Simulating the Relationship between Land Use/Cover Change and Urban Thermal Environment Using Machine Learning Algorithms in Wuhan City, China

Maomao Zhang 1,†, Cheng Zhang 2,†, Abdulla-Al Kafy 3,4 and Shukui Tan 1,*

1 College of Public Administration, Huazhong University of Science and Technology, Wuhan 430079, China; star_mzhang@mails.ccnu.edu.cn
2 College of Mechanical and Electrical Engineering, Hohai University, Changzhou 213022, China; wennie@mails.ccnu.edu.cn
3 Department of Urban & Regional Planning, Rajshahi University of Engineering & Technology, Rajshahi 6203, Bangladesh; abdulla-al.kafy@localpathways.org
4 ICLEI South Asia, Rajshahi City Corporation, Rajshahi 6200, Bangladesh
* Correspondence: tansk@hust.edu.cn
† These authors contributed equally to this work.

Abstract: The changes of land use/land cover (LULC) are important factor affecting the intensity of the urban heat island (UHI) effect. Based on Landsat image data of Wuhan, this paper uses cellular automata (CA) and artificial neural network (ANN) to predict future changes in LULC and LST. The results show that the built-up area of Wuhan has expanded, reaching 511.51 and 545.28 km², while the area of vegetation, water bodies and bare land will decrease to varying degrees in 2030 and 2040. If the built-up area continues to expand rapidly, the proportion of 30~35 °C will rise to 52.925% and 55.219%, and the affected area with the temperature >35 °C will expand to 15.264 and 33.612 km², respectively. The direction of the expansion range of the LST temperature range is obviously similar to the expansion of the built-up area. In order to control and alleviate UHI, the rapid expansion of impervious layers (built-up areas) should be avoided to the greatest extent, and the city’s “green development” strategy should be implemented.

Keywords: land use/land cover changes; urban thermal environment; machine learning algorithms; artificial neural network; Wuhan

1. Introduction

China is going through a stage of rapid urbanization, the proportion of the urban economy in the national economy is constantly increasing [1], and the resource and environmental problems facing cities are becoming increasingly urgent [2,3]. In 2030, China’s urbanization rate will reach 70%, and the percentage of urban population is expected to be as high as 60% [4,5]. The rapid development of urbanization and the rapid expansion of urban population have caused impervious surfaces represented by reinforced concrete to gradually occupy urban space [6]. These building materials have small heat capacity, fast heat absorption, and strong temperature storage, and are the main heat source for the rise of ground temperature [7–9]. The urban heat island (UHI) effect is a product of urbanization [10,11], and is the most prominent feature of the impact of human activities on temperature [12,13]. UHI is one of the most significant features of urban climate, and it is also one of the environmental problems that are ubiquitous in urban ecosystems and attract more and more attention [14,15]. Land use/land cover (LULC) change analysis is an important technology to measure environmental sustainability and ecological environment quality [16,17]. The research on the relationship between LULC and urban surface temperature cannot only analyze the spatial characteristics of urban thermal environment under LULC from the origin of heat island phenomenon, but also provide theoretical basis for...
rational use of land resources, rational planning of urban construction, and weakening of urban heat island effect [18,19].

Since Lake Howard conducted a study of “urban heat island” in the 19th century, taking the temperature and suburbs of central London as an example, people have begun to pay attention to the changes in urban surface temperature [20]. Scholars have carried out a lot of research on the relationship between urban development and thermal environment, and most of them use land surface temperature (LST) as a quantitative indicator of thermal environment [21]. The research content mostly focuses on the following aspects: using impervious surface to characterize urban land cover, and analyzing the impact of land change caused by urban development on LST [22]. There are many LULC indicators that predict changes in LST in a region. Indicators such as the normalized difference vegetation index (NDVI) are weaker predictors of LST, while the normalized difference building index (NDBI) and the normalized difference bare soil index (NDBSI) are considered more reliable by a large number of studies [23,24]. Some scholars have established models of the relationship between Impervious Surface Percentage (ISP) and the index of impervious surface landscape pattern and LST to explore the response law of LST spatial distribution to impervious surface changes [25]. There are also scholars discussing the relationship between human activities and other social and economic factors of urban development and LST. Their research mainly analyzed the influence mechanism of population density, economic development status, intensity and type of human activities on LST [26,27].

With the continuous development of remote sensing technology, it has provided effective help for the inversion of surface temperature, and has accelerated the pace of research on urban heat island effects. The application of remote sensing (RS) and geographic information system (GIS) provides important technical support for estimating urban LULC changes and LST distribution [28]. Landsat image data provide the possibility to describe the changes of LULC and its impact on LST. A lot of research is carried out around LULC and LST prediction, and the choice of prediction method will directly affect the accuracy of prediction [29]. Currently, Markov chain (MC), cellular automata model (CA), logistic regression (LR) and other methods are used in the prediction of changes in LULC and LST [19,30,31]. Artificial neural network (ANN) is an algorithmic mathematical model that is similar to the structure of the brain’s synaptic connections for distributed and parallel information processing, and has the ability of self-learning and self-adaptation [32]. One of the most important advantages of ANN is that it can model complex nonlinear relationships, it has good advantages for the prediction and simulation of LULC and LST [33,34].

