Chin. Geogra. Sci. 2021 Vol. 31 No. 4 pp. 659−670
https://doi.org/10.1007/s11769-021-1215-7

Springer Science Press
www.springerlink.com/content/1002-0063

Predicting Surface Urban Heat Island in Meihekou City, China: A Combination Method of Monte Carlo and Random Forest

ZHANG Yao, LIU Jiafu, WEN Zhuyun

(*College of Tourism and Geographical Sciences*, *Jilin Normal University*, *Siping* 136000, *China*)

Abstract: Given the rapid urbanization worldwide, Urban Heat Island (UHI) effect has been a severe issue limiting urban sustainability in both large and small cities. In order to study the spatial pattern of Surface urban heat island (SUHI) in China's Meihekou City, a combination method of Monte Carlo and Random Forest Regression (MC-RFR) is developed to construct the relationship between landscape pattern indices and Land Surface Temperature (LST). In this method, Monte Carlo acceptance-rejection sampling was added to the bootstrap layer of RFR to ensure the sensitivity of RFR to outliners of SUHI effect. The SHUI in 2030 was predicted by using this MC-RFR and the modeled future landscape pattern by Cellular Automata and Markov combination model (CA-Markov). Results reveal that forestland can greatly alleviate the impact of SUHI effect, while reasonable construction of urban land can also slow down the rising trend of SUHI. MC-RFR performs better for characterizing the relationship between landscape pattern and LST than single RFR or Linear Regression model. By 2030, the overall SUHI effect of Meihekou will be greatly enhanced, and the center of urban development will gradually shift to the central and western regions of the city. We suggest that urban designer and managers should concentrate vegetation and disperse built-up land to weaken the SUHI in the construction of new urban areas for its sustainability.

Keywords: Monte Carlo and Random Forest Regression (MC-RFR); landscape pattern; surface heat island effect; Cellular Automata and Markov combination model (CA-Markov)

Citation: ZHANG Yao, LIU Jiafu, WEN Zhuyun, 2021. Predicting Surface Urban Heat Island in Meihekou City, China: A Combination Method of Monte Carlo and Random Forest. *Chinese Geographical Science*, 31(4): 659−670. https://doi.org/10.1007/s11769-021-1215-7

1 Introduction

With the rapid development of city, urbanization gradually brings various environmental problems, such as urban heat island (UHI), habitat loss, air pollution, and so on (Taha, 1997; WHO, 2010; Mao et al., 2018). UHI affects not only people's normal life, but also the sustainable development of city (Kalkstein and Greene, 1997; Patz et al., 2005). UHI enhancement is the result of changes in urban spatial extent including reduction of vegetation and water body and increase of impervious surface and the emission of artificial heat (Firozjaei et al., 2020). Therefore, studying future changes of UHI by means of urban spatial pattern is of great significance to urban management and planning, which needs extensive attentions.

Accompanied with development of thermal infrared remote sensing in recent years, more and more researchers focus on the Surface Urban Heat Island (SUHI). SUHI has become a research hotspot in the fields of urban climatology, urban ecology, urban planning and urban geography, which is highly related to land cover

Received date: 2020-12-04; accepted date: 2021-04-08
Foundation item: Under the auspices of National Natural Science Foundation of China (No. 41977411, 41771383), Technology Research Project of the Education Department of Jilin Province (No. JJKH20210445KJ)
Corresponding author: LIU Jiafu. E-mail: liujiafu@jlnu.edu.cn

or landscape pattern (Firozjaei et al., 2018). Many SUHI studies have shown that it is feasible to predict LST by urban landscape pattern (Weng, 2001; Voogt and Oke, 2003; Weng et al., 2004; Zhou et al., 2011; Connors et al., 2013; Maimaitiyiming et al., 2014; Zhou et al., 2016; Estoque et al., 2017). Through the analysis of landscape pattern, it is a hot topic to explore the influence of different types, proportions and shapes of landscape on LST. In the past studies, most researchers focused on the historical changes of SUHI and explored the relationships between SUHI and potential environmental factors. Only few researchers explored what will happen to SUHI in future. Therefore, this study analyzed the relationship between landscape pattern and SUHI by combining Monte-Carlo and Random Forest Regression, and predicted the future SUHI of Meihekou City, Northeast China.

Land cover type and other environmental parameters were often used to predict LST through linear prediction model in previous LST prediction. However, linear prediction model can not well fit the nonlinear data. Random Forest Regression (RFR) has been widely used in data mining due to its excellent performance in characterizing nonlinear correlation between different variables (Breiman et al., 1998). Despite the advantages of RFR, the sensitivity of RFR to outliers is low (Segal, 2004). The Monte Carlo acceptance-rejection sampling algorithm can increase the sampling probability of discrete values. Combining Monte Carlo and RFR can solve the problem of low sensitivity of RFR to outliers and ensure the prediction accuracy of discrete values in SUHI (Segal, 2004).

It is difficult to change the existing pattern of SUHI for cities with long development time or large scale. Therefore, from the perspective of long-term sustainable development, it is more important to make good planning for emerging small and medium-sized cities. Meihekou City locating in Jilin Province, Northeast China has developed rapidly in economy and urbanization in recent years. Due to its rapid development of economic level and urban scale, SUHI has brought many eco-environmental issues to residents of the city. These issues are also common in small and medium-sized cities in China and to the world. Therefore, the goals of this study are: 1) to construct the regression relationship between landscape pattern index and LST by a combination method of Monte Carlo and RFR in Meihekou City, 2) to examine LST historical changes and future pattern of LST, and 3) to investigate historical and future SUHI changes in the city. A new LST prediction framework is proposed in this study to explore the future pattern of SUHI in Meihekou City and thus this study provides a good reference for studying SUHI in other medium and small cities and is informative for urban managers or planners to considering SHUI in the construction of new urban land.

2 Materials and Methods

2.1 Study area

Meihekou City in Jilin Province of Northeast China occupies 2175 km^2, extending longitudes between 125°15′E and 126°03′E and latitudes between 42°08′N and 43°02′N, as shown in Fig. 1. Meihekou City is located in the transition zone between Songliao Plain and Changbai Mountainous area, with an average altitude of 340 m and a relatively flat terrain. The climate of the city is characterized by high temperature and rainy in summer and cold and dry in winter. The annual mean temperature of Meihekou City is relatively low. But since 1961, the annual mean temperature has been increasing continuously with changes from 4.8℃ to 5.9℃ (Ren and Wang, 2009).

