically, the value of SVF is generally from 0.3 to 0.8 and F_{veg} from 0.1 to 0.3.

The data from database 2 are used for independent verification of the class predication. The total of 20 stations with varied land use and morphology are evenly distributed in the main urban area of Xi'an (Fig. 1), which could basically represent the spatial characteristics of the city. According to Equation 5, 9 stations are classed as "Normal space", 9 as "Mixed space" and 2 as "Fully-developed space" (Table 7). The results of Equations (1) and (2) are compared in Fig. 11. The extreme UHI intensities (UHI_{max} less than 2–3 K or larger than 6.5–7 K),

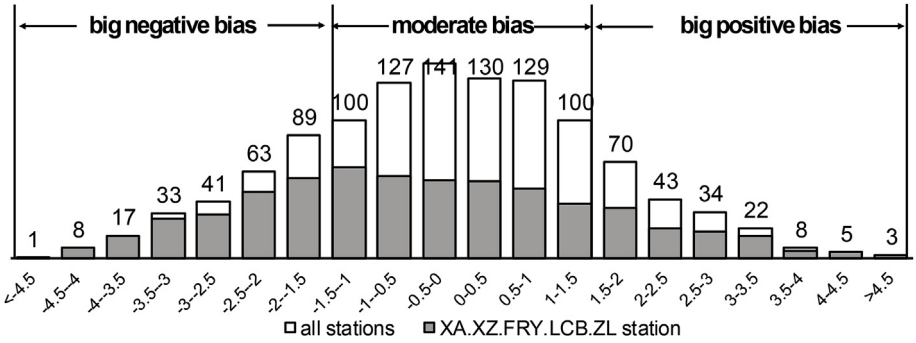

Fig. 6. Bias (=modelled UHI_{max} – observed UHI_{max}) in K for stations in dataset 1, from 2014 to 2016. Numbers denote the frequency of the occurrence of daily UHI_{max} bias in all stations and grey bars denote the frequency of the XA, XZ, FRY, LCB and ZL stations.

coincide in both models because the majority of them are "Normal space" and calculated with the same equation. While the medium UHI intensity (UHI_{max} from 3 to 6 K) are more composed by stations classed as "Mixed space" and "Fully-developed space", and the results of two equations are different. We find that more "Mixed space" in these sites contributes to the overall underestimated results in Equation (1) (RMSE = 1.26 K, MEAE = 1.01 K), and therefore its slope (observe UHI_{max} against the modelled UHI_{max}) is obviously lower than the 1:1 line. However, it has been successfully corrected in Equation (2) and the modelled UHI_{max} is closer consistent with the observed UHI_{max} (RMSE = 0.99 K, MEAE = 0.69 K), with similar performance as in Europe.

5. Discussion

In this paper we apply an equation to diagnose the UHI_{max}, based on routine meteorology data of a rural station and neighborhood-scale urban morphology for Xi'an city. After identifying the additional spatial categories outside the original range of calibration, we provide a modification to extend the European model and keep its accurate performance. Furthermore, the extended equation based on a class prediction with spatial categories could be applied for other cities.

Although many studies try to reveal the correlation between UHI intensity and urban morphology, they typically discuss the roles of different morphological factors separately [18,19]. Bottyan et al. (2010), B. Chun et al. (2014) have successfully utilized statistical models to relate comprehensive morphology with UHI intensity [25,26]. However, these models are derived for a single object and require retuning for application in other cities. In this study, we choose a diagnostic equation established and validated by a number of cities in northwestern Europe and furtherly extend its calibration range based on morphological characteristic in Xi'an China. Results show that the extended equation can be applied in varies urban spaces and it performs reasonably well.

