

# Combining LuoJia1-01 Nighttime Light and Points-of-Interest Data for Fine Mapping of Population Spatialization Based on the Zonal Classification Method

Wei Guo , Jinyu Zhang , Xuesheng Zhao, Yongxing Li, Jinke Liu, Wenbin Sun, and Deqin Fan

**Abstract**—Fine-scale population spatial distribution plays an important role in urban microcosmic research, influencing infrastructure placement, emergency evacuation management, business decisions, and urban planning. In the past, nighttime light (NTL) data were generally used to map the spatial distribution of the population at a large scale because of their low spatial resolution. The new generation of LuoJia1-01 NTL data can be used for fine-scale social and economic analysis with its high spatial resolution and quantitative range. However, due to the geometry and background noise of the data themselves, the accuracy of the original NTL data is still low. Points-of-interest (POI) also can be used to map the population spatialization, but the indicative relationship between the POI and population is not clear, especially in rural and urban areas with different landscape structures. To solve the above-mentioned problems, this study proposes an improved nighttime light (INTL) index to better use the LuoJia1-01 NTL data. Meanwhile, a zonal classification model based on INTL and impervious surface area is proposed to distinguish urban and rural areas. Compared with previous research and existing datasets, our result had the highest accuracy ( $R^2 = 0.86$ ). This study explains that the INTL index is applicable to population spatialization research with the emergence of high-resolution and multispectral NTL satellite data. Moreover, the zonal classification model in this research can significantly improve the accuracy of population spatialization in rural areas. This study provides a possible way to use NTL and POI data in other social and economic spatialization research.

**Index Terms**—Geographically weighted regression (GWR), improved nighttime light (INTL) index, nighttime light (NTL), points-of-interest (POI), population spatialization.

Manuscript received 12 October 2022; revised 17 December 2022; accepted 17 January 2023. Date of publication 19 January 2023; date of current version 6 February 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 41930650, in part by the State Key Laboratory of Geo-Information Engineering and Key Laboratory of Surveying and Mapping Science and Geospatial Information Technology of MNR, CASM (2021-03-04), and in part by The Ningxia Hui Autonomous Region Key Research and Development Project under Grant 2022BEG03064. (Corresponding author: Jinyu Zhang.)

Wei Guo is with the School of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China, and also with the Chinese Academy of Surveying and Mapping, Beijing 100830, China (e-mail: weiguo@cumtb.edu.cn).

Jinyu Zhang, Xuesheng Zhao, Yongxing Li, Jinke Liu, Wenbin Sun, and Deqin Fan are with the School of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China (e-mail: zhangjinyu9908@163.com; zxs@cumtb.edu.cn; liyx\_rs@163.com; liujinkers@163.com; swb@cumtb.edu.cn; deqinfan@cumtb.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3238188

## I. INTRODUCTION

POPULATION data are the basic data sources reflecting human activities and are widely used in ecological and environmental protection, disaster risk assessment, business decision making, and regional planning and development [1], [2], [3], [4]. The census, based on the administrative area as the unit of statistics, is the major way of obtaining population information in most countries. It is authoritative, systematic, and normative [5], [6]. However, the boundaries of the administrative areas are inconsistent with those of natural units (watersheds, natural hazards affected area, etc.) in actual research, and the differences make it difficult to integrate and comprehensive analyze multisource data at a fine scale [7]. Gridded population distribution data can solve the above-mentioned problems [8], [9], [10].

Many methods have been proposed to disaggregate population data from administrative to grid units [11], [12]. One of the most common methods is dasymetric mapping, whose main idea is to establish models between auxiliary data and population data to generate weight layers and disaggregate statistical population data according to the weight layers [13], [14]. Meanwhile, remote sensing data are becoming important auxiliary data because of their large coverage, rapid updating, and public availability [15], [16], [17]. Nighttime light (NTL) data as useful remote sensing data with spatial morphology information and intensity information have been proved to be significantly correlated with large areas of population distribution and have been widely used in generating gridded population distribution [18], [19]. However, the Defense Meteorological Satellite Program and Operational Linescan System (DMSP-OLS) NTL images have some inherent drawbacks and quality problems, such as blooming, saturation, and coarse resolution ( $\sim 1$  km) [20], [21], [22], [23]. The Suomi National Polar-orbiting Partnership (Suomi-NPP) was launched in October 2011, which acquired Visible Infrared Imaging Radiometer Suite Day/Night Band (VIIRS/DNB) NTL data is considered to be the successor product to DMSP/OLS. There is no saturation of VIIRS/DNB NTL data and its spatial resolution ( $\sim 750$  m) has been significantly improved. Compared with DMSP/OLS, VIIRS/DNB NTL data has become much more effective in reflecting social and economic activities. Nevertheless, the moderate spatial resolution