2.2 Data source and processing

The Landsat Thematic Mapper (TM) and Operational Land Imager (OLI) images covering Meihekou City in 2010 and 2017 were obtained from the United States Geological Survey (USGS). Due to the obvious SUHI phenomenon in 2010 and 2017 and the excellent image quality, the images in July 2010 and 2017 were selected as the data source. Considering that the medium resolution satellite can provide a reasonable scale, and TM/OLI has more bands, which can better divide land use data and retrieve land surface temperature, so TM/OLI images is selected. A total of eight scenes of images were used for our study (Table 1). Since July is the hottest month in Meihekou City, and compared with other months, the SUHI phenomenon in July is more obvious and easy to observe. SUHI effect in our study refers only in July (Ren and Wang, 2009).

In order to predict SUHI in 2030, firstly, the land cover type maps of Meihekou City in 2010 and 2017 were obtained by Fractal Net Evolution Approach

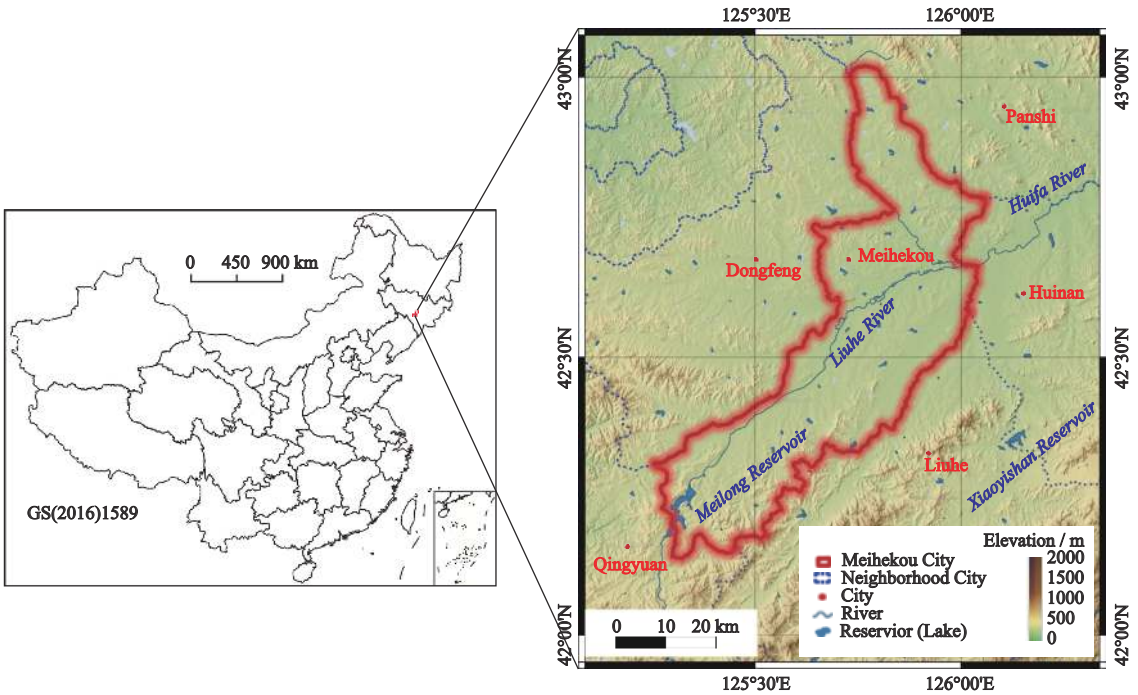

Fig. 1 Geographic location and terrain features of Meihekou City in Northeast China

Table 1 Landsat TM/OLI images used in the study in Meihekou City, Northeast China

Sensor model	ID	Acquisition date	Path	Row
TM	LT05_L1TP_117031_20100729_20161014_01_T1	2010-07-29	117	31
	LT05_L1TP_118031_20100704_20161016_01_T1	2010-07-04	118	31
	LT05_L1TP_117030_20100729_20161014_01_T1	2010-07-29	117	30
	LT05_L1TP_118030_20100704_20161014_01_T1	2010-07-04	118	30
OLI	LC08_L1TP_118030_20170723_20170729_01_T1	2017-07-23	118	30
	LC08_L1TP_118031_20170723_20170729_01_T1	2017-07-23	118	31
	LC08_L1TP_117030_20170716_20170727_01_T1	2017-07-16	117	30
	LC08_L1TP_118031_20170707_20170716_01_T1	2017-07-07	118	31

Notes: TM is Thematic Mapper, OLI is Operational Land Imager

(FNEA) and Classification And Regression Trees (CART). Based on these two datasets of land cover type, the land cover types in 2030 were predicted by Cellular Automata and Markov combination model (CA-Markov). The LST of 2010 and 2017 are obtained by the single window method. Secondly, the land cover type map and LST layer are gridded and divided into 2 km × 2 km grids, which can not only ensure the accuracy of temperature, but also ensure the number of landscape types in a grid. And the landscape pattern indices were calculated according to land cover types after gridding. Thirdly, the MC-RFR was used to get the prediction model. Finally, using the landscape pattern indices of Meihekou City in 2030 combined with the prediction model, the LST in 2030 was generated, and further analyzes were performed. Due to the difference between the maximum temperature and the minimum temperature in each cycle, to study SUHI accurately, the LST of each cycle is normalized. In order to facilitate the comparison between heat islands, they are artificially divided into five levels: strong heat island zone (0.8–1.0), heat island zone (0.6–0.8), normal zone (0.4–0.6), green island zone (0.2–0.4), strong green island zone (0–0.2).

2.2.1 Acquisition of land cover type datasets

FNEA is a multi-scale segmentation algorithm widely used at present (Baatz, 2000). It is also the basis and core content of popular object-oriented image analysis technology. FNEA algorithm uses fuzzy subset theory

to extract interested image objects. In this study, FNEA method in eCognition is used to segment remote sensing images for many times, and it is found that the segmentation effect is better when the scale parameter is given by 60, shape is 0.6, and compactness is 0.5. At the same time, this study selected more than 600 samples, set the threshold of each node in CART through visual interpretation, and obtained the final classification result through decision tree (. In order to ensure the accuracy of the classification results, 2500 sample points were selected and compared with the local topographic map of the current year. Finally, 1997 sample points were classified correctly, and the accuracy rate reached 79.88%, Kappa coefficient is 0.747, which confirmed that the classification accuracy was basically reliable.