It is important to note that the additional space categories in Xi'an did not just appear here but widely exist in developing cities in China. The "Fully-development space" typically represents the spatial form of newly emerging urban centers. The agglomeration of business offices and commercial facilities contributes to the gathering of crowd, which aggravates the severity of thermal uncomfortable caused by UHI effects. Meanwhile, according to the specific situation of each city, the "Mixed spaces" could be summarized as old city area and villages inside city. These spaces have similar layout of regionally high density and ventilation lacking, which leading to higher UHI intensity. Therefore, as the core area of city renewal, it is necessary to pay attention to these spaces and improve its thermal comfort.

Moreover, as mentioned in section 3.2, this study uses a comprehensive index Ai to represent the neighborhood-scale morphology of urban area and establish its internal relation with UHI_{max}. Although Ai is consisted of three key morphological indicators, SVF, F_{veg} and F_{build}, not all potentially important factors were taken in to account, considering the simplicity of data acquisition and calculation. Similar to our research, Lin et al. (2018), Ignatius et al. (2016) have adopted more variables as pavement area ratio, building height, areas of the building elevation, the surface albedo as predicting factors to downscale coarse resolutions to fine spatial and temporal resolutions required for building energy simulations [27,28]. However, the surface albedo is difficult to obtain in Xi'an, and the pavement ratio, building height and areas of the building elevation appeared to be poor governing parameters for our results.

Table 3
Large Bias (> 1.5K or < -1.5K) distribution with stations in database1.

BIAS (K)	XA	XZ	FRY	LCB	ZL	XQ	FQL	AZZX	XGX	Grand Total
< -4.5	0.0%	0.0%	0.0%	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
−4.5–4	0.0%	0.0%	0.0%	0.0%	7.3%	0.0%	0.0%	0.0%	0.0%	0.7%
-4–3.5	0.0%	0.0%	1.6%	0.7%	12.8%	0.0%	0.0%	0.0%	0.0%	1.5%
−3.5–3	0.0%	0.0%	2.3%	2.9%	20.2%	0.0%	2.3%	0.0%	0.8%	2.8%
-3–2.5	0.0%	0.0%	7.0%	4.4%	15.6%	0.0%	4.7%	0.7%	1.5%	3.5%
−2.5–2	0.0%	0.0%	21.1%	7.4%	10.1%	3.0%	7.0%	0.0%	1.5%	5.4%
-2–1.5	0.8%	1.5%	13.3%	15.4%	15.6%	4.4%	9.4%	5.2%	4.5%	7.6%
1.5–2	15.5%	12.2%	0.0%	0.7%	0.0%	7.4%	5.5%	8.9%	3.0%	6.0%
2–2.5	7.0%	9.2%	0.0%	0.7%	0.0%	5.9%	3.1%	3.7%	3.0%	3.7%
2.5–3	6.2%	9.2%	0.0%	0.0%	0.0%	3.7%	1.6%	0.0%	5.3%	2.9%
3–3.5	5.4%	7.6%	0.0%	0.0%	0.0%	0.7%	0.8%	2.2%	0.0%	1.9%
3.5–4	2.3%	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	0.0%	0.7%
4–4.5	1.6%	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%
> 4.5	1.6%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%
BIAS < −1.5	0.8%	1.5%	45.3%	31.6%	81.7%	7.4%	23.4%	5.9%	8.3%	21.6%
BIAS > 1.5	39.5%	43.5%	0.0%	1.5%	0.0%	17.8%	10.9%	16.3%	11.3%	15.9%
Grand Total	40.3%	45.0%	45.3%	33.1%	81.7%	25.2%	34.4%	22.2%	19.5%	37.5%