still limits the ability of VIIRS/DNB NTL data to capture urban detail, which restricts the mapping of population spatial distribution at fine scale [24], [25]. Launched in 2018, the professional NTL satellite LuoJia1-01 can provide remote sensing images with a spatial resolution of 130 m, which can convey more detailed structural information within cities [26], [27]. Furthermore, it does not have a saturation effect, and the “blooming” effect is reduced due to the increase in spatial resolution. Therefore, it has a more accurate description of the extent and intensity of human activity and has greater advantages in social and economic research. Previous studies have mostly focused on the direct use of LuoJia1-01 NTL images [23], [28]. However, several papers pointed out that when high quantification level NTL images are used, the contrast of the low and high areas in the image is enhanced [18], [29], [30]. The views suggest that direct use of LuoJia1-01 NTL images would result in overestimation of the population in high light areas and underestimation of the population in low light areas. Therefore, the accuracy of the population spatialization is not high when directly using LuoJia1-01 NTL data, and the data quality needs to be improved. In this context, one of the contributions of this study is to propose an improved nighttime light (INTL) index to improve the accuracy of the population spatialization.

On the other hand, the geospatial Big Data with specific geographical coordinates and semantic information can be a useful supplement to remote sensing data [4], [31], [32]. As a representative of geospatial Big Data, points-of-interest (POI) data have the characteristics of large volume, wide variety, and high positioning accuracy [33], [34], [35]. Yang et al. [36] conducted experiments in Zhejiang and concluded that the combination of POI and multisource remote sensing data significantly improved the accuracy of the population spatialization. However, the relationship between the POI and population may vary from one region to another because different regions or environmental conditions have different spatial structures and population aggregation patterns [19]. For example, obvious differences exist in architectural patterns and population distribution between urban and rural areas. In this research, we proposed a threshold based on the INTL and the impervious surface area (ISA) index to effectively identify areas with large structural differences. We performed zonal classification based on the above-mentioned index and established different models.

In summary, the main objective of this study was to improve the accuracy of the population estimation. We established INTL index to make better use of the LuoJia1-01 NTL images. At the same time, we used a zonal classification model to ensure the POI had a better population interpretation effect. In this research, we constructed a geographically weighted regression (GWR) to obtain the population from street scale to grid level based on the above-mentioned model. This result not only provided a fine scale population distribution but could also benefit other social and economic spatialization research.

## II. STUDY AREA AND DATASETS

### A. Study Area

Beijing is a world-famous ancient capital and a modern international city [37]. By 2020, the city is expected to cover

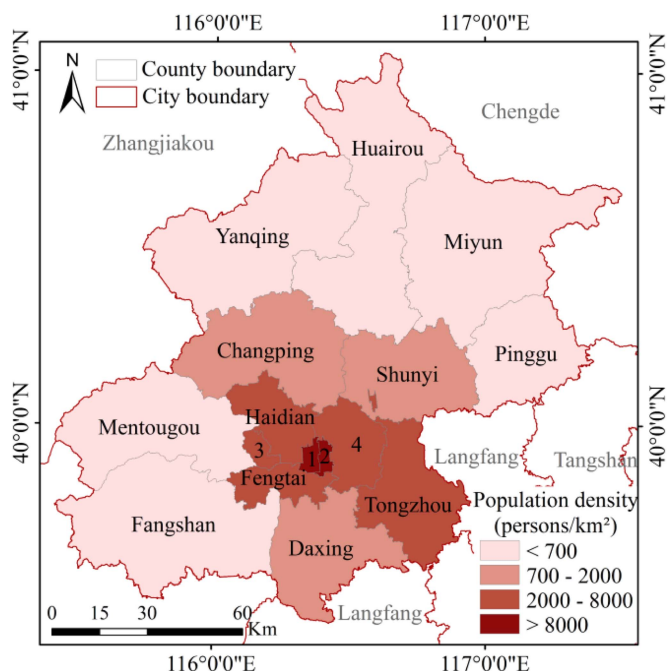


Fig. 1. Population density of districts in Beijing. 1, 2, 3, and 4 represent Xicheng, Dongcheng, Shijingshan, and Chaoyang, respectively.

an area of 16410.54 km<sup>2</sup> (Fig. 1). The resident population of Beijing is approximately 21893100 (approximately 1344 persons/km<sup>2</sup>). The level of economic development and social functions of each region in Beijing are different, resulting in varying population densities. For example, the Xicheng district has the highest population density of 21 818 persons/km<sup>2</sup>, while the Yanqing district has the lowest density of 173 persons/km<sup>2</sup>. The population distribution of Beijing has spatial heterogeneity and complex characteristics. Therefore, selecting Beijing as the study area has reference significance for other large cities with a very large population, strong spatial heterogeneity, and rapid development in terms of population spatialization.