Although the current research has achieved many results, there are still the following problems: (1) The research method used is relatively simple, and the relevant analysis methods are not screened and selected, and the decision-making of the analysis model needs to be improved. The research methods are mostly based on correlation analysis and ordinary linear regression analysis, ignoring the spatial heterogeneity of urban development and LST, and it is difficult to deeply explore the impact of spatial variability of urban development on LST. (2) The research elements are not synergistic. Many studies only conduct a single analysis of land cover or climate environment, and fail to study the common time changes and impact relationship between the two. It is difficult to provide a reasonable reference for environmental governance in different regions. The possible innovations in this research are as follows. (1) This article is based on remote sensing image data of the past two decades (2000, 2011, 2020). Through comparative analysis of multiple methods, CA and ANN models are used to predict the growth trend of urban land cover and LST in 2030 and 2040. (2) Analyze the synergy between LULC and LST under different time conditions, and quantify the impact of urban LULC changes on urban thermal environment.
2. Materials

2.1. Study Area

This study uses the main urban areas of Wuhan (Jiang’an District, Jianghan District, Qiaokou District, Hanyang District, Wuchang District, Hongshan District, and Qingshan District) as the case study area. Wuhan is an important industrial city in China, as well as the central economic center and regional transportation hub. It has jurisdiction over 7 central urban areas and 6 remote urban areas. Wuhan is located at 113°41′–115°05′ east longitude and 29°58′–31°22′ north latitude (Figure 1). It belongs to the humid north subtropical monsoon climate, with abundant rainfall, sufficient sunshine, hot summers and cold winters. Generally, the average annual temperature is 15.8–17.5 °C. During the year, the average temperature in January is the lowest at 0.4 °C; the average temperature in July and August is the highest at 28.7 °C. The summer is extremely long for 135 days. Because Wuhan is located at 30 degrees north latitude, the sun can reach 38 degrees at noon in summer. It is also located inland and far from the ocean. The terrain is like a basin. The rain is concentrated in the rainy season in early summer, and heat collection is easy to dissipate. There are many rivers and lakes, so there is more water vapor at night. Coupled with the urban heat island effect and the control of the subtropical high during the summer drought, it is very muggy. Wuhan is known as the “Three Big Stoves”, and the urban heat island effect is significant.

Figure 1. Location map of the study area (a) and elevation map of the study area (b).

2.2. Data Sources

In this study, the Landsat multispectral remote sensing image data are all sourced from the official website of U.S. Geological Survey (http://earthexplorer.usgs.gov/ (14 June 2020). Landsat 4–5 Thematic Mapper (TM) and Landsat 8 Operational Land Imager (OLI) image data are used as research data to monitor LULC changes, LST distribution and related indexes (Table 1). In order to truly reflect the condition of the ground objects, it eliminates the problems of image data overlap and radiant brightness distortion caused by factors such as the sun’s height, terrain, atmosphere, and the photoelectric system of the sensor itself. It needs to pre-process the image with radiometric calibration, atmospheric correction, band removal and other preprocessing [35]. This study used ENVI 5.3 to preprocess the image, and the data analysis was done using ArcGIS 10.3, IBM SPSS 21 and other software. Wuhan city administrative division vector data and meteorological station data are auxiliary data. The vector data of the administrative division of Wuhan City is used for the division of the study area. The weather station data is used for the temperature verification of the inversion, and comes from the China Meteorological Data Network (http://data.cma.cn/ (25 May 2020)).
Table 1. Description of the collected seasonal satellite image from USGS Earth Explorer.

<table>
<thead>
<tr>
<th>Date of Acquisition</th>
<th>Sensor</th>
<th>Path/Row</th>
<th>Multi-Spectral Band Resolution</th>
<th>Cloud Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>27 July 2000</td>
<td>TM</td>
<td>123/39</td>
<td>30 m</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>8 June 2011</td>
<td>TM</td>
<td>123/39</td>
<td>30 m</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>3 August 2020</td>
<td>OLI_TIRS</td>
<td>123/39</td>
<td>30 m</td>
<td>&lt;10%</td>
</tr>
</tbody>
</table>

3. Methods

3.1. LULC Classification

The research applied the support vector machine (SVM) supervised classification method in ENVI 5.3 software. In this supervised classification, images are classified using polygons (training samples/signatures), which represent separate sample regions of different LULC types to be classified [36]. The images are divided into four types of land cover, namely water body (including rivers, lakes, ponds and any other wetlands), vegetation (including woodland, shrubland and any other vegetation surface), bare land (including sand, bare soil, and no vegetation land) and built-up area (including buildings, roads and any other impervious surfaces).

3.2. Accuracy of LULC Data

In this paper, about 50 samples are collected for each LULC land use type in the ENVI 5.3 software platform, and the LULC map of the corresponding year is generated. The accuracy of the LULC classification map in each issue is mainly determined by randomly selecting 350 sample data and analyzing them in the Global Positioning System (GPS) and Google Earth images and ground truth data [37,38]. This is one of the important methods for evaluating the accuracy of image classification. Among them, the overall accuracy, user accuracy, producer accuracy, and kappa statistic formula are as follows [39].

\[
\text{Overall accuracy} = \frac{\text{Total number of corrected classified pixels(diagonal)}}{\text{Total number of reference pixels}} \times 100 \quad (1)
\]

\[
\text{User accuracy} = \frac{\text{Number of correctly classified pixels in each category(diagonal)}}{\text{Total number of reference pixels in each category(row total)}} \times 100 \quad (2)
\]

\[
\text{Producer accuracy} = \frac{\text{Number of correctly classified pixels in each category(diagonal)}}{\text{Total number of reference pixels in each category(column total)}} \times 100 \quad (3)
\]

\[
\text{Kappa coefficient}(T) = \frac{\text{Total number of sample of corrected sample}}{\text{Total number of sample}^2 - \sum(\text{col}.\text{tol}.\text{row tol})} \times 100 \quad (4)
\]

