Analysis of Land Surface Temperature Driving Factors and Spatial Heterogeneity Research Based on Geographically Weighted Regression Model

Yin Zhi, Liang Shan, Lina Ke, and Ruxin Yang

School of Geography, Liaoning Normal University, Dalian 116029, China

Correspondence should be addressed to Liang Shan; sl-lnnu@lnnu.edu.cn and Lina Ke; linake@lnnu.edu.cn

Received 29 April 2020; Revised 4 July 2020; Accepted 20 July 2020; Published 6 October 2020

Abstract

Acceleration of urbanization has brought about a series of problems, which include irreversible changes to urban surfaces and continuous increases in land surface temperatures (LSTs). In this context, analysis of the driving factors and spatial heterogeneity of urban LST is of considerable importance for mitigating urban heat island effects and promoting healthy and comfortable urban living environments. This study explored the relationship between the spatial characteristics and driving factors of the LST by using a geographically weighted regression (GWR) model to analyze multisourcedata from the Xigang District of Dalian City. The results showed that the urban heat island effect in Xigang District is significant, with LSTs generally above 28°C at the end of August, mostly concentrated in the range of 38–40°C. The highest LST values were detected in northern port and harbor areas; the lowest LST values occurred in mountainous forest areas. The global Moran’s I value was 0.994, which was indicative of a very high positive correlation, and local Moran’s I values formed H-H and L-L type clusters concentrated in the northern harbor area and southern mountainous area, respectively. Finally, the GWR model could reflect the spatial heterogeneity of the relationships between LST and its driving factors well. Among these, in terms of natural physical factors, digital elevation model, normalized difference vegetation index, and modified normalized difference water index data were found to be negatively correlated with LSTs in most cases; in the social dimension, the point-of-interest number and building-coverage ratio were generally positively correlated with LSTs.

1. Introduction

Since the initiation of China’s economic reforms in the 1970s, the country has experienced rapid urbanization and numerous economic leaps, which have been accompanied by a surge in population growth, shortages of resources and space for living and commercial production, large-scale conversion of vegetation cover into impervious surfaces, and reductions in ecological space. The intensification of human activities and the consequential increase in industrial and residential energy consumption have led to the emergence of the urban heat island (UHI) effect and a decline in the quality of the thermal environment [1]. In view of frequent extremely high-temperature weather events in recent years, urban land surface temperatures influence both the health and comfort of urban residents and adversely destabilize ecosystems. The temperature also affects energy consumption, air quality, and vegetation phenology [2] and has far-reaching impacts on public health, urban settlements, and microclimates [3].

In the early 19th century, Howard et al. [4] began to observe differences in land surface temperatures (LSTs) in urban and suburban areas. Nowadays, thermal-infrared remote sensing technology has become an effective tool for the study of LSTs because of its ability to detect the thermal radiation emitted by ground objects. The measurement of changes in surface temperature has advanced from using observation data from meteorological stations to conducting multisource data analyses; this has alleviated the problem of limited coverage by meteorological stations. Many retrieval algorithms have been proposed for different sensors. Based on the required channels, these algorithms can be...
categorized into single-channel, split-window, and multi-channel [5–7]. Different remote sensing sensors are used at different scales; for example, low-resolution remote sensing images, such as those of the National Oceanic and Atmospheric Administration (NOAA) and Moderate Resolution Imaging Spectroradiometer (MODIS), are generally used for interprovincial and other large-scale retrievals; medium-high-resolution remote sensing data, such as from Landsat or SPOT (Satellite Pour l’Observation de la Terre), are often used for small-scale retrievals or detailed studies that require high-resolution imagery [8–10]. Studies have been conducted on daytime, nighttime, seasonal, and annual LST trends [11]; the effects of two-dimensional (2D) and three-dimensional (3D) urban morphology on LST [12, 13]; the effects of architectural forms and spaces on LST [14, 15]; changes in LST along urban–rural gradients [16]; the driving factors of LST; the impacts of the UHI effect on urban air circulation and vegetation phenology [17]; and the simulation of urban thermal environments by using models such as the Weather Research and Forecasting (WRF) Model and Urban Canopy Model (UCM) [18].