### B. Datasets

The data in this article are divided into remote sensing data, geospatial Big Data, and validation data. Table I lists all the data used in this study, including the remote sensing image data (LuoJia1-01 NTL, Landsat 8), the geospatial Big Data (POIs, OpenStreetMap data), the census data, and the validation data (WorldPop and LandScan). The LuoJia1-01 NTL satellite provided images with a spatial resolution of 130 m, which can more accurately represent the spatial information inside cities and has irreplaceable advantages in spatial scale analysis. The Landsat 8 satellite has a 16-day revisit period. The Operational Land Imager (OLI) includes nine bands with a spatial resolution of 30 m, including a panchromatic band with a spatial resolution of 15 m [38]. Two cloud-free images of Landsat 8 from January 2020 were selected in this study to extract ISA information. The POIs for this study were obtained from the Amap service, which is the most used and largest online mapping service in China. We obtained more than 3 00 000 records in the study area, including catering services, public facilities, companies,

TABLE I  
DATA SOURCE DESCRIPTION

Data	Description	Source	Year
Luojia1-01	130-m spatial resolution; a representation of the intensity of human activity	Hubei Provincial Data and Application Centre ( <a href="http://59.175.109.173:8888">http://59.175.109.173:8888</a> )	2018
Landsat 8	30-m spatial resolution; without clouds or less than 1% of clouds	US Geological Survey ( <a href="http://www.gscloud.cn/">http://www.gscloud.cn/</a> )	2020
Points-of-interest	A good reflection of human activities at the micro scale	Amap service, China ( <a href="https://www.amap.com/">https://www.amap.com/</a> )	2020
OpenStreetMap data	Volunteered geographic information, used to eliminate non residential areas	OpenStreetMap Contributors ( <a href="https://www.openstreetmap.org/">https://www.openstreetmap.org/</a> )	2020
Population data	Seventh national census of China; district level and subdistrict level	Beijing Municipal Bureau Statistics ( <a href="http://tjj.beijing.gov.cn/">http://tjj.beijing.gov.cn/</a> )	2020
WorldPop	9 arc-seconds spatial resolution	WorldPop dataset for Beijing ( <a href="https://www.worldpop.org/">https://www.worldpop.org/</a> )	2020
LandScan	1 km spatial resolution	LandScan dataset for Beijing ( <a href="https://landscan.ornl.gov/">https://landscan.ornl.gov/</a> )	2020

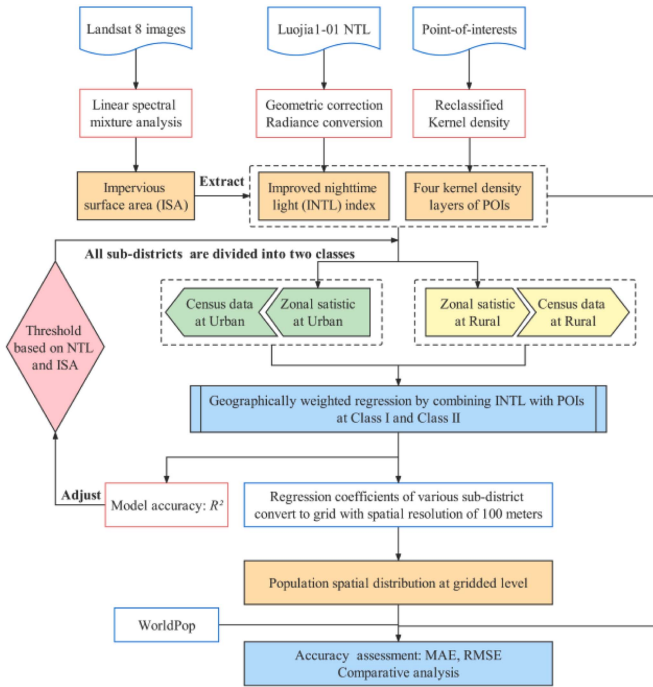


Fig. 2. Framework for mapping the population distribution of Beijing.

transportation facilities, commercial institutions, commercial spaces, scientific, educational and cultural services, medical care services, accommodation services, etc. OpenStreetMap (OSM) is a type of volunteered geographic information. After years of accumulation, OSM now has detailed geographic information for most regions of the world. This study used OSM data to eliminate the nonresidential areas, such as roads and scenic spot information.

### III. METHOD

The target of the study was to generate an accurate population distribution map with a spatial resolution of 100 m. Fig. 2 illustrates the flowchart of the entire study. The following four steps were taken:

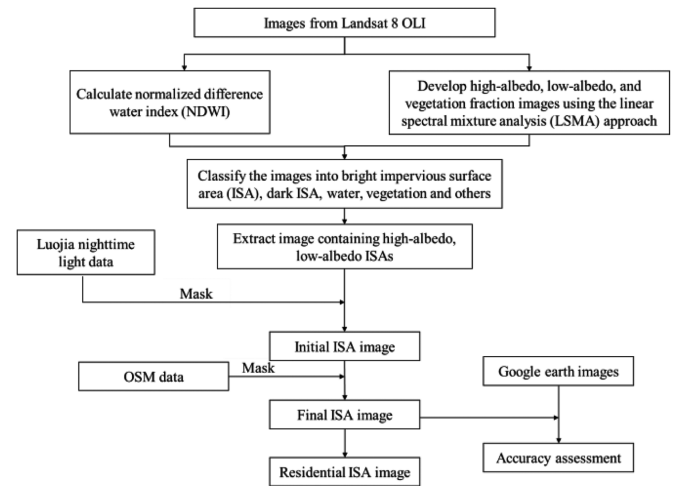


Fig. 3. Framework for obtaining the residential ISA.