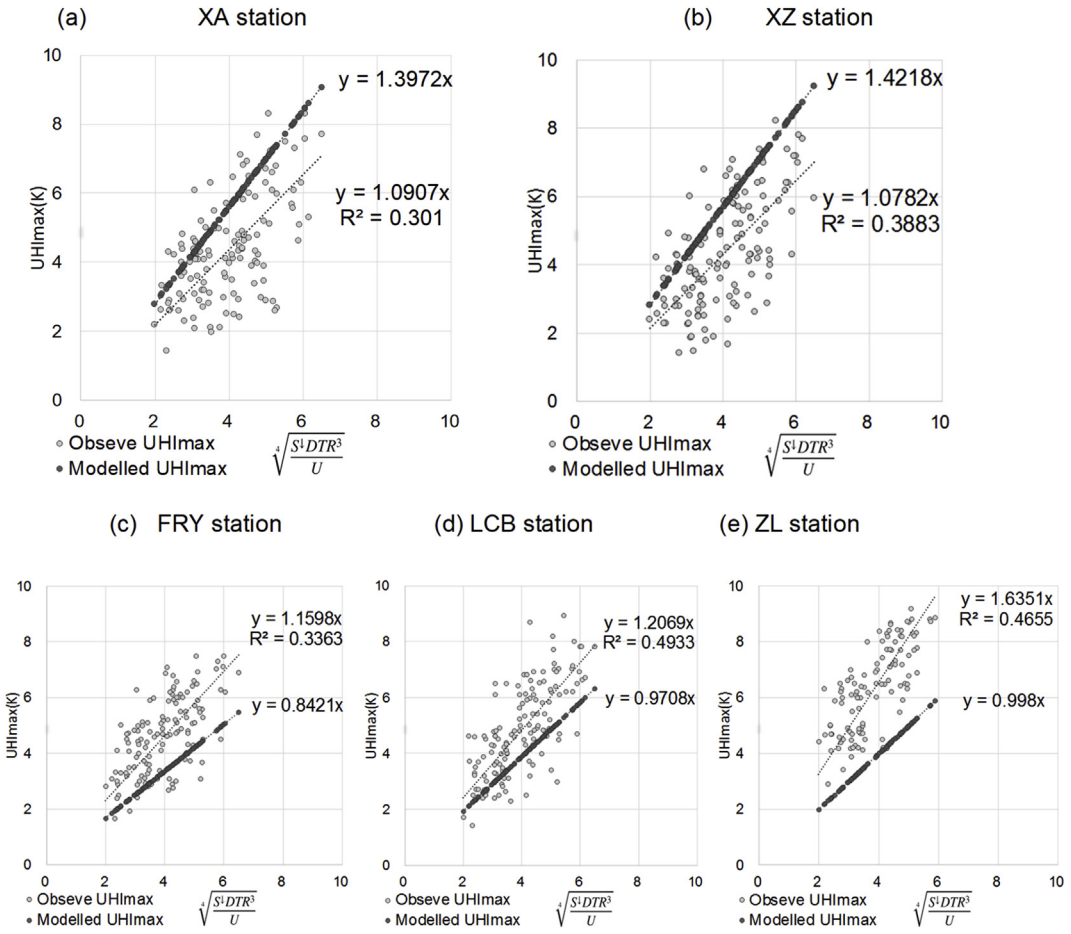

Fig. 7. The observed UHI_{max} (light grey dots) and modelled UHI_{max} (dark grey dots) against the meteorology part in Equation(1(a)): (a)the XA station, (b)the GXYZ station, (c)the FRY station, (d)the LCB station,(e)the ZL station. The light grey lines represent *Observed Ai* and its value are presented in the equations with R^2. The dark grey lines represent *Modelled Ai* and its value are presented in the equations without R^2.

Finally, all stations used for the original equation are located in Europe and in a mild maritime climate [15]. Hence, the range of weather types is large and year-round data are selected. However, Xi'an as the object in our study is located in transitional zone between temperate and subtropical monsoon climate zone, with relatively hot summers and cold winters and the UHI intensity varies greatly in different seasons [17]. Therefore, in our study, we only use the meteorology data of summer, which has the most prominent UHI effects. The applicability of the extended equation for other seasons in Xi'an could be considered as a next step. Meanwhile, we could also compare the performance of the extended equation for other urban areas, which have similar morphological characteristics but being located in different climate zones than Xi'an.