Presently, studies on the driving factors of LST generally focus on meteorological characteristics, landscape features, remote sensing spectral information, land use types, and urban morphology [19, 20]. Climatic factors such as wind speed, precipitation, sunshine duration, and climate all have an impact on LST, but these are not the root cause of the UHI effect. Reductions in natural features (e.g., water bodies and vegetation cover), the expansion of urban land (building density), and the intensification of human activities all accelerate heat accumulation and contribute to a significant increase in the LST [21, 22].

Extracting relevant indices from remote sensing images has become an effective tool in LST research. In urban systems, surface cover is mainly composed of vegetation, water bodies, and impervious surfaces. The normalized difference vegetation index (NDVI) is often used to extract green vegetation information [23]; the normalized difference moisture index (NDMI) or modified normalized difference water index (MNDWI) are used to extract data on water bodies [24]; and the normalized difference built-up index (NDBI) [25, 26] is used to extract data on impervious surfaces. In addition, some studies have used land use data to calculate transfer matrix and landscape index data to study the effects of urban morphology on LST. These studies have proposed that changes in land cover are the primary factor driving changes in LST and have suggested that urban planning can be used to alleviate the UHI effect [27–29]. In the social dimension, statistical data such as the gross domestic product (GDP), population size, and air quality are often used for research. Finally, regression and correlation analysis are useful tools that can be used to analyze the relationships between LST and its driving factors; however, to some extent, correlation coefficients are affected by the research scale [16].

Variables vary from location to location. Therefore, ordinary statistical analyses have limitations while carrying out the driving factor analysis on LSTs. In regard to the measure of spatial autocorrelation, the global autocorrelation approach was developed by Moran; later, some scholars proposed the local autocorrelation Moran index and Geary index [30]. Based on the idea of local smoothing, Brunsdon et al. [31] embedded the spatial position of data into regression parameters for the first time, and they proposed a geographically weighted regression model to describe the characteristics of regional nonstationary data; this was continuously improved upon afterwards [32]. Moreover, for the study of spatial heterogeneity, methods such as the semivariance function and spatial expansion model can also be used [33].

In this study, a multisource data analysis was conducted to investigate the impacts of natural physical and social factors on LST. The relationships between different driving factors and the spatial characteristics and heterogeneity of LST in Xigang District, located in the coastal city of Dalian, were analyzed. This study aims to provide an effective tool for urban planning and environmental management in order to alleviate problematic UHI effects.

2. Data and Research Methods

2.1. Study Area. Xigang District (38°57′–38°51′N and 121°34′–121°39′E) is one of the main urban areas of Dalian, Liaoning Province. The study area (Figure 1) covers seven streets within a built-up village area having a resident population of 300,000. Dalian is a well-developed and functionally diverse coastal city, with a monsoon-influenced humid climate. The air is humid in summer, with an annual average temperature fluctuation of 10°C and an annual maximum temperature of 35°C. The topography is characterized by higher elevations in a mountainous area to the south and lower elevations to the north. The driving factors of LST analyzed in this study change gradually and are representative of the study area.

2.2. Data Sources and Processing. This study was conducted by using Landsat 8 remote sensing imagery, point-of-interest (POI) data, building-structure data, meteorological data, and digital elevation model (DEM) data (Table 1). The POI and building-structure data were taken from Baidu’s Open Platform. After cropping the remote sensing image to fit the administrative boundary of the study area, we applied radiometric and atmospheric corrections; then, relevant indices and LST were extracted from the processed image. Considering that the spatial distribution of UHI does not show seasonal variations, but the UHI intensity tends to be higher in summer [34], and taking into account weather conditions during image capture, a remote sensing image captured at the end of August was selected for use.