- 1) Data preprocessing, which included extracting the ISA, preprocessing the Luojia1-01 NTL images, and processing the POI data.
- 2) Based on the NTL and ISA, we obtained the results of the zonal classification.
- 3) We constructed regression models using GWR with the INTL and POI as independent variables and the census data as the dependent variable. The population was disaggregated into grids, according to the regression coefficients.
- 4) The accuracy of our model was assessed.

#### A. Extracting the Residential ISA From the Landsat 8 Images

The ISA represents major sites of human activities. There are already many methods to accurately extract the ISA using Landsat [39], [40], [41], [42]. Fig. 3 shows the framework for extracting the residential ISA.

In this study, we used the linear spectral mixture analysis method to analyze the images, and three fraction images were generated, representing different endmembers [39]. The three



endmembers were high-albedo objects, low-albedo objects, and vegetation, respectively. We generated high-albedo and low-albedo images according to the selected endmembers. The ISA exists in both high-albedo and low-albedo images. At the same time, shallow water, wet soil, and impurities were also assigned to the two images. Therefore, subsequent removal was necessary. We used the normalized difference water index to remove the water [43]. The NTL was used to remove the remaining non-ISA, such as landslides, poor vegetation coverage, and water, which can easily be misclassified as ISA [44]. The roads, squares, parking lots, airports, etc., were classified as ISA, but they are not places where resident populations live. Therefore, in order to estimate the population more accurately, these areas in the ISA needed to be removed via the OSM data [8]. Finally, we used Google Earth images to evaluate the accuracy of the ISA. The verification samples were selected using the random sampling method. A total of 600 samples were selected, including 400 non-ISA samples and 200 ISA samples. The accuracy was evaluated using an error matrix [45], [46].

### B. Luojia1-01 Preprocessing and the Development of an INTL Index

The Luojia1-01 satellite, launched in June 2018, provided images with a spatial resolution of 130 m. The data are freely available from the Hubei Provincial Data and Application Centre (<http://59.175.109.173:8888>). We selected the cloud-free image on 6 September 2018 and the image was a single-phase image. According to the information provided by the open download website, we converted the digital numbers (DNs) to radiance values after absolute radiation correction [26], [27]. Compared with the DMSP-OLS NTL data (6 bit), Luojia1-01 NTL data (15 bit) have a higher quantification, which creates a large difference between the maximum and minimum values. The Luojia1-01 NTL data were the system's geometric correction product, and the positioning accuracy was relatively low [47]. Moreover, there was still residual background noise in areas such as forests and water bodies. These phenomena will have a negative impact on data processing and are not suitable for direct population spatial research. Therefore, we processed the Luojia1-01 images as follows:

- 1) *Reprojection and resampling*: we reprojected the extracted images to the WGS 1984 UTM zone 50N. Next, the nearest neighbor method was used for resampling the images to a resolution of 100 m.
- 2) *Geometric correction*: taking the Landsat 8 image as the reference image, by selecting several evenly distributed road intersections on the image, geometric correction processing was carried out to keep the image offsetting within one cell size.
- 3) After absolute radiation correction using (1), we converted the DN values to radiance values.

$$L = DN^{\frac{3}{2}} \cdot 10^{-10} \quad (1)$$

where  $DN$  is the brightness value of the pixel, and  $L$  is the radiant value, the unit of which is  $W/(m^2 \cdot sr \cdot \mu m)$ .

- 4) *Denoise*: in order to remove background noise, we used (2) to unify the NTL data. Subsequently, there were still some extremely high values. In order to reduce their influence, we used the threshold method to remove the background noise and the extremely high values. Here, we set the maximum threshold to 900, and the minimum threshold was set to 0.92 [27]

$$L' = L \cdot \omega \cdot 10^5 \quad (2)$$

where  $\omega$  is the bandwidth; the range is 460~980 nm;  $\omega = 0.52 \mu m$ ; and  $L'$  is the radiance after the unified dimension.

It is worth noting that in some human residential areas, such as residential quarters, shanty towns, and especially rural resident areas, the light sources mainly come from indoor lighting. Compared with areas that directly emit light, the radiance value in the residential area is low; hence, the intensity of human activities in these areas could be underestimated, resulting in the inaccuracy of the population distribution [48]. Therefore, we proposed an INTL index to enrich the spatial variation information of these areas, which can reflect the population activity level more effectively. In this study, we chose logarithms and square roots to smooth some of the extremely high brightness values. We added a value of 1 in the formula to keep the value positive.

The INTL can be demonstrated using the following formula:

$$INTL = \ln \left( 1 + \sqrt{NTL} \right) \quad (3)$$

The NTL is the radiance value of the Luojia1-01 NTL data after preprocessing.