Table 5
Summary of modifications of *Modelled Ai*.

Adding F_{build}	**Equation 2** Fully-developed space: Ai = 2-SVF- F_{veg} - F_{build} Mixed space: Ai = 2-SVF- F_{veg} + F_{build} Normal space: Ai = 2-SVF- F_{veg}
Replacing SVF	**Equation 3** All space: Ai = 2-2* F_{veg} −2* F_{build}
Replacing F_{veg}	**Equation 4** Fully-developed and Normal space: Ai = 2-2*SVF + F_{build} Mixed space: Ai = 2-2*SVF+2* F_{build}

6. Conclusion

Based on the existing research methods of UHI intensity, this study proposes a simple and accurate equation to diagnose the UHI_{max} in

Table 4
The correlation coefficient of observe Ai and morphology indexes.

Morphology indexes	Correlation coefficient	Unit	Definition
F_{build}	0.67	m^2/m^2	The ratio of buildings' footprint area to the total plan area [22].
Population density	−0.47	persons/km^2	The ratio of population to sum area of this sub district.
complete aspect ratio	0.06	m^2/m^2	The ratio of the total 3-D external surface area of all elements including the ground surrounding them to the total plan area they occupy [23].
FAR	0.24	m^2/m^2	The ratio of gross floor area to the total plan area they occupy.
$4\sqrt{PQb}$	0.24	m^2/m^2	The ratio of the sum of building perimeters to the block circumference [24]
S/A	−0.30	m^2/m^2	The ratio of the total surface area of all buildings to the total plan area [24].
Fimper	0.14	m^2/m^2	The ratio of the sum imperious surface area to the total plan area.

Table 6
Summary of different Ai according to modifications.

Site	Type	Observe Ai	Equation 1 Ai	Equation 2 Ai	Equation 3 Ai	Equation 4 Ai
XA	Fully-developed	1.09	1.40	1.15	1.23	1.31
XZ	Fully-developed	1.08	1.42	1.13	1.14	1.41
FRY	Mixed	1.16	0.84	1.06	0.97	0.71
LCB	Mixed	1.21	0.97	1.21	1.07	0.87
ZL	Mixed	1.64	1.00	1.45	0.84	1.15
XQ	Normal	0.97	1.07	1.07	1.12	0.72
FQL	Normal	1.29	1.27	1.27	1.04	1.27
AZZX	Normal	1.19	1.28	1.28	1.10	1.14
XGX	Normal	1.17	1.25	1.25	1.23	1.11

Observed Ai is calculated by the linear fitting of observed data according to Equation (1a).
Equation1, 2, 3, 4 Ai is calculated by Equation(1b), (2), (3), (4).