2.2.2 Inversion of LST data by Single Window Method

The single window algorithm was proposed by Prof. Qin Zhihao in 2001 and has been continuously improved for LST retrieval of Landsat series satellites (Qin et al., 2001). This method avoids the dependence of real-time sounding data and only needs one thermal infrared channel. The traditional ground temperature data are mainly obtained through ground monitoring stations. Although the accuracy is high, it has some disadvantages such as high cost, long time-consuming and obtaining discrete point data. Therefore, only by using the results of discrete point data for spatial interpolation can we obtain the LST in a large continuous time range. The error between the obtained surface temperature and the actual value is large, and there is uncertainty. The single-window algorithm directly covers the influence of atmosphere and surface in the model equation. Compared with other radiative transfer models, the single window algorithm is more suitable for LST inversion from Landsat data. In this study, the atmospheric upward and downward brightness and atmospheric transmittance were obtained by NASA (http://atmcorr.gsfc.nasa.gov/). With the help of ENVI software and the parameters in Professor Qin's manuscript, the surface temperature was calculated.

2.2.3 Prediction of land cover types by CA-Markov Model in 2030

CA (Cellular Automata) and Markov are both dynamic models. Markov model is mainly used for numerical prediction, while CA model is mostly used to simulate the change of complex spatial system. Combining the advantages of both models, CA-Markov model can balance the historical trend of land pattern evolution and dynamically simulate the future land cover pattern.

CA can predict time and space simultaneously. The model has three discrete variables including state, domain space and time.

$$S_{(t+1)} = f(S_{(t)}, N) \quad (1)$$

where S is the set of cell states, t and $t+1$ represent two different times, N is the neighborhood of a cell; f is the cell transformation rule of local space.

Markov model can capture and describe the land cover change trend between two different time states, and set this change trend as the evolution rule of the next stage to predict the land cover pattern in the future. This rule is created by the conversion probability matrix of land cover change between two time points.

$$S'_{(t+1)} = P_{ij} S'_t \quad (2)$$

where S'_t, $S'_{(t+1)}$represents the state of land cover system at t and t+1 respectively, P_{ij} represents the state transition matrix, that is, the transition probability of landscape type i to landscape type j.

2.2.4 Landscape pattern indices

The sample set calculates the landscape pattern indices by using fragstats from three levels of landscape patter*n* patch, type, and landscape. The characteristics of landscape pattern in the study area were taken as regression samples. Referring to the relevant literature and the current situation of the study area (Weng, 2001; Weng et al., 2004; Xiao, 2007; Weng, 2009; Myint et al., 2010; Zhou et al., 2011; Connors et al., 2013; Maimaitiyiming et al., 2014; Zhou et al., 2016; Estoque et al., 2017), Shannon Evenness Index (SHEI) was selected from the landscape level to reflect the evenness of different types of landscape in the land landscape. In the type level, the Class Area (CA), Percentage of Landscape (PLAND), Largest Patch Index (LPI), Edge density (ED) and Landscape Division Index (Division) were selected. Detailed calculation methods refer to the official website of fragstats (http://www.umass.edu/landeco/research/fragstats/fragstats.html).

2.3 Monte Carlo and Random Forest Regression Combination

2.3.1 RFR algorithm

In the 1980s, Breiman and others invented the algorithm of classification tree, which used repeated bin-

ary data to classify or regress and greatly reduced the amount of calculation (Breiman et al., 1998). In 2001, Breiman combined the classification trees into RFR (Breiman, 2001), that is, the use of variables (columns) and the use of data (rows) were randomized to generate a number of classification trees, and then summarized into a comprehensive decision tree $\{h(x,\theta_k), t = 1, 2, 3, ...,N\}$. θ_k is independent random variable, x is input variable, and N is the number of regression decision trees. The mean value of K regression trees $h(x,\theta_k)$ is used as the result of regression prediction.

In the process of generating RFR, the probability that each sample will not be extracted is $p = (1-1/N)^N$, when N approaches infinity, p tends to 0.368. This shows that about 36.8% of the samples in the sample set N are not selected, which is called Out Of Bag data (OOB). Because the OOB data is unbiased, the accuracy of the model can be evaluated by calculating the OOB error. RFE model uses the mean square residuals of OOB data to evaluate the influence of independent variables on dependent variables, which is called the Variable Importance Measure (VIM) (Grömping, 2009).

2.3.2 Monte Carlo method

Monte Carlo method, also known as statistical simulation method, is widely used in physics, chemistry, economics and information technology fields through repeated random sampling to simulate the probability and statistics of objects. Acceptance-rejection sampling is a random sampling method for complex problems (Casella et al., 2004). First, an auxiliary recommendation distribution G is determined to generate candidate samples. Uniform distribution or normal distribution can be selected. Second, determine the uniform distribution of another auxiliary $U(0, 1)$. Third, set a constant c. c needs to satisfy the minimum value of $c \times g(x) \geq f(x)$, and the closer to 1, the better. Finally, if $U \leq \frac{f(Y)}{cg(Y)}$, let $X = Y$, cycle calculation is performed.

The sample set Y is obtained from the sample G of the proposed distribution, and the sample set u is obtained from the sample $U(0,1)$.

2.3.3 Monte Carlo-Random Forest Regression Combination Model

In the RFR regression algorithm, in order to solve the over fitting problem and improve the prediction accuracy, each decision tree is obtained by a bootstrap random sampling, and its node splitting feature is obtained by combining the training subset obtained from the sampling with the random subspace method (Segal, 2004; Grömping, 2009). Bootstrap random sampling is to sample all the samples as a whole, which makes the RFR extremely insensitive to outliers in the sample data. However, the objectives of SUHI study are extremely high value and very low value, which belong to abnormal value compared with the whole, so it is necessary to improve bootstrap in RFR. In this study, the acceptance rejection sampling method in Monte Carlo is introduced before bootstrap in RFR. Generate sample sets for all levels of data. Then, the RFR was used to calculate, in order to achieve better results. The flow chart is shown as Fig. 2. In order to test the stability of the model and the rationality of the evaluation results, MC-RFR model, RFR model and Linear Regression model were compared. The parameters of RFR are consistent with MC-RFR, i.e., k = 500. In order to compare the fitting ability and accuracy of the three models, the *RMSE* and R^2 values of each model were calculated. In order to avoid randomness, the three models were run for 50 times.