Cities are a product of human construction activity in the natural environment; therefore, when considering the driving factors of LST changes, both social and natural factors should be selected and analyzed comprehensively. Consequently, the following five factors were selected for the analysis (Table 2): average DEM value, average NDVI value, average MNDWI value, POI number (NUMPOI), and building-coverage ratio (AV)
As the study area is coastal and within a maritime climate, it is not sufficient to consider whether a location is a water body or not; it is also necessary to consider the differences in humidity between different locations in the study area. For this reason, the MNDWI, which is more effective in urban areas, was used to extract moisture conditions [24]. Vegetation cover was extracted by using the NDVI, and terrain features were derived from a DEM.

For the social dimension, the effects of population density and human economic activity, as well as the number of man-made structures on the LST in the study area, were considered. Population density and economic activity of each grid-cell are reflected by its number of POI, and the proportion of man-made structures is reflected by \( A_V \). POI data are a type of point-like big data that describe real features. POI data can be used to record spatial and attribute

**Figure 1:** Location of the study area. (a) Digital elevation model (DEM) of the study region; (b) and (c) study area location within China and within Dalian, respectively.

**Table 1:** Data information.

<table>
<thead>
<tr>
<th>Type</th>
<th>Sensor and resolution (m)</th>
<th>Data source</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TIRS 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Street center</td>
<td>1:10000</td>
<td>Dalian land resources and housing bureau</td>
<td></td>
</tr>
<tr>
<td>POI</td>
<td>Baidu map (<a href="https://www.baidu.com">https://www.baidu.com</a>)</td>
<td></td>
<td>March 2019</td>
</tr>
<tr>
<td>Building data</td>
<td>Baidu map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meteorological data</td>
<td><a href="https://rp5.ru">https://rp5.ru</a></td>
<td></td>
<td>March 2019</td>
</tr>
<tr>
<td>DEM</td>
<td>30</td>
<td>China Academy of Sciences (<a href="http://www.gscloud.cn">http://www.gscloud.cn</a>)</td>
<td>28 August 2019</td>
</tr>
</tbody>
</table>

*DEM, digital elevation model; POI, point-of-interest number.*

As the study area is coastal and within a maritime climate, it is not sufficient to consider whether a location is a water body or not; it is also necessary to consider the differences in humidity between different locations in the study area. For this reason, the MNDWI, which is more effective in urban areas, was used to extract moisture conditions [24]. Vegetation cover was extracted by using the NDVI, and terrain features were derived from a DEM.
information of human activities, where POI spatial information can represent the physical locations of human activities on a microscale, and POI attribute information reflects, to a certain extent, economic and social activity [35, 36]. The number of man-made structures is represented by \( A_V \); \( A_Y \) can accurately reflect different factors over a period of time without the problem of multicollinearity between different factors.

### 2.3. Research Methods

#### 2.3.1. Surface Temperature Retrieval.

The highly specialized mono-window algorithm of Tan et al. [37] was used for LST retrieval from the Landsat image. This algorithm produces highly accurate results with very few parameters, and it incorporates surface and atmospheric effects directly into the algorithm, as shown in the following equations:

\[
T_s = \frac{a(1 - C - D) + \left[(b - 1)(1 - C - D) + 1\right]T_b - DT_a}{C} + 273.15, \tag{1}
\]

\[
C = \varepsilon r, \tag{2}
\]

\[
D = (1 - \varepsilon)[1 - (1 - \varepsilon)r], \tag{3}
\]

where \( T_s \) is the retrieved LST (°C); \( a \) and \( b \) are constants (based on the existing research, \( a = -67.355531 \) and \( b = 0.458606 \)); \( T_b \) is the brightness temperature (K); \( T_a \) is the effective mean atmospheric temperature (K); \( C \) and \( D \) are intermediate variables, which can be derived from the surface emissivity; \( \varepsilon \) is the atmospheric transmittance; and \( r \) represents the thermal-infrared band, which was calculated according to the atmospheric parameters on NASA’s official website.