3.3. Inversion of Surface Temperature

According to the characteristics of different thermal infrared remote sensing data, there are three main types of land surface temperature retrieval algorithms: atmospheric correction method, single-channel algorithm and split window algorithm [20,26,40]. In this paper, the atmospheric correction method suitable for Landsat satellite imagery is used to retrieve the surface temperature of the thermal infrared bands of TM and TIRS. Landsat8 TIRS has two thermal infrared bands, b10 and b11. In this study, b10, which has a similar spectral range to the TM thermal infrared band, is selected for the inversion of land surface temperature [19,41]. The principle of the atmospheric correction method is to first estimate the influence of the atmosphere on the surface thermal radiation, and then subtract this part of the atmospheric influence from the total thermal radiation observed by the satellite sensors to obtain the surface thermal radiation intensity. Finally, this heat radiation intensity will be converted into the corresponding surface temperature [42]. The thermal infrared radiance value \( L_A \) received by the satellite sensor is composed of three parts: the upward radiance of the atmosphere \( L_A \), the true radiance of the ground reaches the satellite sensor energy after passing through the atmosphere; the downward radiation
of the atmosphere reaches the ground and the reflected energy $L_{↓}$ [36]. The formula of the thermal infrared radiance value $L_{λ}$ received by the satellite sensor is (radiation transfer equation):

$$L_{λ} = \left[ \varepsilon \times B(T_s) + (1 - \varepsilon) \times L_{↓} \right] \times \tau + L_{↑}$$

(5)

In Formula (5), $\varepsilon$ is the surface emissivity, $T_s$ is the true surface temperature (K), $B(T_s)$ is the black body thermal radiation brightness, and $\tau$ is the transmittance of the atmosphere in the thermal infrared band. Before retrieving the surface temperature, the remote sensing image needs to be preprocessed by radiometric calibration and atmospheric correction. Radiation calibration refers to the received remote sensing data, usually gray value (DN value), which will be converted into radiation or ground reflectivity [43]. Atmospheric correction is to eliminate the influence of atmospheric molecular factors and aerosol scattering on the ground object reflection, and accurately obtain the ground object reflectance information. This article uses the FLASH module in ENVI 5.3 to perform atmospheric correction on the TM/TIRS images in the study area. The specific steps of surface temperature inversion are shown in Appendix A.

3.4. The Prediction of LULC Map

This paper uses the MOLUSCE (Land Use Change Assessment Module) plug-in in the QGIS 2.18 software to predict future LULC changes [20]. The plug-in includes several steps starting from the input module, regional change analysis, modeling method, simulation and verification. We load the independent variable and the dependent variable into the input module, where the independent variable is the LULC raster data of the main urban area of Wuhan, and the dependent variable includes the raster data of elevation, slope, distance from residential, commercial, road and water body. In order to predict the LULC maps in 2030 and 2040, the LULC maps in the main urban area of Wuhan in 2000, 2011 and 2020 are used. We analyze the area change in the MOLUSCE plug-in, calculate the LULC change between two time periods, and generate a transition matrix and a LULC change graph. In the modeling method stage, the ANN model is used to predict the LULC change conversion potential, and the maximum iteration is set to 1000. We set its neighborhood pixels to 9 (3 × 3) units to define the maximum iteration of the model and neighborhood pixels for neural network analysis and prediction [44]. In order to verify the accuracy of the predicted LULC map, we use Kappa statistics for verification.

3.5. The Prediction Process of LST

In the practice of case studies, more machine learning methods are generally used, including feedforward neural network (FFNN), radial basis function neural network (RBFNN), Kohonen self-organizing neural network, recurrent neural network (RNN), convolutional neural network (CNN) and modular neural network (MNN) [45–47]. Before predicting the LST, we need to input parameters such as the LST, LULC map, NDBI, NDBSI image, latitude and longitude of the past years into the model to identify hidden patterns in the data set and generate predictions by moving along the time series data. The LST maps of the main urban area of Wuhan in 2000 and 2011 were used to simulate the LST maps in 2030 and 2040. This article mainly uses two machine learning methods, ANN and Elman, and we run them in MATLAB 2016b to predict the LST in the main urban area of Wuhan in 2030 and 2040. In order to choose a more appropriate prediction algorithm, we compare the prediction performance of the two algorithms by the correlation coefficient (R) and root mean square error (RMSE) of the 2020 forecast and the actual interpreted LST data [31,48]. The Formulas are as (6) and (7).

$$RMSE = \sqrt{\frac{\sum (T_{obs} - T_{model})^2}{n}}$$

(6)
The correlation coefficient results show that the results of the fitting are significantly different (Figure 2). In the prediction of LST, the R of ANN and Elman algorithms are 0.885 and 0.080, respectively, and the corresponding RMSE are 0.513 and 0.705, respectively. It shows that the prediction results of the ANN algorithm are better and more robust, and are more suitable for predicting the LST in this study area. It shows that the prediction result of the ANN algorithm is more robust and more suitable for predicting the LST in this study area [20].

Through the correlation coefficient results, it is found that the results of the fitting are significantly different (Figure 2). In the prediction of LST, the R of ANN and Elman algorithms are 0.885 and 0.080, respectively, and the corresponding RMSE are 0.513 and 0.705, respectively. It shows that the prediction results of the ANN algorithm are better and more robust, and are more suitable for predicting the LST in this study area. It shows that the prediction result of the ANN algorithm is more robust and more suitable for predicting the LST in this study area [20].

\[
R = \frac{\sum (T_{\text{obs}} - T_{\text{obs}}) (T_{\text{model}} - T_{\text{model}})}{\sqrt{\sum (T_{\text{obs}} - T_{\text{obs}})^2 \sum (T_{\text{model}} - T_{\text{model}})^2}}
\]

(7)