### C. Kernel Density Processing for the POI Data

The POIs in our study were obtained from the Amap service (<https://www.amap.com>), which is the most used and largest online mapping service in China. In 2020, there were more than 3 00 000 records in the study area, covering 23 categories. The POI has the characteristic of reflecting human activities at the micro scale, which makes it possible to describe more fine-grained land use types.

First, data cleaning was carried out to remove incomplete and repeated data from the POIs, and 9 03 544 effective POI records were retained. Based on previous studies and the actual condition of the study area [3], [49], we reclassified these POIs into four categories, as shown in Table II. These four categories were based on convenient transportation, complete living facilities, and complete services. POIs are point data with discrete and discontinuous characteristics; the kernel density method was used to transform the discrete POI into smooth and continuous density raster data. After kernel density processing, the POI reflected the spatial aggregation state quantitatively, which is convenient for statistical analysis. This study selected the bandwidth of 1.5 km to generate the POI density layers with a 100 m spatial resolution.

The Luojia1-01 NTL, the residential ISA, and the four categories of POI density layers were converted to the WGS-84 coordinate system and projected into the Universal Transverse Mercator coordinate system. Then, the nearest neighbor method was used to resample the images to a resolution of 100 m. In the process of population spatial distribution, the principle is that

TABLE II  
 CATEGORIES OF POIS

Reclassified categories	POI categories
Residence	Accommodation Service, Commercial House
Institution	Governmental Organization & Social Group, Medical Service, Science/Culture & Education Service, Finance & Insurance Service
Service	Food & Beverages, Shopping, Sports & Recreation
Transportation	Pass Facilities, Place Name & Address, Road Furniture, Transportation Service

the population should be distributed in human settlements. In this study, we used the ISA to extract the NTL and POI density layers.

#### D. Threshold Based on the INTL and ISA Index

Due to the spatial heterogeneity between urban and rural areas, the performance of modeling factors shows discrepancy; the same value is different in terms of the population distribution and human activity intensity, which leads to the low accuracy of the model. Therefore, we proposed a threshold based on the INTL and ISA index (TBNII), representing the intensity of the INTL in subdistricts, to classify each subdistrict into urban and rural areas. The formula for the TBNII is as follows:

$$TBNII_i = \left( \frac{TNL_i}{ISA_i} \right) \bigg/ \left( \frac{TNL}{ISA} \right) \quad (4)$$

where  $TNL_i$  represents the total value of INTL in the  $i$ th subdistrict, and  $ISA_i$  represents the total area of the ISA in the  $i$ th subdistrict. The higher the TBNII, the greater the activity intensity of the subdistrict.

#### E. Constructing the Population Model

According to previous studies, the distribution of the population with spatial nonstationarity leads to poor accuracy of those models. GWR is an effective way to examine spatial nonstationarity [50], [51]. Compared with global linear regression, the GWR model takes into account the spatial location of data in the regression process [52], [53]. Only data within a certain distance of the regression point are considered, and the closer the distance, the greater the effect on the regression point. Therefore, there are different regression coefficients at each regression point, which means that each regression point has an independent regression equation.

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^n x_{ij} \beta_j(u_i, v_i) + \varepsilon_i \quad (5)$$

$(i = 1, 2, \dots, m, j = 1, 2, \dots, n)$

where  $i$  illustrates the  $i$ th subdistrict;  $m$  is the number of subdistricts participating in the regression;  $j$  illustrates the  $j$ th independent variable;  $n$  is the number of independent variables;

$(u_i, v_i)$  indicates the spatial geographic center coordinate of the  $i$ th subdistrict;  $\beta_0(u_i, v_i)$  represents the constant term;  $x_{ij}$  illustrates the  $j$ th variable in the  $i$ th subdistrict;  $\varepsilon_i$  illustrates the error term of the  $i$ th subdistrict,  $\varepsilon_i \sim N(0, \sigma^2)$ ; and  $\beta_j(u_i, v_i)$  represents the local regression coefficients. The regression coefficients are estimated by the following equation:

$$\hat{\beta}_j(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} (X^T W(u_i, v_i) Y) \quad (6)$$

where  $\hat{\beta}_j(u_i, v_i)$  is the estimation of  $\beta_j(u_i, v_i)$ ;  $X$  illustrates a vector of independent variables;  $Y$  illustrates a vector of the dependent variables; and  $W(u_i, v_i)$  is a spatial weight matrix, whose diagonal elements are  $w_{lc}$  [54]. In this study, the matrix was calculated by Gauss's function, represented by the following equation:

$$w_{lc} = \exp \left( - \left( \frac{d_{lc}}{b} \right)^2 \right) \quad (7)$$

where  $d_{lc}$  illustrates the Euclidian distance of the subdistrict geographic center between  $l$  and  $c$ . With the change in  $d_{lc}$ ,  $w_{lc}$  varies. That is, the closer the distance, the higher the weight.  $b$  illustrates the bandwidth, which determines the maximum distance from the  $i$ th region center. Different bandwidths have different weight decay rates. With the increase in distance, the weight with a large bandwidth decays slowly, while the weight with a small bandwidth decays quickly. In this study, the adaptive bandwidth was adopted to determine the number of nearby subdistricts. Cross validation was used to optimize the bandwidth [55].