3 Results

3.1 Spatial and temporal patterns of land cover in 2010, 2017 and 2030

Among our tests, the classification achieved the best results when scale parameter is given by 60, shape is 0.6, and compactness is 0.5. The overall classification accuracy is 79.88%, and kappa coefficient is 0.747. Classification result has good consistency with actual samples and meet the accuracy requirements. According to Table 2 and Fig. 3, urban land in Meihekou City increased from 159 km^2 in 2010 to 168 km^2 in 2017 and to 178 km^2 in 2030. The area of the central city gradually expanded, and vertical expansion to the north and south. At the same time, towns in the northeast and southeast are expanding. Agricultural land in the central and eastern regions has expanded notably, with the total agricultural land area increasing from 1237 km^2 in 2010 to 1296 km^2 in 2030. The area of water body and grassland also increased slightly. The total woodland area decreased by about 100 km^2, of which the southern forestland area decreased markedly. Increase of grassland in the south can also be attributed to the degradation of forestland.

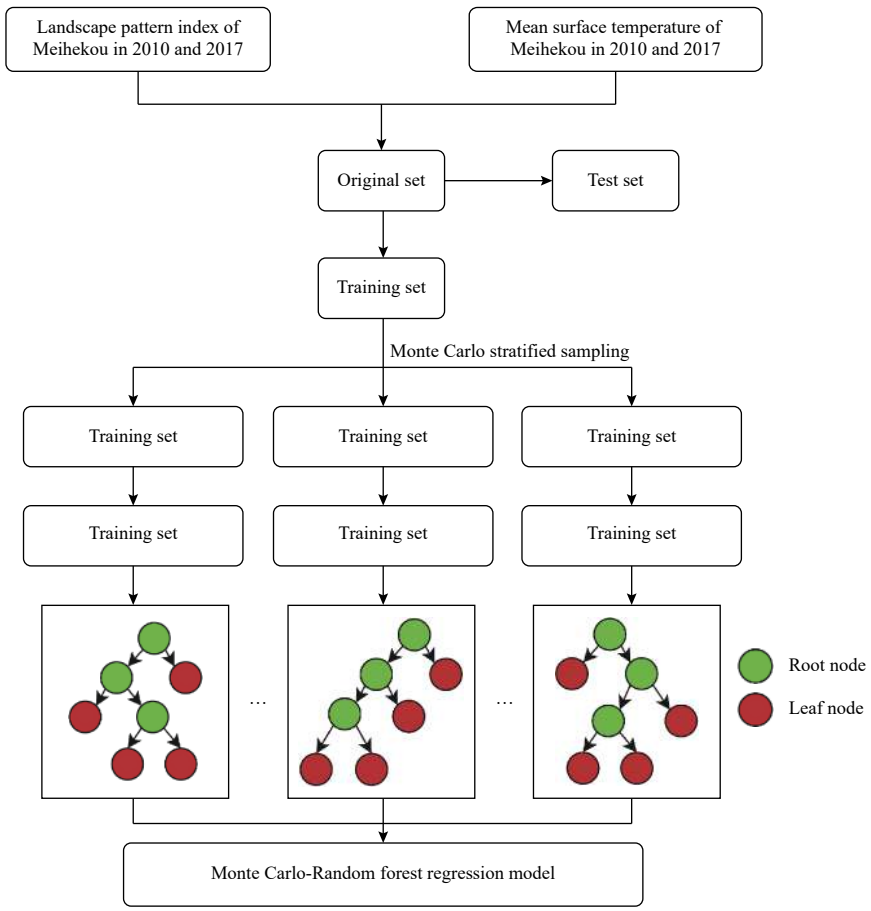

Fig. 2 Construction of the MC-RFR (Monte Carlo and Random Forest Gegression) Forecasting LST (Land Surface Temperature) Model

Table 2 Area of different land cover types in Meihekou City, Northeast China in different years / km^2

Land cover types	2010	2017	2030
Agricultural land	1236.78	1261.51	1326.41
Forestland	687.83	652.62	573.49
Grassland	9.08	9.10	9.46
Water body	84.79	85.70	89.45
Urban land	159.20	167.97	177.79
Unused land	3.92	4.68	5.00
Total	2181.59	2181.59	2181.59

3.2 Spatiotemporal patterns of LST in 2010 and 2017

As revealed in Fig. 4, ranges of high temperature in the urban center have an obvious expansion from 2010 to 2017. A large range of areas in the northeast region have changed from medium low temperature to medium high temperature. The temperature of a large number of low-temperature areas in the central and southern zone has increased and transformed into medium temperature areas. However, the low temperature areas in the south and north area increased gradually, which is supposed to be caused by the policy of returning agricultural land to forest. Overall, the surface temperature in 2017 was significantly higher than that in 2010.

3.3 Method accuracy comparison and prediction

Average values of the MC-RFR, RFR and Linear Regression indices were compared in Table 3. MC-RFR and RFR provide better fitting ability for the Training set compared with Linear Regression. The accuracy of MC-RFR is significantly higher than that of traditional RFR, indicating that MC-RFR is more suitable for studying the regression between landscape pattern and LST in Meihekou City.