3.6. ANN Model

Based on the existing remote sensing image data, we build a 100 × 100 fishing net in ArcMap software to generate sampling points. Then extract the LULC pixel value to the center of the grid and use this value as one of the input parameters of ANN [26]. Similarly, we extracted the LST, NDBI, NDBSI, longitude and latitude of the main urban area of Wuhan in 2000, 2011 and 2020 respectively. Next, to predict the future LST of the main urban area in 2030 and 2040, a sequence of past years LST, LULC, NDBI, BDBSI images at an interval of 10 and 20 years, Latitude and longitude were provided as input parameters to the ANN model to identify hidden patterns in the data set and generate predictions by moving along the time series data [48,49]. The formulas are as follows.

\[
\begin{align*}
LST (t + 10) &= f [LST(t), LST(t - 10), LULC(t), LULC(t - 10), NDBI(t), NDBI(t - 10), NDBSI(t), NDBSI(t - 10)] \\
LST (t + 20) &= f [LST(t), LST(t - 20), LULC(t), LULC(t - 20), NDBI(t), NDBI(t - 20), NDBSI(t), NDBSI(t - 20)]
\end{align*}
\]

(8) (9)

In the Formulas (8) and (9), the \( t = 2020 \).

4. Results and Analysis

4.1. Variation in Past LULC Patterns

Based on the ENVI 5.3 platform, we used the SVM algorithm to classify the remote sensing image data over the years (Figure 3). The accuracy of the LULC classification map is described in Table 2. The percentages of overall accuracy and Kappa coefficient of LULC classification maps in 2000, 2011 and 2020 are 98.12%, 87.97%, 91.37% and 89.68%, 91.25%, and 90.18%, respectively. The percentages of overall accuracy and Kappa coefficient for all years are greater than 87.97%, indicating that the classification shows high accuracy [50,51].
On the whole, the vegetation area is showing a continuous decreasing trend, the built-up area is continuously expanding, and the water body and bare land area are also showing a continuous decreasing trend.

![LULC classification of the study area using the SVM algorithm.](image)

**Table 2.** Accuracy assessment of classified LULC maps using SVM algorithm.

<table>
<thead>
<tr>
<th>Year</th>
<th>User Accuracy (%)</th>
<th>Producer Accuracy (%)</th>
<th>Overall Accuracy (%)</th>
<th>Kappa Statistics (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>87.56</td>
<td>89.51</td>
<td>98.12</td>
<td>89.68</td>
</tr>
<tr>
<td>2011</td>
<td>89.16</td>
<td>90.33</td>
<td>87.97</td>
<td>91.25</td>
</tr>
<tr>
<td>2020</td>
<td>91.08</td>
<td>88.24</td>
<td>91.37</td>
<td>90.18</td>
</tr>
</tbody>
</table>

Statistics on the area of various land types are shown in Table 3. It shows that the area of vegetation land decreased from 370.675 km² in 2000 to 283.554 km² in 2011, with a decrease of 23.50%. In 2020, the area of vegetation land further reduced to 246.859 km², with a decrease of 12.94%. From the two stages, compared with the previous 10 years, the decline of vegetation area in the next 10 years has slowed by nearly 50%. The reason may be the continuous expansion of construction land in the main urban area from 2000 to 2020, which has squeezed some space for vegetation.
Table 3. Changes in LULC class from 2000 to 2020.

<table>
<thead>
<tr>
<th>Class</th>
<th>2000</th>
<th>2011</th>
<th>2020</th>
<th>2000–2020</th>
<th>Rate of Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>370.675</td>
<td>283.554</td>
<td>246.859</td>
<td>−123.816</td>
<td>−50.16</td>
</tr>
<tr>
<td>Built-Up Area</td>
<td>267.683</td>
<td>388.852</td>
<td>452.111</td>
<td>184.428</td>
<td>40.79</td>
</tr>
<tr>
<td>Water Body</td>
<td>320.742</td>
<td>289.869</td>
<td>263.898</td>
<td>−56.844</td>
<td>−21.54</td>
</tr>
<tr>
<td>Bare Land</td>
<td>8.365</td>
<td>5.192</td>
<td>4.599</td>
<td>−3.766</td>
<td>−81.88</td>
</tr>
</tbody>
</table>

From the two stages, compared with the previous 10 years, the decline of vegetation area in the next 10 years has slowed by nearly 50%. The reason may be the continuous expansion of construction land in the main urban area from 2000 to 2020, which has squeezed some vegetation. The decline in the next 10 years is mainly due to the government’s implementation of the concept of “green water and green mountains are golden mountains and silver mountains” in the new round of urban planning. More attention has been paid to the healthy and sustainable development of urban green development, the addition of “pocket parks” and the strengthening of greening renovation in urban gray areas.

The urban construction area has continued to increase, expanding from 267.683 km² in 2000 to 388.852 km² in 2011, and reaching 452.111 km² in 2020. During this period, the area of urban construction land increased by 40.79%, which is closely related to the rapid development of urbanization in Wuhan during this period. Due to the continuous economic and social development and the continuous population expansion, the construction area of the main urban area, transportation construction land and public activity space needs to continue to rise, causing more construction land to squeeze the phenomenon of other types of land. During the period, the area of bare land decreased by 81.88%. The area of bare land decreased from 8.365 km² in 2000 to 5.192 km² in 2011, and further decreased to 4.599 km² in 2020. It may be due to the increased intensity of the landscape design and renovation projects on both sides of the Yangtze River during this period, which led to the reduction of the bare land in the river bed during this period.

Similarly, the area of the water body has dropped from 320.742 km² in 2000 to 263.898 km² in 2020, and the area has decreased by −56.844 km², a decrease of 23.578%. This may be because in order to increase and supplement the area of farmland in the Jianghan Plain, there have been many blind reclamation activities before 2011, such as lake filling and weir filling for farmland. In addition, due to the continuous expansion of the built-up area in the main urban area, some rivers and lakes are shrinking.