The population GWR model can be expressed as follows:

$$\begin{aligned} POP_i = & \beta_0(u_i, v_i) + INTL_i \cdot \beta_1(u_i, v_i) \\ & + RE_i \cdot \beta_2(u_i, v_i) + IN_i \cdot \beta_3(u_i, v_i) \\ & + SE_i \cdot \beta_4(u_i, v_i) + TR_i \cdot \beta_5(u_i, v_i) + \varepsilon_i \end{aligned} \quad (8)$$

where  $POP_i$  illustrates the  $i$ th subdistrict;  $RE_i$  is the value of residences in the  $i$ th subdistrict;  $IN_i$  is the value of institutions in the  $i$ th subdistrict;  $SE_i$  is the value of services in the  $i$ th subdistrict; and  $TR_i$  is the value of transportation in the  $i$ th subdistrict.

Therefore, we obtained the regression coefficients for independent variables on each street, which are  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  in (8).

Disaggregating population from subdistrict-level to grid-level is based on the following assumption: the modeled relationship between population and independent variables derived from district-level data retains at grid-level.

The trained model and regression coefficients were used to redistribute the total population. We calculate the population in each grid with 100 m resolution by multiplying coefficients with the values of the corresponding independent variables at each grid.

#### F. Accuracy Assessment

The coefficient of determination ( $R^2$ ), mean absolute error (MAE), and root mean square error (RMSE) were used to assess

TABLE III  
EVALUATION OF THE EXTRACTION ACCURACY OF THE RESIDENTIAL ISA

Category	Reference data		Precision		Overall accuracy
	ISA	Others	User's accuracy	Producer's accuracy	
ISA	172	28	86%	90.05%	92.17%
Others	381	19	95.25%	93.15%	

the accuracy of different models and compare the predicted values with the subdistrict census data. The subdistrict level accuracy was assessed using China's Seventh Census data as the reference data

$$R^2 = 1 - \frac{\sum_{i=1}^n (POP_i^{estimated} - POP_i^{census})^2}{\sum_{i=1}^n (POP_i^{estimated} - \overline{POP}_i)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (POP_i^{estimated} - POP_i^{census})^2}{n}} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |POP_i^{estimated} - POP_i^{census}| \quad (11)$$

where  $n$  is the total number of subdistrict units;  $POP_i^{estimated}$  represents the estimated data of the subdistrict;  $POP_i^{census}$  denotes the census data of the subdistrict; and  $\overline{POP}_i$  denotes the average value of the estimated data.

Meanwhile, to highlight that our new population map holds a competitive accuracy, we introduced WorldPop and LandScan for validation.

#### IV. RESULTS

##### A. Residential ISA Distribution Analysis and Accuracy Evaluation

The ISA is a key factor of land use and land cover, which is representative of the building footprint and the human settlement. However, the ISA information contains other areas such as roads and parking lots [2], [8]. To obtain more accurate data, we trimmed the refined ISA to obtain the residential ISA by excluding roads, parking lots, and other nonresidential data, which were obtained from OSM. It is important to extract the ISA accurately. As shown in Table III, the accuracy of the ISA was good with an overall accuracy of 92.17%. The results of the residential ISA are shown in Fig. 4.

The fractional ISA has a value from 0 to 1, representing the percentage of land occupied, and was also used as a variable in this study. The ISA was abundant in the relatively flat central, southern, and eastern parts of Beijing. In the mountainous areas of the north and west, the distribution of ISAs was sparse.

##### B. Evaluation of the INTL and NTL

Fig. 5 presents the radiation value variations of the NTL and INTL along a latitude transect in Beijing. Fig. 5(a) presents the

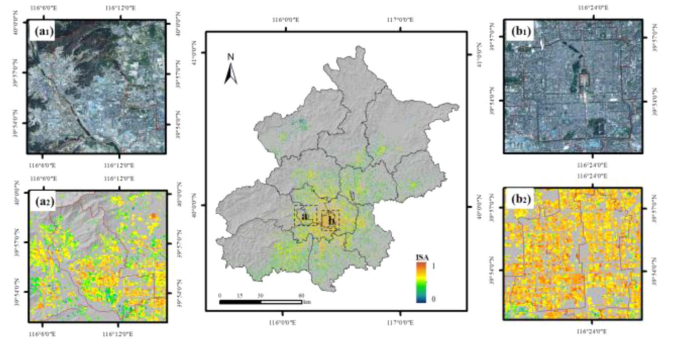


Fig. 4. Residential ISA distribution in Beijing: (a<sub>1</sub>) and (a<sub>2</sub>) are the 30-m spatial resolution color synthetic Landsat 8 OLI image and ISA distribution map of Shijingshan district, respectively; (b<sub>1</sub>) and (b<sub>2</sub>) are the 30-m spatial resolution color synthetic Landsat 8 OLI image and ISA distribution map of Dongcheng district and Xicheng district, respectively.