In order to compare the performance of the three models more accurately, this study uses the cross validation method to extract 30 values from the sample sets as the test sets for accuracy simulation evaluation, and cir-

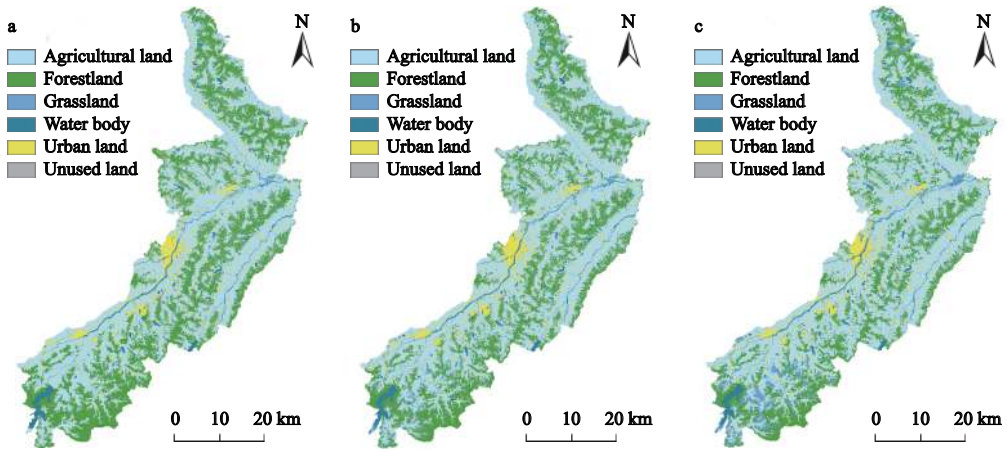

Fig. 3 Spatial patterns of land cover in Meihekou City, Northeast China. (a) 2010, (b) 2017, (c) 2030

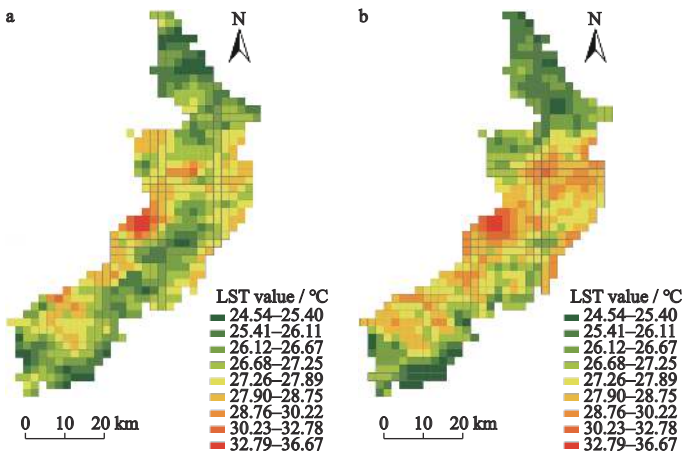

Fig. 4 Spatial patterns of LST in Meihekou City, Northeast China. (a) 2010 and (b) 2017

Table 3 Comparison of performance indices of different models

Model	Monte Carlo-Random Forest Regression		Random Forest Regression		Linear regression	
	Training set	Test set	Training set	Test set	Training set	Test set
R^2	0.831	0.822	0.796	0.818	0.537	0.603
RMSE	0.131	0.126	0.252	0.193	0.474	0.531

cularly extract 50 times. Fig. 5 is the result of one of them. It can be seen that RFR has a large deviation at extremely high value and extremely low value, and the prediction result is close to the real result only when it is located in the average value area. Although the MC-RFR has a deviation in the extremely high value region and the extremely low value region, the prediction result is relatively close. The difference between the two models is 30%. The results of Linear Regression are not ideal. The precision of MC-RFR and RFR is about 0.78 and 0.54, and the precision of Linear Regression is only 0.32. In addition, the prediction model is proved to be significant and credible. However, it can be seen from Fig. 6 that the MC-RFR still has some shortcomings in dealing with outliers, which needs to be improved in the follow-up study.

Compared Fig. 6a with Fig. 4, the temperature in the center study area has been significantly increased and the scope has gradually expanded in 2030. A number of medium and high-temperature regions have been transformed into high-temperature regions and connected with each other by 2030. There are new high temperature areas in the north and south of the city center, which can be inferred that the expansion of the city is in the

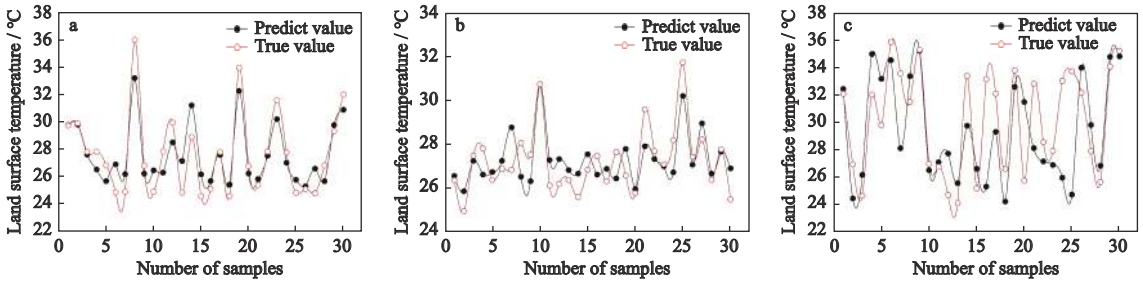

Fig. 5 Verification results of the landscape pattern and heat island effect test sets. (a) Monte Carlo-Random Forest Regression, (b) Random Forest Regression, (c) Linear regression

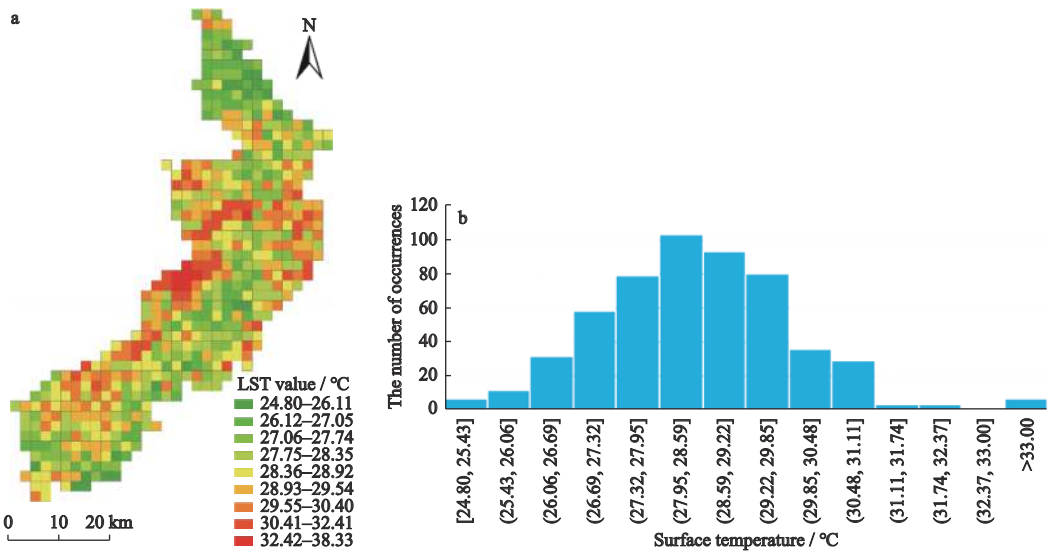

Fig. 6 Spatial pattern of LST in Meihekou City, Northeast China in 2030. (a) Spatial variation, (b) Frequency

north-south direction. From the Fig. 6b, it can be found that most of the current temperature distribution range is between 26.69℃–29.85℃, but the frequency of 28.59℃–29.85℃ is greater than 26.06℃–27.32℃, indicating that the middle high temperature region is far larger than the middle low temperature region.