#### 2.3.2. Moran’s I Index.

Based on the Third Law of Geography, the similarities and differences in spatial localization or stratification of different areas correspond to the similarities and differences of target variables in these areas [38]. Global Moran’s I quantitatively describes the spatial correlation of LST at a global range. Local Moran’s I [39] can be used to further explain clusters of LST values at a local range and to determine whether any LST values are abnormal, which provides the basis for analyzing the spatial heterogeneity of LST.

#### 2.3.3. Geographically Weighted Regression.

The geographically weighted regression (GWR) model allows for the analysis of data characterized by spatial nonstationarity and for exploring the spatial heterogeneity of the relationship between LST and its driving factors [40]. Based on the existing research [41] and actual conditions in the study area, we divided the study area into a 30 m × 30 m grid, and DEM, NDVI, MNDWI, NUMPOI, \( A_V \), and LST were extracted for each grid-cell. To improve the accuracy of the GWR model, the conventional ordinary linear least squares (OLS) model was used for preliminary tests of the driving factors. The variance inflation factor (VIF) test was performed to ensure that there was no multicollinearity between the variables. When VIF < 7.5, the GRW can be carried out by using the following equation:

\[
y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{n} \beta_n(u_i, v_i)x_{in} + \theta_i, \tag{4}
\]

where \( y_i \) is the observed variable; \( \beta_0(u_i, v_i) \) is the regression constant of the sample point at the \( (u_i, v_i) \) coordinate; \( \beta_n(u_i, v_i) \) is the regression parameter and is a function of the geographic location of variable \( n \) at the sample point; \( n \) is the number of factors; \( x_{in} \) is the value of the independent variable \( x_n \) at the sample point; and \( \theta_i \) is the random error.

### 3. Results

#### 3.1. Spatial Differentiation of LSTs.

Figure 2 shows the retrieved LST of the Xigang District. The maximum temperature difference was 15°C. It was evident that the UHI effect was significant in the study area, and the overall LSTs rose gradually from the south to the north, with the highest LST values occurring close to the harbor and the lowest LST values occurring in mountain forest areas. The LST was generally above 28°C; only 0.20% of the LST values fell below 28°C. The proportions were all between 10% and 20% in the range greater than 40 and that at 28–40°C divided into 2-degree increments. The majority of LST values fell within the range of 38–40°C and covered an area of 4,86.18 hectares, accounting for 19.10% of the land in Xigang District, and the
36–38°C range covered an area of 4,61.61 hectares, accounting for 11.66% of the land in Xigang District; these data indicate that the UHI effect is significant in the study area (Table 3).

3.2. Spatial Heterogeneity of LSTs. The univariate autocorrelation value of LST in the study area was 0.994, and the standard deviation was 4.139, which indicates an exceedingly high positive correlation. According to preliminary judgments, urban population flow, similarities of artificial material media in terms of its heat conduction properties, and influences of coastal city humidity made the data highly spatially autocorrelated. It was evident from Figure 3 that local spatial autocorrelation values formed an H-H (High-High) cluster in Xianglujiao Township, Zhanbei Township, and Beijing Township (northern Xigang), and an L-L (Low-Low) cluster was present in Bayi Township, Baiyun Township, and People’s Square Township (southern Xigang); these findings are consistent with the LST values, and no abnormal values were detected.

3.3. GWR Analysis of Driving Factors. The regression parameters of the LST established by OLS and factors are as follows (Table 4): the VIF of each factor was less than 3.500 (Table 5), so there was no redundant problem, but the interpretation of OLS on this problem was not very ideal, at just 0.706, and the residual spatial autocorrelation result was 0.77.