4.2. Analysis of LST Changes

Due to the limitation of the research location, it is difficult to obtain the specific temperature of a specific area. However, the use of satellite imagery and remote sensing technology can effectively estimate the urban LST on the micro and macro scales. We estimated the summer LST in the main urban area of Wuhan based on Landsat’s thermal induction zone, and combined the characteristics of the LST results to divide it into four types: <25 °C, 25~30 °C, 30~35 °C, >35 °C (Figure 4). In 2000, the highest temperature in the main urban area of Wuhan was 32.1 °C and the lowest temperature was 16.4 °C. The remote sensing image of 2000 was collected on July 27, which was the mid-summer time, but the temperature was relatively low at this time. In 2011, the highest temperature in the main urban area was 36.3 °C and the lowest temperature was 22.3 °C. Compared with 2000, the lowest and highest temperatures increased by 4.2 and 5.1 °C respectively. The remote sensing images in 2011 were collected on June 8, which was in the early summer period, and the highest temperature at this time node reached 36.3 °C. In 2020, the minimum and maximum temperatures in the main Wuhan urban area will be 20.6 and 37.3 °C, respectively. This is compared with 2000 and 2011. Except for a slight drop in the minimum temperature, the temperature in the main urban area in 2020 will be higher overall. Compared with
other cities at the same latitude, the LST of the main urban area for three years is generally higher, which also verifies Wuhan’s reputation as China’s “three major stoves”.

![Figure 4. LST variation in the study area.](image)

The area of four types of temperature is counted in ArcGIS 10.3 (Table 4). It reveals that the summer LST in the main urban area of Wuhan is mainly in the two intervals of 25–30 °C and 30–35 °C from 2000 to 2020. These two temperature ranges account for more than 77%, but the area with a temperature >35 °C accounts for less than 3%. In 2000, the area with the temperature less than 25 °C was 152.015 km², which only accounted for 15.713% of the main urban area. The area in this temperature range in 2011 and 2020 are 139.273 and 126.448 km², respectively.

<table>
<thead>
<tr>
<th>Year</th>
<th>&lt;25 °C Area (Km²)</th>
<th>&lt;25 °C Ratio (%)</th>
<th>25~30 °C Area (Km²)</th>
<th>25~30 °C Ratio (%)</th>
<th>30~35 °C Area (Km²)</th>
<th>30~35 °C Ratio (%)</th>
<th>&gt;35 °C Area (Km²)</th>
<th>&gt;35 °C Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>152.015</td>
<td>15.713</td>
<td>501.872</td>
<td>51.874</td>
<td>313.581</td>
<td>32.413</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>139.273</td>
<td>14.396</td>
<td>398.596</td>
<td>41.200</td>
<td>423.066</td>
<td>43.729</td>
<td>6.533</td>
<td>0.675</td>
</tr>
<tr>
<td>2020</td>
<td>126.448</td>
<td>13.069</td>
<td>315.661</td>
<td>32.627</td>
<td>515.108</td>
<td>53.242</td>
<td>10.251</td>
<td>1.059</td>
</tr>
</tbody>
</table>

4.3. Changes in LST under Different LULC

We conducted statistics on the LST area of Wuhan’s main urban area under different LULCs in ArcGIS 10.3 (Table 5). It shows that the highest temperature range in the main urban area in 2000 was 25–30 °C, and the proportion of water body in this range reached 20.580%, followed by vegetation, which accounted for 18.694%. The temperature range with the second highest proportion is 30–35 °C, of which the built-up area accounts for
15.917%, and the vegetation accounts for 14.616%. Compared with 2000, the situation in 2011 is slightly different. This mainly shows that the highest temperature range in the main urban area is 30–35 °C, and the built-up area accounts for as high as 30.937%. In the range of 25–30 °C, the proportions of vegetable and water body are still significant, 16.709% and 15.910% respectively. In 2020, the temperature range with the highest proportion of the main urban area is 25–30 °C, and the proportion of built-up area in this range reaches the maximum, which is 36.606%, followed by vegetation, with a proportion of 10.239%.

Table 5. Cross-linkage of LST variation over different LULC classes.

<table>
<thead>
<tr>
<th>LULC</th>
<th>&lt;25 °C</th>
<th>25–30 °C</th>
<th>30–35 °C</th>
<th>&gt;35 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Area</td>
<td>Ratio (%)</td>
<td>Area</td>
<td>Ratio (%)</td>
</tr>
<tr>
<td>Vegetation</td>
<td>33.214</td>
<td>3.433</td>
<td>183.474</td>
<td>18.964</td>
</tr>
<tr>
<td>Built-up Area</td>
<td>11.435</td>
<td>1.182</td>
<td>114.847</td>
<td>11.871</td>
</tr>
<tr>
<td>Water Body</td>
<td>105.413</td>
<td>10.896</td>
<td>199.106</td>
<td>20.580</td>
</tr>
<tr>
<td>Bare Land</td>
<td>1.953</td>
<td>0.202</td>
<td>4.445</td>
<td>0.459</td>
</tr>
</tbody>
</table>