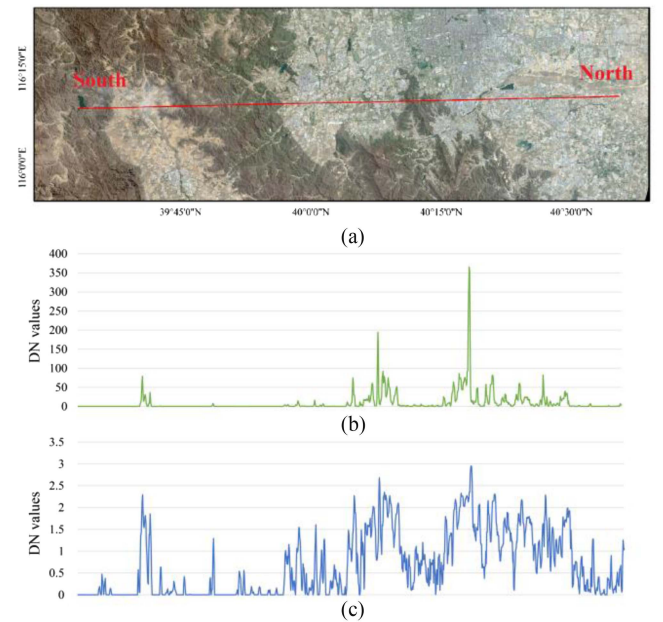


Fig. 5. Radiation value variations of (a) Landsat image, (b) LuoJia1-01 NTL, and (c) INTL along a latitude transect in Beijing.

large difference between the low regions and the high regions (the highest was close to 400). Because of the large difference in the LuoJia1-01 NTL data values, the urban and the rural/suburb population distribution was not well represented. Fig. 5(b) shows that the INTL effectively smoothed the extremely high values and improved the low values. It also better distinguished the low-value area from the nonvalue area and reduced the large difference between the high and low values. Thus, the INTL significantly improved the accuracy of the population distribution.

##### C. Results of the Zonal Classification Based on TBNII

We backward the optimal value of TBNII from the accuracy of population spatialization. If  $R^2$  is higher, it means that the model is less confusing, which represents the best differentiation effect of population aggregation pattern. That is, the discrimination



TABLE IV  
COMPARISON OF THE MAE AND RMSE

Model	Main data source	Main thoughts	$R^2$	MAE ( $10^4$ )	RMSE ( $10^4$ )
M1(LandScan)	NTLs, roads, slope, elevation, land cover	Multivariate dasymetric mapping	0.359	3.29	5.09
M2(WorldPop)	Roads, land cover, urban areas, NTLs, infrastructure	Semiautomated dasymetric mapping	0.470	3.55	6.23
M3(Others)	NTL, POI	Directly using Luojia1-01 NTL data	0.708	2.12	3.12
M4	INTL, POI	An Improved NTL index	0.766	1.81	2.80
M5(Ours)	INTL, POI	Zone classification strategy	0.858	1.64	2.30

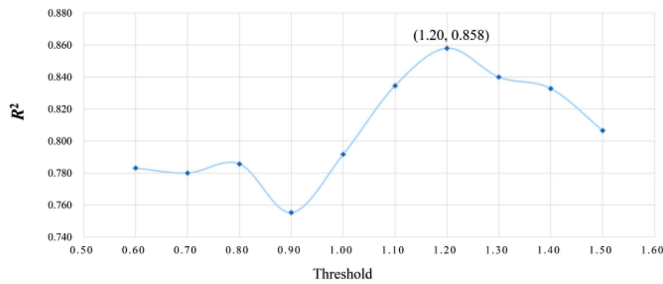


Fig. 6. Model accuracy ( $R^2$ ) under different TBNII values. The optimal threshold for the model to reach the highest accuracy (0.858) was 1.2.

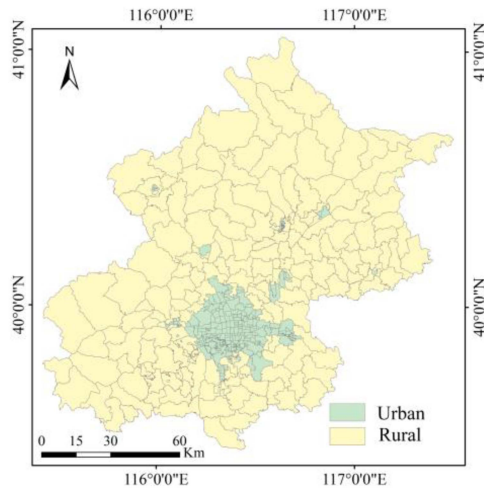


Fig. 7. Results of the zonal classification when the TBNII was 1.2.

between rural and urban works is better. In this study, 0.1 was used as the interval unit to traverse different thresholds; the  $R^2$  fluctuated significantly with the adjustment of the threshold. As shown in Fig. 6, the fluctuating  $R^2$  first increased and then decreased. When the TBNII value was 1.2, the model had the highest accuracy with an  $R^2$  at 0.858. Thus, we used 1.2 as the threshold to classify the urban (TBNII>1.2) and rural (TBNII <1.2) areas. The results of the regional classification based on the TBNII are shown in Fig. 7, the central urban area was transformed into one class and the surrounding areas into the other class, which was consistent with the development of Beijing.