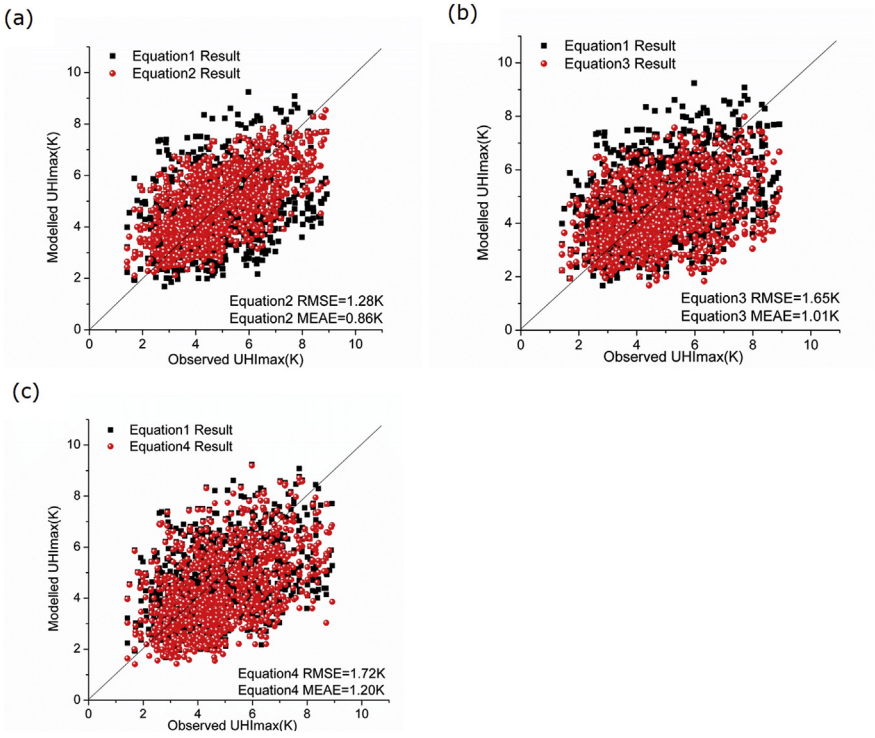

Fig. 8. The observed UHI_{max} against the results from three different equations: (a)Equation (2), (b)Equation(3) and (c) Equation(4). The black dots show the results of Equation(1) and red dots show the results of Equation (2) (3) (4). The solid line is the 1:1 line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

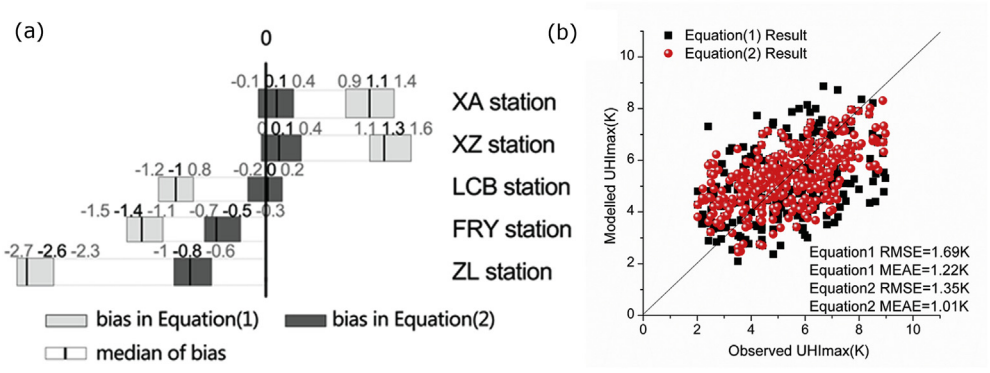

Fig. 9. (a) 95% confidence interval of bias of the XA, XZ, LCB, FRY and ZL stations. The dark grey bars stand for bias in Equation (1) and the light grey bars for bias in Equation (2). Fig. 9(b). The observed UHI_{max} against the predicted UHI_{max}. The black dots are the modelled UHI_{max} of Equation (1) and the red dots of Equation (2). The solid line is the 1:1 line. Summer climate data of 2017 (Jun, July, August) in dataset 1 are used for validation, and rainy (R_{sum} > 3) and fog (daily maximum relative humidity > 80%) events have been excluded. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Xi'an, which is a typical Chinese city with prominent UHI effects. Based on an evaluation of 3-year summer data, we find that the sum of morphological parameters of Xi'an are outside the original range of calibration, which leads to substantial large biases in the results. Thereto our optimization is aimed to find an extended equation that comprehensively expresses all spatial categories by adding a key morphological indicator F_{build} into it. The validation with independent data in 2017 supports the accurate performance of this new calibrated equation. Furtherly, a class prediction with three spatial categories are proposed based on control ranges of SVF + F_{veg}, which also validated by independent data of 20 stations in 2018.