2011

<table>
<thead>
<tr>
<th>LULC</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>30.521</td>
<td>3.155</td>
<td>161.651</td>
<td>16.709</td>
<td>90.757</td>
<td>9.381</td>
<td>0.817</td>
<td>0.084</td>
</tr>
<tr>
<td>Built-up Area</td>
<td>5.120</td>
<td>0.529</td>
<td>80.276</td>
<td>8.298</td>
<td>299.307</td>
<td>30.937</td>
<td>3.978</td>
<td>0.411</td>
</tr>
<tr>
<td>Water Body</td>
<td>103.554</td>
<td>10.704</td>
<td>153.924</td>
<td>15.910</td>
<td>32.361</td>
<td>3.345</td>
<td>0.013</td>
<td>0.001</td>
</tr>
<tr>
<td>Bare Land</td>
<td>0.078</td>
<td>0.008</td>
<td>2.765</td>
<td>0.286</td>
<td>0.641</td>
<td>0.066</td>
<td>1.708</td>
<td>0.177</td>
</tr>
</tbody>
</table>

2020

<table>
<thead>
<tr>
<th>LULC</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
<th>Area</th>
<th>Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>32.321</td>
<td>3.341</td>
<td>114.685</td>
<td>11.854</td>
<td>99.057</td>
<td>10.229</td>
<td>0.915</td>
<td>0.095</td>
</tr>
<tr>
<td>Built-up Area</td>
<td>6.921</td>
<td>0.715</td>
<td>85.258</td>
<td>8.812</td>
<td>354.149</td>
<td>36.606</td>
<td>5.783</td>
<td>0.598</td>
</tr>
<tr>
<td>Water Body</td>
<td>86.244</td>
<td>8.914</td>
<td>112.779</td>
<td>11.657</td>
<td>61.785</td>
<td>6.386</td>
<td>3.058</td>
<td>0.316</td>
</tr>
<tr>
<td>Bare Land</td>
<td>0.962</td>
<td>0.099</td>
<td>2.894</td>
<td>0.299</td>
<td>0.241</td>
<td>0.025</td>
<td>0.502</td>
<td>0.052</td>
</tr>
</tbody>
</table>

It is worth noting that the built-up area occupies a larger area in the three temperature ranges of 25–30 °C, 30–35 °C and >35 °C. This feature is particularly obvious in 2020. Compared with other land types, the proportion of built-up area in these three temperature ranges is relatively high, which are 34.098%, 11.287% and 0.098%, respectively. This may be because the built-up area is mostly urban asphalt or concrete roads and buildings. They will produce long-term infrared heat waves and high thermal radiation, which will cause the LST of the main urban area to rise, and the rise of LST is an important factor in causing UHI. It is well known that vegetation and water body have a regulating effect on urban LST. Therefore, we can moderately increase the water body and vegetation space in order to regulate and reduce the UHI effect.

4.4. Prediction and Analysis of LULC

The LULC and DEM data of the main urban area of Wuhan in 2000, 2011, and 2020, as well as the power factor raster data of Slope and Road were added to the QGIS 2.18-MOLUSCE plug-in. We use the ANN simulation function of the MOLUSCE plug-in to predict LULC in 2030 and 2040. Before that, we need to use the 2000 LULC and 2011 LULC to predict the LULC of the main urban area in 2020. Subsequently, we tested the predicted results with the real 2020 LULC results in the MOLUSCE plug-in. The QGIS-MULUSCE Plugin module verification result shows that the correctness of LULC in 2020 is predicted to reach 84.018%, Kappa (overall) is 0.756, Kappa (histo) is 0.898, and Kappa (location) is 0.843. It shows that the prediction results of this model are robust and suitable for predicting future LULC under this standard [52].

The LULC in the main urban area of Wuhan in 2030 and 2040 is predicted in QGIS-MOLUSCE, and the area of various types of LULC is calculated (Figure 5 and Table 6). Comparing the LULC of 2030 and 2040 to ArcGIS 10.3 and the comparison of the LULC of 2020, it is found that the built-up area of Wuhan’s main urban area has further expanded.
In 2030 and 2040, the built-up area will be expanded from 2000 to 511.510 km\(^2\) and 545.282 km\(^2\). In contrast, in 2030 and 2040, the areas of vegetation, water body and bare land in the main urban area are shrinking, and the reductions are 18.76%, 4.03%, 53.28% and 24.78%, 11.10%, and 58.74%, respectively.

![Figure 5. Predicted LULC classes.](image)

**Table 6.** Predicted LULC changes in the study area for 2030 and 2040.

<table>
<thead>
<tr>
<th>LULC Class</th>
<th>Area(Km(^2))</th>
<th>2020~2030</th>
<th>2020~2040</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>370.6749</td>
<td>283.554</td>
<td>200.559</td>
</tr>
<tr>
<td>Built-up Area</td>
<td>267.6834</td>
<td>388.8522</td>
<td>452.111</td>
</tr>
<tr>
<td>Water Body</td>
<td>320.742</td>
<td>289.8693</td>
<td>263.898</td>
</tr>
<tr>
<td>Bare Land</td>
<td>8.3646</td>
<td>5.1921</td>
<td>2.13</td>
</tr>
</tbody>
</table>

4.5. Prediction and Analysis of LST

Predicting and simulating the LST have an important guiding role in the evaluation of the future UHI in the study area and the formulation of the adjustment mechanism. As described in Section 3.5 of the article. The more parameters entered in the model can improve the prediction accuracy of the ANN. In this paper, NDBI, NDBSI, LULC, latitude and longitude are used as supporting parameters and input into the ANN model for prediction. In the prediction of LST, the R and RMSE of the prediction result are 0.885 and 0.513, respectively, indicating that the prediction result of the ANN algorithm is robust. Then, the summer LST of the main urban area of Wuhan in 2030 and 2040 was predicted, and the area of each temperature interval was counted in ArcGIS 10.3 (Figure 6 and Table 7).
Figure 6 and Table 7 show that the temperature range of <25 °C and 25~30 °C in the main urban area will be tightened in 2030 and 2040. The range of 30~35 °C and >35 °C continues to expand, of which 30~35 °C will rise from 32.413% in 2000 to 52.925% and 55.219% in 2030 and 2040, respectively. The temperature range of >35 °C has been expanded from 2.095 km² in 2020 to 15.264 km² in 2030 and 33.612 km² in 2040. From the perspective of extreme temperature, the future temperature in the main urban area will rise to varying degrees. In 2030, the minimum temperature in the main urban area in the summer of 2040 will be 22.1 °C and 23.6 °C, and the maximum temperature will be 38.5 °C and 39.1 °C. The expansion trend of the main urban area in the range of 30~35 °C and >35 °C is somewhat similar to the expansion trend of built-up area in 2040. This shows that if the development of rapid urbanization is not rationally curbed, it will inevitably further aggravate the urban heat island effect.