3.4 SUHI Effect based on LST

The strength of SUHI is actually the distance between LST and average LST. Because the distance between LST and average LST is related to time scale, according to a large number of remote sensing experimental analysis and related research, the heat island can be divided into five grades: strong green island, green island, middle, heat island and strong heat island (Xiao, 2007; Sun et al., 2012; Song et al., 2014). It can be seen from Fig. 7 that the strong heat island zone in Meihekou City changed little from 2010 to 2017, but the area change showed an upward trend from 2017 to 2030, with the area changing from 11.68 km^2 to 42.76 km^2. From 2010 to 2030, the heat island zone increased from 29.42 km^2 in 2010 to 44.85 km^2 in 2030. The normal zone showed a trend of decline first and then rise. The green island zone decreased by half from 1516.03 km^2 in 2010 to 782.20 km^2 in 2030. The strong green island zone increased steadily year by year, from 570.07 m^2 in 2010 to 938.04 km^2 in 2017 and then to 1195.87 km^2. In 2017, a strong heat island point appeared in the north central region, but disappeared in 2030, followed by a larger range of median islands. In 2010 and 2017, the normal zone of the central and western regions will be expanded by 2030. The southwest strong green island zone is also increasing year by year.

4 Discussion

4.1 Relationship between landscape pattern and land surface temperature

Land cover changes usually caused significant ecologic-

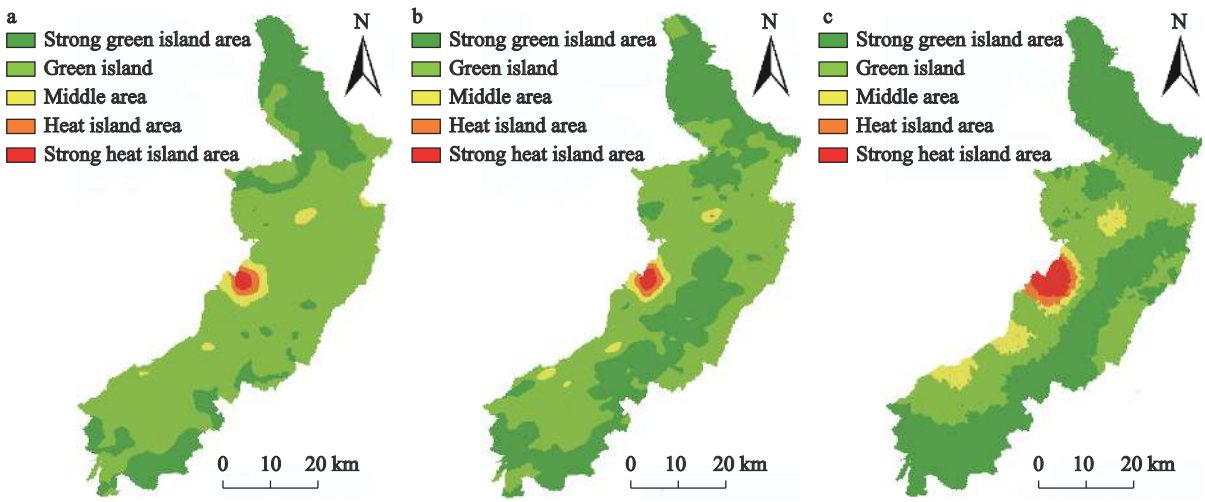

Fig. 7 Spatial and temporal distribution pattern of SUHI in Meihekou City, Northeast China, (a) 2010, (b) 2017, and (c) 2030

al changes (Mao et al., 2019). Landscape types contribute more to the study of the relationship between landscape pattern and LST than landscape configuration. The main reason for this situation is that each landscape type has a different impact on LST. For example, many studies have shown that the LST of greenbelt and forestland is lower than that of impervious surface (Weng et al., 2004; Sun et al., 2012; Song et al., 2014). Peng et al. found in the study of Beijing metropolitan area in 2016 that when forestland or grassland occupied more than 70% of the whole patch area, its LST decreased significantly (Peng et al., 2016). At the same time, Estoque et al. (2017) also proposed that the area of forestland and grassland in the whole patch should be increased if the SUHI effect is to be alleviated. However, both large cities and small and medium-sized cities in this study should focus on improving the single landscape type. For example, Du et al. analyzed the impact of the preparation of isolation zone in Zhuhai City on LST in 2016, and concluded that landscape configuration is also very important for LST (Du et al., 2014). Therefore, there is a great heterogeneity between landscape pattern and LST (Cheng et al., 2015). Generally speaking, the increase of fragmentation degree of strong impact sources (such as impervious surface) is beneficial to alleviate SUHI effect. However, the impact of weak sources (forestland, grassland, water body, *etc*.) on SUHI is mainly manifested in that the LST decreases with the increase of the percentage of patch area. With the increase of the degree of fragmentation of strong impact sources, it is bound to lead to the participation of other sources, which will lead to the decrease of LST (Zhou et al., 2011).

4.2 Significance of Monte Carlo and random forest combination model

Choosing appropriate methods or models is the key to study the relationship between landscape pattern indices and LST. Traditional methods such as Linear Regression or Artificial Neural Network (ANN) are often used to study landscape pattern and LST, but the spatial autocorrelation caused by spatial independence of dependent variables can not be considered in traditional Linear Regression, which leads to large model error (Segal, 2004; Grömping, 2009). At the same time, in the study of heat island effect, although the effect of ANN is better, its implementation is troublesome, which is not conducive to the operation in the real environment (Ashtiani et al., 2014). In this study, Monte Carlo and random forest regression are combined and compared with random forest regression and Linear Regression. We found that the prediction accuracy of random forest regression is significantly improved compared with Linear Regression, but the extreme value and extremely low value of random forest are not even as good as Linear Regression because of its insensitivity to outliers. In order to solve this problem, Monte Carlo and random forest regression are combined to solve the problem of insensitive outliers. Therefore, the MC-RFR combination method is a promising method to solve the relationship between landscape pattern indices and LST, which is of great significance to the study of SUHI.