After eliminating collinearity by using the VIF test, a GWR model with an explanatory power of 89.50% was established (see Table 6 for model parameters). The residual spatial autocorrelation result was 0.77. The GWR model was used to obtain the regression coefficients of the driving factors in each grid (Figure 4). The statistical values (average, maximum, minimum, median, upper, and lower quartiles of each factor) are listed in Table 7. The coefficients of all five factors had both positive and negative symbols, thus demonstrating the spatial nonstationarity of the relationships between LST and the variables. In addition to NUMPOI reaching significance at the 0.05 level, the remaining four factors all reached significance at the 0.01 level.

It is evident from Figure 4 that the regression coefficient values for the DEM were concentrated between −0.05 and 0.00. In addition, in the northern harbor area and southern coastal edge area, the DEM coefficient values were greater than zero; coefficients were normally distributed and negatively correlated with the LST. It was found that, for the NDVI, with the exception of the maximum value, all other coefficient values were negative. The maximum and minimum values were extremely far apart, and the coefficient values were relatively scattered, but the positive values of the lowest interval were sparsely distributed and concentrated mainly around the pier in the north. The vegetation cover had the strongest inhibiting effect on the LST in the southeast and northwest; from there, the inhibiting effect dissipated, with some fluctuations, toward the north and central parts of the study area, where it mostly showed a negative correlation with LST. Similar to the NDVI, all of the MNDWI regression coefficient values were negative, with the exception of the maximum value, which was positive. The spatial distribution of the coefficient values was relatively scattered, with positive values distributed mostly in the central and southernmost areas of the study area; in the remaining areas, the correlation with the LST was negative. The POI coefficient values in the range of 0.0–0.5 were mostly concentrated in the central part of Xigang; other coefficient values were sparsely scattered throughout the study area. For the most part, POI numbers showed a positive correlation with the LST. The $A_V$ coefficient values were all positive with the exception of the minimum value. The coefficient values increased from west and south to north and east, most of them being positive, with some negative values distributed primarily in the west and south.

It was found that among the five factors considered in this study, most of the DEM, NDVI, and MNDWI values were negatively correlated with LST, whereas the $\text{NUMPOI}$ and $A_V$ values were positively correlated with LST. Based on the absolute values of the coefficients, it was found that, in terms of natural factors, the NDVI had the strongest effect on LST in the study area, followed by the MNDWI; the DEM had the weakest effect. In terms of social factors, the effect of $A_V$ on LST was stronger than that of $\text{NUMPOI}$.

4. Discussion

4.1. Improvements to the Model Accuracy. Multiple studies have explored LST driving factors by using correlation indices to represent the relationships between natural conditions and human activities. Among the adopted indices,
the NDBI [25], which is used for impervious surfaces, has often been found to have multicollinearity with the NDVI. Statistical data used to characterize the intensity of human activity, such as GDP, population size, and air quality index data, are difficult to apply spatially in urban areas with good accuracy; a problematic time lag is also common. The correlation coefficient is affected to a certain extent by the research scale [41]. Therefore, to improve the accuracy of the model, the research scale was reduced to 30 m × 30 m. In this study, the building-coverage ratio was selected to study the impact of man-made structures on the LST in order to eliminate the problem of multicollinearity. In addition, the POI number, which is characterized by strong temporal sensitivity and spatial continuous, was used to estimate the population density and economic activities in each grid-cell of the study area.

The selection of a regression model greatly affects the results of correlation analysis between different factors. For example, traditional regression models, such as the OLS, consider the overall space to be homogeneous and thereby ignore the spatial heterogeneity and complexity of driving factors [40]. In this study, the OLS model was used for pretesting. It was found that the explanatory power of the OLS model was 70.61%, and the residual spatial autocorrelation result was 0.77. The VIF test value was less than 3.5, which was within the acceptable range. The p values of all driving factors were statistically significant at less than 0.01. On this basis, the GWR was used to carry out a local regression analysis to establish the relationships between different regression parameters and their corresponding geographic locations and thus provide for a more accurate model of the relationships between the LST and different geographic factors. It was found that the GWR model was significantly better than the OLS model, with an explanatory power of 89.50%, and the residual spatial autocorrelation result was 0.67. The spatial autocorrelation value of the residual was still high, but it showed improvements compared to the OLS.