5. Discussion and Conclusions

With the development of GIS and RS technology, although the application of CA and ANN models can provide effective methods for the prediction of LULC and LST, the accuracy of the prediction results needs to be improved [53]. The prediction results of CA and ANN models may only predict the development trend as a whole, and their ability to identify the relationship between variables is limited. This is not enough to clearly predict the development process of LULC and LST. In addition, in order to obtain better prediction results, we try to find remote sensing image data with consistent time and high resolution,
but there are still some differences in time and the resolution of the image is not high \cite{51,53}. Therefore, these shortcomings should be considered to further improve the accuracy of the forecast in the following research and provide a more reliable reference for the green development and sustainable development of the city.

This research is based on remote sensing image data of LULC and LST in the main urban area of Wuhan in 2000, 2011 and 2020. We use the CA and ANN models in the Molusce plug-in of the QGIS 2.18 software platform to predict the LULC and LST in 2030 and 2040. We can get the following conclusions.

1. In the past 20 years in Wuhan, with the rapid development of urbanization, the land use structure of the main urban area has undergone great changes. The area of built-up area in the main urban area of Wuhan continues to expand. Compared with 2000, the area of built-up area in 2020 has increased by 40.79%.

2. The expansion of built-up area in the main urban area is mainly by replacing or squeezing vegetation, water bodies and bare land. The area of bare land and vegetation decreased significantly, with a decrease of 81.88% and 50.16%, respectively. According to the forecast results, the built-up area will be further expanded in 2030 and 2040.

3. From 2000 to 2020, except for the slight decrease in the minimum temperature in the summer of 2010 in the main urban area of Wuhan, the overall temperature has shown a gradual upward trend. In 2030 and 2040, the predicted extreme temperature of LST has shown signs of increasing compared with previous years. The 30–35 °C and >35 °C areas in the main urban area continue to expand. The direction of the expansion range of the LST temperature range is obviously similar to the expansion of the built-up area.

**Author Contributions:** M.Z., Conceptualization, Methodology, Data Curation, Resources, Writing—Original Draft, Writing—Review and Editing; S.T., Supervision; C.Z. and A.-A.K., Writing—Review and Editing, Technological Guidance. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study due to the fact that the analyzed datasets were properly anonymized, and no participant could be identified.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the reviewers for their expertise and valuable input.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A**

Inversion of surface temperature.

According to the characteristics of different thermal infrared remote sensing data, there are three main types of land surface temperature retrieval algorithms: atmospheric correction method, single-channel algorithm and split window algorithm \cite{26,54}. In this paper, the atmospheric correction method suitable for Landsat satellite imagery is used to retrieve the surface temperature of the thermal infrared bands of TM and TIRS. Landsat8 TIRS has two thermal infrared bands, \( b_{10} \) and \( b_{11} \). In this study, \( b_{10} \), which has a similar spectral range to the TM thermal infrared band, is selected for the inversion of land surface temperature \cite{19,55}. The principle of the atmospheric correction method is to first estimate the influence of the atmosphere on the surface thermal radiation, and then subtract this part of the atmospheric influence from the total thermal radiation observed by the satellite sensors to obtain the surface thermal radiation intensity. Finally, this heat radiation intensity will be converted into the corresponding surface temperature \cite{42,56}. The thermal infrared radiance value \( L_{\lambda} \) received by the satellite sensor is composed of three parts: the upward
radiance of the atmosphere $L^\uparrow$, the true radiance of the ground reaches the satellite sensor energy after passing through the atmosphere; the downward radiation of the atmosphere reaches the ground and the reflected energy $L^\downarrow$ [57]. The formula of the thermal infrared radiance value $L_\lambda$ received by the satellite sensor is (radiation transfer equation):

$$L_\lambda = [\varepsilon \times B(T_s) + (1 - \varepsilon) \times L^\downarrow] \times \tau + L^\uparrow$$  \hspace{1cm} (A1)

In Formula (A5), $\varepsilon$ is the surface emissivity, $T_s$ is the true surface temperature (K), $B(T_s)$ is the black body thermal radiation brightness, and $\tau$ is the transmittance of the atmosphere in the thermal infrared band. Before retrieving the surface temperature, the remote sensing image needs to be preprocessed by radiometric calibration and atmospheric correction. Radiation calibration refers to the received remote sensing data, usually gray value (DN value), which will be converted into radiation or ground reflectivity [58]. Atmospheric correction is to eliminate the influence of atmospheric molecular factors and aerosol scattering on the ground object reflection, and accurately obtain the ground object reflectance information. This article uses the FLASSH module in ENVI 5.3 to perform atmospheric correction on the TM/TIRS images in the study area. The specific steps for inversion of surface temperature are as follows.

$$NDVI = (B_{Nir} - B_{Red}) / (B_{Nir} + B_{Red})$$  \hspace{1cm} (A2)

$$F_V = (NDVI - NDVI_S) / (NDVI_V - NDVI_S)$$  \hspace{1cm} (A3)

In Formula (A2), $B_{Nir}$ and $B_{Red}$ represent the reflectivity of the image in the near-infrared band and the red band, respectively. In Formula (A3), $F_V$ is the vegetation coverage, and $NDVI_S$ is the $NDVI$ value of non-vegetation pixels. $NDVI_V$ represents the $NDVI$ value of pure vegetation pixels.