To summarize, this study proposes an extended equation, of which the calibration is expanded from Europe to China. This equation could be used as an effective way to evaluate daily neighborhood-scale UHI

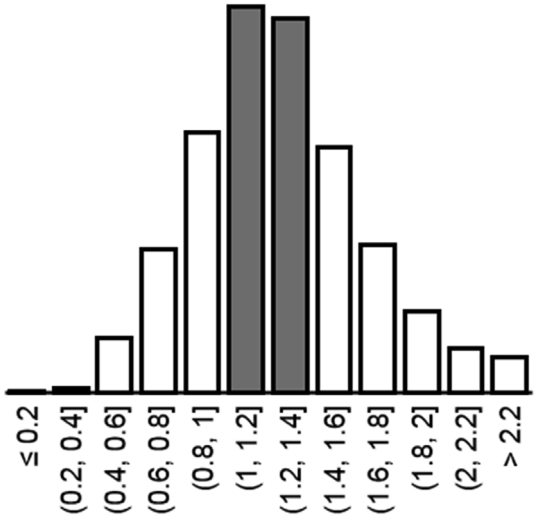

Fig. 10. The histogram of *Daily Ai* and dark grey bars present the most reliable range of Ai.

Table 7
Summary of spatial morphological indicators and different Ai for stations in database 2.

Name	TYPE	F_{veg} + SVF	Equation1 Ai	Equation 2 Ai
XA	Fully-developed	0.60	1.40	1.15
XZ	Fully-developed	0.58	1.42	1.13
FRY	Mixed	1.16	0.84	1.06
LCB	Mixed	1.03	0.97	1.21
SZF	Mixed	1.09	0.91	1.10
HZZX	Mixed	1.05	0.95	1.17
NM	Mixed	1.02	0.98	1.27
YXL	Mixed	1.02	0.98	1.09
DZKD	Mixed	1.00	1.00	1.25
SLP	Mixed	1.00	1.00	1.35
MTK	Mixed	1.00	1.00	1.16
XQ	Normal	0.93	1.07	1.07
FQL	Normal	0.73	1.27	1.27
AZZX	Normal	0.72	1.28	1.28
XGX	Normal	0.75	1.25	1.25
XASW	Normal	0.74	1.26	1.26
BHZX	Normal	0.65	1.35	1.35
SB	Normal	0.62	1.38	1.38
BG	Normal	0.61	1.39	1.39
BQZF	Normal	0.70	1.31	1.31

intensity in urban areas. Here, we also provide some suggestions about the wider use. Firstly, based on the evaluation results, a real-time dynamic map of UHI_{max} could be built to estimate the UHI intensity inside the city, which can provide reference for thermal environment monitoring. Also, considering the equation revealing the internal relation between urban morphology and UHI intensity, it can also provide suggestions on urban construction, regional renewal and future development characteristics by changing the morphological index of blocks.

Acknowledgements

This research was based on the project of "Comprehensive Survey of Human Settlement Quality and Preparation of Urban Climatic Maps Collection for Typical Cities" [grant IDs: 2013FY112500] and was financially supported by the Ministry of Science and Technology of the People's Republic of China. The authors gratefully acknowledge support from the The Netherlands Organisation for Scientific Research(NWO) grant 864.14.007 and the financial support from China Scholarship Council for Xi Zhang.