### 4.2. Study Limitations
This study examined the relationships between LST and natural and social driving factors; namely, the terrain topography, vegetation coverage, humidity, POI number, and building-coverage ratio. The building-coverage ratio was selected instead of the commonly used NDBI to eliminate the need for multispectral calculations and address the problem of multicollinearity between the NDBI and NDVI, with man-made structures and vegetation coexisting in the same space. However, one of the limitations of the building-coverage ratio is that it only accounts for building structures, while ignoring impervious surfaces such as roads and public squares; furthermore, it also does not reflect the effect of building height on the ventilation corridors [42] in a three-dimensional urban space [43, 44].

As the study area is located in a coastal city with no plains, but with numerous hills, and owing to its monsoon-influenced climate and sloping topography, there were significant differences in the LST between sunny and shady sides of hills, at the foots and tops of hills, and in coastal and urban-center areas. In this study, changes in LST were analyzed by using data collected in summer, when the UHI effect is the strongest, but without considering daytime and nighttime variations and seasonality [34]. The effects of fluctuations in the time series data [45] of each driving factor were also ignored.
5. Conclusions

In this study, the spatial distribution characteristics of LST in the Xigang District of Dalian, a typical coastal city in China, and its natural and social influencing factors were analyzed based on multivariate data, including remote sensing imagery, building-structure vectors, meteorological data, POI data, and DEM data. The results are as follows:

1. The urban heat island effect in Xigang is significant; LSTs in the study area increase gradually from south to north; LST temperatures at the end of August generally remain above 28°C, with most values concentrated in the range of 38°C to 40°C.

2. The global autocorrelation value was 0.994, which is indicative of a very high positive correlation; local autocorrelation values form H-H and L-L type clusters concentrated in the northern harbor area and southern mountainous area, respectively.

3. The GWR produced significantly better results than the OLS; regression coefficients had both positive and negative values and thus indicate spatial heterogeneity. The DEM, NDVI, and MNDWI were mostly negatively correlated with LST, whereas NUMPOI and Av were mostly positively correlated with LST. Among the natural factors, NDVI had the strongest effect on LST, followed by the MNDWI; the DEM had the weakest effect. Among the social factors, the effect of Av on LST was more significant than that of NUMPOI.

The issue of world climate change is a topic that has been discussed since the Industrial Revolution. From this article, it is evident that natural factors contribute significantly to...
suppress the increase in LSTs. In particular, it is necessary to construct public areas covered by vegetation, such as forested roadways, parks, roof gardens, and artificial wetlands, which can increase urban ecological sensitivity, promote ecological cycles, and fix carbon dioxide. In the social dimension, the increase in the proportion of artificial buildings will lead to increases in the surface temperature. We should, therefore, strive to find new building materials, build sponge cities, and reduce the heat absorption of impervious surfaces. The POI density was found to have a little effect on the surface temperature, but the CO₂ produced by human activities and life processes is the primary cause of atmospheric warming. Furthermore, for the sustainable progress of mankind, the development and use of new energy sources will play a particularly crucial role. This will remain an enduring challenge for the future.

Data Availability
Remote sensing imagery was acquired from USGS (https://glovis.usgs.gov). POI and building data were taken from the Baidu Map (https://www.baidu.com). Meteorological data were queried at RP5 (https://rp5.ru). DEM was downloaded from the China Academy of Sciences website (http://www.gscloud.cn).

Conflicts of Interest
There are no conflicts of interest to declare.

Acknowledgments
This research study was supported by the National Natural Science Foundation of China (Grants nos. 41771178, 41630749, and 41471140) and Liaoning Province Innovative Science Foundation of China (Grants nos. 41771178, 41630749, and 41471140) and Liaoning Province Innovative Talents Support Program (Grant no. LR2017017). The authors would like to thank all experts for their contributions to the urban thermal environment and urban planning research.

References