4.3 Surface urban heat island mitigation strategy and adaptation

According to the observation, the area of urban land-based strong heat island in the middle of Meihekou City has increased. The area is relatively dense buildings, less green space and water, resulting in higher temperature than other areas. The normal zone in the central region also increased, indicating that villages and towns are gradually developing. However, due to the sparse buildings in villages and towns and the low reflectivity of roads, it presents the characteristics of intermediate value, but it is still higher than that of green space and agricultural land. The strong green island zone in the north and South increased significantly, and the vegetation coverage was much higher than that in the middle of Meihekou City, because the strong green island zone increased with the increase of agricultural land and forestland (Zhou et al., 2011; Zhou et al., 2016). However, it is worth noting that the heat island zone is still expanding, and the reasonable layout of new urban land should be strengthened.

At present, many scholars interpret SUHI from different perspectives and start to study how to deal with and improve SUHI. For example, the ability of vegetation covered roofs, roofs of special materials and roads in adapting to and mitigating SUHI, and the influence of urban three-dimensional structure on air temperature are also the current research directions (Ferguson et al., 2008; Hoverter, 2012; Rajagopalan et al., 2014; Hoag, 2015; Qin, 2015; Sharma et al., 2016). At the same time, there is another direction, that is, the direction of urban greening, to increase and improve the area and structure of green space and water body in the city. From the study results, reducing the proportion of construction land and unused land in the whole patch and aggregation degree is the fundamental way to solve SUHI. However, in the process of urban development, the reduction of urban construction land is not in line with the actual development needs, so we need to start from the landscape configuration, increase the fragmentation degree of urban land in patches, and increase the interval of forestland, grassland or water body to reduce SUHI. In addition, adjusting the spatial distribution of strong heat sources is also a strategy. In the future urban planning, some landscapes can be added to absorb heat and facilitate heat dissipation, including forestland and water body. At the same time, in order to obtain rich landscape diversity, multiple buffer zones can be put in urban land to reduce SUHI brought by large-scale built-up areas. In addition, saving and intensive use of construction land, tapping the inherent potential of construction land, and reasonably adjusting the urban spatial form are of great significance to alleviate the SUHI effect (Weng, 2009; Myint et al., 2010; Estoque et al., 2017).

In the process of urbanization, different landscape types are inevitably disturbed by human beings. Human activities will lead to changes in the distribution and area of all kinds of landscapes, and then affect the size and strength of heat island. If according to a certain development strategy, reasonable planning of the distribution of all kinds of landscape in the city can give full play to the various characteristics of the landscape, and effectively slow down the phenomenon of heat island. But landscape type is not enough. Humidity, wind speed, building area, vegetation cover types, and even building height and material are the factors affecting SUHI effect (Hoverter, 2012). In this study, the landscape level is simply studied. The influencing factors and the landscape configuration combination should be refined and improved in the follow-up study (Qin, 2015; Sharma et al., 2016). Only in this way can we put forward more reasonable guidance for urban planning, which is of great significance to solve the SUHI effect.

5 Conclusions

In this study, a combination method of Monte Carlo and Random Forest Regression (MC-RFR) was used to evaluate SUHI in Meihekou City in Northeast China. Based on the land cover data in 2010 and 2017, land cover types in 2030 are predicted by using CA-Markov algorithm, and the future SUHI effect was further evaluated. The prediction results show that urban development center of Meihekou City will shift from the central urban area to the central and western regions in 2030. The strong heat island points in the northern part of 2017 will change to a larger normal zone in 2030, indicating that the development process will change from centralized development to surrounding development. The green island zone is increasing year by year, which indicates that the protection of agricultural land and forestland is increasing. We believe that urban managers should reduce the impact of SHUI by concentrating vegetation and dispersing the built urban land (such

as buildings and roads) in the construction of new urban land, so as to provide a more livable urban living environment. The study could provide a reference to evaluate SUHI in other cities.

References

Ashtiani A, Mirzaei P A, Haghighat F, 2014. Indoor thermal condition in urban heat island: comparison of the artificial neural network and regression methods prediction. *Energy and Buildings*, 76: 597–604. doi: 10.1016/j.enbuild.2014.03.018

Baatz M, Schape A, 2000. Multi-resolution segmentation: an optimization approach for high quality multi-scale image segmentation. Heidelberg: Angewandte Geographische Informationsverarbeitung.

Breiman L, 1998. Arcing classifier (with discussion and a rejoinder by the author). *The Annals of Statistics*, 26(3): 801–849. doi: 10.1214/aos/1024691079

Breiman L, 2001. Random forests. *Machine Learning*, 45(1): 5–32. doi: 10.1023/A:1010933404324

Casella G, Robert C P, Wells M T, 2004. Generalized accept-reject sampling schemes: a festschrift for Herman Rubin. *Institute of Mathematical Statistics*, 342–347. doi: 10.1214/lnms/1196285403

Cheng X Y, Wei B S, Chen G J et al., 2015. Influence of park size and its surrounding urban landscape patterns on the park cooling effect. *Journal of Urban Planning and Development*, 141(3): A4014002. doi: 10.1061/(ASCE)UP.1943-5444.0000256

Connors J P, Galletti C S, Chow W T L, 2013. Landscape configuration and urban heat island effects: assessing the relationship between landscape characteristics and land surface temperature in Phoenix, Arizona. *Landscape Ecology*, 28(2): 271–283. doi: 10.1007/s10980-012-9833-1

Du D Y, Li A H, Zhang L L, 2014. Survey on the applications of big data in Chinese real estate enterprise. *Procedia Computer Science*, 30: 24–33. doi: 10.1016/j.procs.2014.05.377

Estoque R C, Murayama Y, Myint S W, 2017. Effects of landscape composition and pattern on land surface temperature: an urban heat island study in the megacities of Southeast Asia. *Science of the Total Environment*, 577: 349–359. doi: 10.1016/j.scitotenv.2016.10.195

Ferguson B, Fisher K, Golden J et al., 2008. *Reducing Urban Heat Islands: Compendium of Strategies-cool Pavements*. Washington, DC, United States: Environmental Protection Agency.