Secondly, the surface specific emissivity thermal infrared radiance is calculated. The surface emissivity refers to the ratio of the amount of radiation emitted by the surface to the amount of radiation emitted by a black body at the same temperature. There are obvious differences in the surface emissivity of various land use types. Qin et al. proposed to divide the types of ground features into three types: water bodies, natural surfaces, and towns, and then calculate their respective surface emissivity values [36]. The structure of the water surface is simple, and the specific emissivity in the thermal band is very high, so the default emissivity value of the water body is 0.995.

$$\varepsilon_{surface} = 0.9526 + 0.0614F_V - 0.0461F_V^2$$  \hspace{1cm} (A4)

$$\varepsilon_{building} = 0.9589 + 0.086F_V - 0.0671F_V^2$$  \hspace{1cm} (A5)

where, the $\varepsilon_{surface}$ is the specific emissivity of the natural surface pixel; $\varepsilon_{building}$ is the specific emissivity of the town. The calculation formula for the radiance $B(T_S)$ of a blackbody at a temperature of $T$ in the thermal infrared band is as follows:

$$B(T_S) = \left[ L_\lambda - L^\uparrow - \tau (1 - \varepsilon) L^\downarrow \right] / \tau \cdot \varepsilon$$  \hspace{1cm} (A6)

We enter the imaging time and intermediate latitude and longitude in NASA (http://atmcsc.gsfc.nasa.gov/) (16 June 2020) to obtain the parameters in the formula.

Finally, the true surface temperature is retrieved. Because of the difference between the brightness temperature and the true temperature of the earth’s surface [46–48]. After obtaining the radiance of a black body with a temperature of $T_S$ in the thermal infrared band, it is necessary to obtain the true surface temperature $T_S$ according to the inverse function of Planck’s formula [49,59]. The formula is as follows:

$$T_S = \frac{K_2}{\ln \left( \frac{K_1}{\pi(T_S)} + 1 \right)} - 273.15$$  \hspace{1cm} (A7)

$$\varepsilon_{surface} = 0.9526 + 0.0614F_V - 0.0461F_V^2$$  \hspace{1cm} (A4)

$$\varepsilon_{building} = 0.9589 + 0.086F_V - 0.0671F_V^2$$  \hspace{1cm} (A5)

where, the $\varepsilon_{surface}$ is the specific emissivity of the natural surface pixel; $\varepsilon_{building}$ is the specific emissivity of the town. The calculation formula for the radiance $B(T_S)$ of a blackbody at a temperature of $T$ in the thermal infrared band is as follows:

$$B(T_S) = \left[ L_\lambda - L^\uparrow - \tau (1 - \varepsilon) L^\downarrow \right] / \tau \cdot \varepsilon$$  \hspace{1cm} (A6)

We enter the imaging time and intermediate latitude and longitude in NASA (http://atmcsc.gsfc.nasa.gov/) (16 June 2020) to obtain the parameters in the formula.

Finally, the true surface temperature is retrieved. Because of the difference between the brightness temperature and the true temperature of the earth’s surface [46–48]. After obtaining the radiance of a black body with a temperature of $T_S$ in the thermal infrared band, it is necessary to obtain the true surface temperature $T_S$ according to the inverse function of Planck’s formula [49,59]. The formula is as follows:

$$T_S = \frac{K_2}{\ln \left( \frac{K_1}{\pi(T_S)} + 1 \right)} - 273.15$$  \hspace{1cm} (A7)
In Formula (A7), $K_1$ and $K_2$ are constant terms, and different image data have different values. For TM data, $K_1 = 607.76$, $K_2 = 1260.56$, and for TIRS data, $K_1 = 774.89$, $K_2 = 1321.08$.

References


6. Cai, Z.; Liu, Q.; Cao, S. Real Estate Supports Rapid Development of China’s Urbanization. *Land Use Policy* 2020, 95, 104582. [CrossRef]


27. Li, J.; Li, J.; Yuan, Y.; Li, G. Spatiotemporal Distribution Characteristics and Mechanism Analysis of Urban Population Density: A Case of Xi’an, Shaanxi, China. Cities 2019, 86, 62–70. [CrossRef]


42. Luo, Y.; Sun, W.; Yang, K.; Zhao, L. China Urbanization Process Induced Vegetation Degradation and Improvement in Recent 20 Years. Cities 2021, 114, 103207. [CrossRef]

43. Liu, L.; Zhang, Y. Urban Heat Island Analysis Using the Landsat TM Data and ASTER Data: A Case Study in Hong Kong. Remote Sens. 2011, 3, 1535–1552. [CrossRef]


45. Hanberry, B.B. Timing of Tree Density Increases, Influence of Climate Change, and a Land Use Proxy for Tree Density Increases in the Eastern United States. Land 2021, 10, 1121. [CrossRef]


47. Sejati, A.W.; Buchori, I.; Rudiarto, I. The Spatio-Temporal Trends of Urban Growth and Surface Urban Heat Islands over Two Decades in the Semarang Metropolitan Region. Sustain. Cities Soc. 2019, 46, 101432. [CrossRef]


52. Santamouris, M. Recent Progress on Urban Overheating and Heat Island Research. Integrated Assessment of the Energy, Environmental, Vulnerability and Health Impact. Synergies with the Global Climate Change. *Energy Build.* 2020, 207, 109482. [CrossRef]