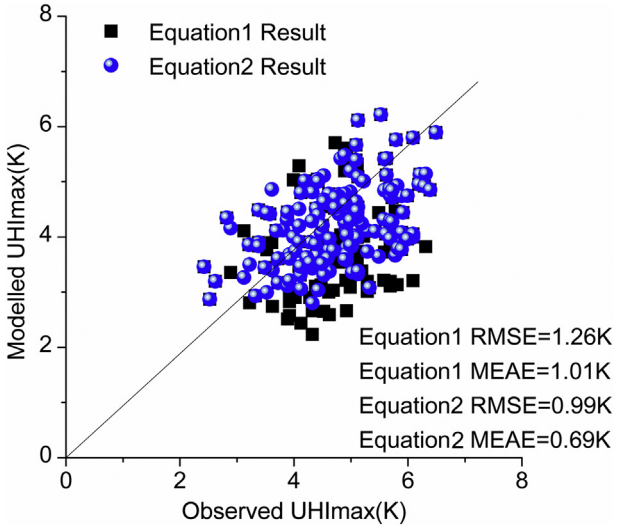

Fig. 11. The observe UHI_{max} against the modelled UHI_{max}. The black dots are the modelled UHI_{max} of Equation (1) and the blue dots of Equation (2). The solid line is the 1:1 line. Summer climate data of 2018 (16–25 July 2018) in dataset 2 are used for validation. Rainy (Rsum > 3) and fog (daily maximum relative humidity > 80%) events have been excluded.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.buildenv.2019.05.004.

References

[1] I. Khaliq, C. Hof, R. Prinzinger, K. Böhning-Gaese, M. Pfenninger, Global variation in thermal tolerances and vulnerability of endotherms to climate change, Proc. R. Soc. B 281 (2014) 20141097 1789 https://doi.org/10.1098/rspb.2014.1097.
[2] T.R. Oke, The energetic basis of the urban heat island, Q. J. R. Meteorol. Soc. 108 (455) (1982) 1–24 https://doi.org/10.1002/qj.49710845502.
[3] T.R. Oke, Street design and urban canopy layer climate, Energy Build. 11 (1–3) (1988) 103–113 https://doi.org/10.1016/0378-7788(88)90026-6.
[4] T.R. Oke, The heat island of the urban boundary layer: characteristics, causes and effects, Wind Climate in Cities, Springer, Dordrecht, 1995, pp. 81–107.
[5] T. Kjellstrom, I. Holmer, B. Lemke, Workplace heat stress, health and productivity–an increasing challenge for low and middle-income countries during climate change, Glob. Health Action 2 (1) (2009) 2047 https://doi.org/10.3402/gha.v2i0.2047.
[6] L. Martinelli, T.P. Lin, A. Matzarakis, Assessment of the influence of daily shadings pattern on human thermal comfort and attendance in Rome during summer period, Build. Environ. 92 (2015) 30–38 https://doi.org/10.1016/j.buildenv.2015.04.013.
[7] J. Tan, Y. Zheng, X. Tang, C. Guo, L. Li, G.,. Song, H. Chen, The urban heat island and its impact on heat waves and human health in Shanghai, Int. J. Biometeorol. 54 (1) (2010) 75–84.
[8] J. Yang, P. Yin, J. Sun, B. Wang, M. Zhou, M. Li, Q. Liu, Heatwave and mortality in 31 major Chinese cities: definition, vulnerability and implications, Sci. Total Environ. 649 (2019) 695–702 https://doi.org/10.1016/j.scitotenv.2018.08.332.
[9] P.A. Mirzaei, F. Haghighat, Approaches to study urban heat island–abilities and limitations, Build. Environ. 45 (10) (2010) 2192–2201 https://doi.org/10.1016/j.buildenv.2010.04.001.
[10] S. Koopmans, R. Ronda, G.J. Steeneveld, A. Holtslag, A. Klein Tank, Quantifying the effect of different urban planning strategies on heat stress for current and future climates in the agglomeration of the hague (The Netherlands), Atmosphere 9 (9) (2018) 353 https://doi.org/10.1002/joc.859.
[11] G.J. Steeneveld, S. Koopmans, B.G. Heusinkveld, L.W.A. Van Hove, A.A.M. Holtslag, Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in The Netherlands, J. Geophys. Res.: Atmosphere 116 (D20) (2011), https://doi.org/10.1029/2011JD015988.
[12] J.A. Voogt, T.R. Oke, Thermal remote sensing of urban climates, Remote Sens. Environ. 86 (3) (2003) 370–384 https://doi.org/10.1016/S0034-4257(03)00079-8.
[13] G.J. Steeneveld, S. Koopmans, B.G. Heusinkveld, N.E. Theeuwes, Refreshing the role of open water surfaces on mitigating the maximum urban heat island effect, Landsc. Urban Plann. 121 (2014) 92–96 https://doi.org/10.1016/j.landurbplan.2013.09.001.
[14] P.A. Mirzaei, Recent challenges in modeling of urban heat island, Sustain. Cities Soc. 19 (2015) 200–206 https://doi.org/10.1016/j.scs.2015.04.001.
[15] N.E. Theeuwes, G.J. Steeneveld, R.J. Ronda, A.A. Holtslag, A diagnostic equation for the daily maximum urban heat island effect for cities in northwestern Europe,