Firozjaei M K, Fathololoumi S, Kiavarz M et al., 2020. Modelling surface heat island intensity according to differences of biophysical characteristics: a case study of Amol City. *Ecological Indicators*, 109: 105816. doi: 10.1016/j.ecolind.2019.105816

Firozjaei M K, Kiavarz M, Alavipanah S K et al., 2018. Monitoring and forecasting heat island intensity through multi-temporal image analysis and cellular automata-Markov chain modelling: a case of Babol city. *Ecological Indicators*, 91: 155–170. doi: 10.1016/j.ecolind.2018.03.052

Grömping U, 2009. Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, 63(4): 308–319. doi: 10.1198/tast.2009.08199

Hoag H, 2015. How cities can beat the heat: rising temperatures are threatening urban areas, but efforts to cool them may not work as planned. *Nature*, 524(7566): 402–405. doi: 10.1038/524402a

Hoverter S P, 2012. Adapting to urban heat: a tool kit for local governments. Georgetown Climate Center, August 2012.

Kalkstein L S, Greene J S, 1997. An evaluation of climate/mortality relationships in large US cities and the possible impacts of a climate change. *Environmental Health Perspectives*, 105(1): 84–93. doi: 10.1289/ehp.9710584

Maimaitiyiming M, Ghulam A, Tiyip T et al., 2014. Effects of green space spatial pattern on land surface temperature: implications for sustainable urban planning and climate change adaptation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 89: 59–66. doi: 10.1016/j.isprsjprs.2013.12.010

Mao D H, Wang Z M, Wu J G et al., 2018. China's wetlands loss to urban expansion. *Land Degradation & Development*, 29(8): 2644–2657. doi: 10.1002/ldr.2939

Mao D H, He X Y, Wang Z M et al., 2019. Diverse policies leading to contrasting impacts on land cover and ecosystem services in Northeast China. *Journal of Cleaner Production*, 240: 117961. doi: 10.1016/j.jclepro.2019.117961

Myint S W, Brazel A, Okin G et al., 2010. Combined effects of impervious surface and vegetation cover on air temperature variations in a rapidly expanding desert city. *GIScience & Remote Sensing*, 47(3): 301–320. doi: 10.2747/1548-1603.47.3.301

Patz J A, Campbell-Lendrum D, Holloway T et al., 2005. Impact of regional climate change on human health. *Nature*, 438(7066): 310–317. doi: 10.1038/nature04188

Peng J, Xie P, Liu Y X et al., 2016. Urban thermal environment dynamics and associated landscape pattern factors: a case study in the Beijing metropolitan region. *Remote Sensing of Environment*, 173: 145–155. doi: 10.1016/j.rse.2015.11.027

Qin Y H, 2015. A review on the development of cool pavements to mitigate urban heat island effect. *Renewable and Sustainable Energy Reviews*, 52: 445–459. doi: 10.1016/j.rser.2015.07.177

Qin Z H, Karnieli A, Berliner P A, 2001. Mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *International Journal of Remote Sensing*, 22(18): 3719–3746. (in Chinese). doi: 10.1080/01431160010006971

Rajagopalan P, Lim K C, Jamei E, 2014. Urban heat island and

wind flow characteristics of a tropical city. *Solar Energy*, 107: 159–170. doi: 10.1016/j.solener.2014.05.042

Ren Limin, Wang Gang, Guo Lihong, 2009. Discussion on climate change in Meihekou city in recent 50 years and its influence on agricultural production. Changchun, Jilin, China: Collection of papers on climate change of the 26th annual meeting of China Meteorological Society. (in Chinese)

Segal M R, 2004. Machine learning benchmarks and random forest regression. Center for Bioinformatics and Molecular Biostatistics. April 14, 2003.

Sharma A, Conry P, Fernando H J S et al., 2016. Green and cool roofs to mitigate urban heat island effects in the Chicago metropolitan area: evaluation with a regional climate model. *Environmental Research Letters*, 11(6): 064004. doi: 10.1088/1748-9326/11/6/064004

Song J, Du S H, Feng X et al., 2014. The relationships between landscape compositions and land surface temperature: quantifying their resolution sensitivity with spatial regression models. *Landscape and Urban Planning*, 123: 145–157. doi: 10.1016/j.landurbplan.2013.11.014

Sun Q Q, Wu Z F, Tan J J, 2012. The relationship between land surface temperature and land use/land cover in Guangzhou, China. *Environmental Earth Sciences*, 65(6): 1687–1694. doi: 10.1007/s12665-011-1145-2

Taha H, 1997. Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. *Energy and Buildings*, 25(2): 99–103. doi: 10.1016/S0378-7788(96)00999-1

Voogt J A, Oke T R, 2003. Thermal remote sensing of urban climates. *Remote Sensing of Environment*, 86(3): 370–384. doi: 10.1016/S0034-4257(03)00079-8

Weng Q H, Lu D S, Schubring J, 2004. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment*, 89(4): 467–483. doi: 10.1016/j.rse.2003.11.005

Weng Q, 2001. A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. *International Journal of Remote Sensing*, 22(10): 1999–2014. doi: 10.1080/713860788

Weng Q H, 2009. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4): 335–344. doi: 10.1016/j.isprsjprs.2009.03.007

WHO, 2010. The world health report: health systems financing: the path to universal coverage: executive summary. World Health Organization, 2010.

Xiao R B, Ouyang Z Y, Zheng H et al., 2007. Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences*, 19(2): 250–256. doi: 10.1016/S1001-0742(07)60041-2

Zhou D C, Zhang L X, Hao L et al., 2016. Spatiotemporal trends of urban heat island effect along the urban development intensity gradient in China. *Science of the Total Environment*, 544: 617–626. doi: 10.1016/j.scitotenv.2015.11.168

Zhou W Q, Huang G L, Cadenasso M L, 2011. Does spatial configuration matter: understanding the effects of land cover pattern on land surface temperature in urban landscapes. *Landscape and Urban Planning*, 102(1): 54–63. doi: 10.1016/j.landurbplan.2011.03.009