Int. J. Climatol. 37 (1) (2017) 443–454 https://doi.org/10.1002/joc.4717.
[16] X. Yang, L. Yao, C. Zhu, T. Jin, L.L. Peng, Analysis of local heat islands in nanjing, China, based on the local climate zone scheme, Procedia Eng. 205 (2017) 2501–2508 https://doi.org/10.1016/j.proeng.2017.09.980.
[17] M. Gao, H. Shen, X. Han, H. Li, L. Zhang, Multiple timescale analysis of the urban heat island effect based on the Community Land Model: a case study of the city of Xi'an, China, Environ. Monit. Assess. 190 (1) (2018) 8.
[18] S. Liu, Z. Zang, W. Wang, Y. Wu, Spatial-temporal evolution of urban heat Island in Xi'an from 2006 to 2016, Phys. Chem. Earth (2018) Parts A/B/C https://doi.org/10.1016/j.pce.2018.11.007.
[19] C. Qiuji, L. Chuting, Land surface temperature retrieval based on landsat ETM +/TM–Taking xi'an city as an example, Open Cybern. Syst. J. 9 (1) (2015).
[20] Xi'an Statistics Bureau, Xi'an Statistical Yearbook 2017, Xi'an Press, 2017.
[21] I.D. Stewart, T.R. Oke, Local climate zones for urban temperature studies, Bull. Am. Meteorol. Soc. 93 (12) (2012) 1879–1900 https://doi.org/10.1175/BAMS-D-11-00019.1.
[22] M.B. Pont, P. Haupt, Spacematrix: Space, Density, and Urban Form, NAi, 2010.
[23] T.R. Oke, Mills G. Christen, et al., Urban Climates, Cambridge University Press, 2017.
[24] J.Y. Yang, X. Sun, x Shi, Coupling Mechanism between Thermal Environment and Space Form and Optimization Design in City Center, Southeast University Press, 2017.
[25] Z. Bottyán, J. Unger, A multiple linear statistical model for estimating the mean maximum urban heat island, Theor. Appl. Climatol. 75 (3–4) (2003) 233–243 https://doi.org/10.1007/s00704-003-0735-7.
[26] B. Chun, J.M. Guldmann, Spatial statistical analysis and simulation of the urban heat island in high-density central cities, Landsc. Urban Plann. 125 (2014) 76–88 https://doi.org/10.1016/j.landurbplan.2014.01.016.
[27] F.Y. Lin, K.T. Huang, T.P. Lin, R.L. Hwang, Generating hourly local weather data with high spatially resolution and the applications in bioclimatic performance, Sci. Total Environ. 653 (2019) 1262–1271 https://doi.org/10.1016/j.scitotenv.2018.10.433.
[28] M. Ignatius, N.H. Wong, S.K. Jusuf, The significance of using local predicted temperature for cooling load simulation in the tropics, Energy Build. 118 (2016) 57–69 https://doi.org/10.1016/j.enbuild.2016.02.043.
[29] A. Matzarakis, F. Rutz, H. Mayer, Modelling Radiation fluxes in simple and complex environments – basics of the RayMan model, Int. J. Biometeorol. 54 (2010) 131–139.