#### D. Model Comparison and Accuracy Assessment

M1 and M2 are two commonly used products of population spatial distribution. M3 represents the model combined original Luojia1-01 NTL data with POI data. M4 represents the improved model, which performs logarithm and square root processing according to the characteristics of Luojia1-01 NTL data. M5 represents the model proposed in this study.

Table IV shows the accuracy comparison of population mapping models. Overall, M5 had the highest adjusted R-squared and the lowest MAE and RMSE. The results showed that the INTL index improved the model accuracy, and the zoning strategy also significantly improved the model accuracy. Meanwhile, the above-mentioned three models were all better than WorldPop and LandScan in terms of the population spatialization.

Fig. 8 shows the scatterplots between the estimated population and the census data at the subdistrict level of Beijing. M5 had the minimum deviation between the estimated population and the census data in urban areas (i.e., subdistricts with a population greater than  $20 \times 10^4$ ). In the rural areas with a smaller population, M5 also performed best.

There was a mismatch between Luojia1-01 NTL data and other data of the time-scale, because we had to use the night light data from 2018. Nevertheless, in the population spatialization model proposed by the study, the Luojia1-01 NTL data was not a single independent variable and was only used as a weighting layer for redistribute the total population, which reduced the influence of the temporal difference to a certain extent.

#### E. Mapping Results of the Population Spatial Distribution

The spatial distribution result of the gridded population in Beijing in 2020 is shown in Fig. 9. The grid-level population data better expressed the spatial heterogeneity of the population distribution. The population was mainly concentrated in the central urban areas. The spatial distribution of the population in Beijing was uneven. The population density gradually decreased from the central city to the surrounding areas. The central, eastern, and southern regions were flat places with dense populations, while the western and northern mountainous areas were sparsely populated, determined by the characteristics of Beijing’s natural environment and socioeconomic development. Compared with the central, southern, and eastern plains, the mountainous areas in the west and north are not suitable for large populations to inhabit.

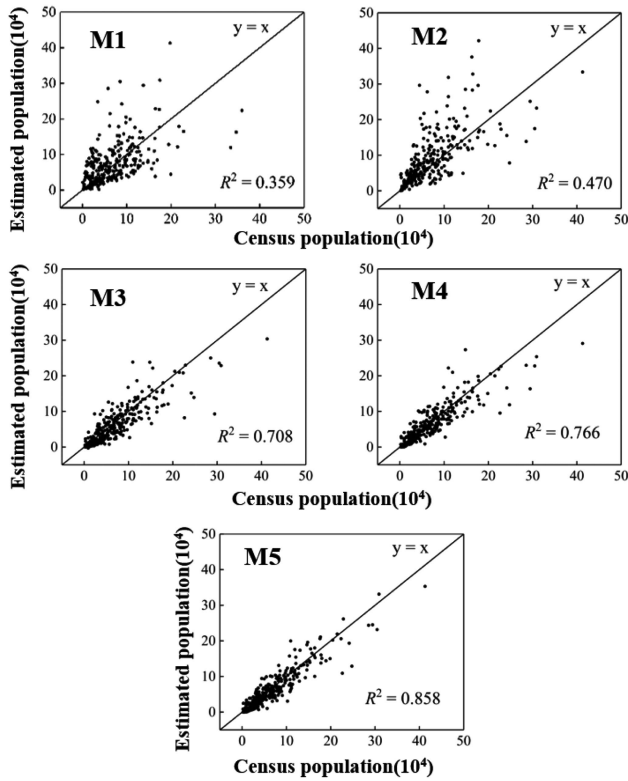


Fig. 8. Scatterplots between the estimated and census data from the different models. Each point denotes a subdistrict in Beijing. (a) M1. (b) M2. (c) M3. (d) M4. (e) M5.

## V. DISCUSSION

### A. Efficacy and Potential of the INTL

The Luojia1-01 NTL performs better when reflecting social and economic activities because of its high spatial resolution, no saturation, and lower blooming effect. At present, many studies have used Luojia1-01 NTL images to analyze socioeconomic parameters such as the gross domestic product (GDP) [56], urban activities [47], poverty estimation [57], and urban transformation [58]. However, there are still some limitations in Luojia1-01 NTL, such as the existence of noise background, which creates a large difference between the extremely high and low values. Therefore, we needed to process the extremely high values and background noise to enable the Luojia1-01 NTL image to perform better.

To accomplish this goal, we used the square root and logarithm transformation to suppress the high values and enhance the low values in the Luojia1-01 NTL, which we called the INTL index. This index better represented the spatial distribution of the population. As indicated in Table V, the model based on the INTL had a higher accuracy than the model based on the NTL. The INTL has two main purposes. The first purpose is to enhance the low nightlight information, especially in rural areas. INTL can provide richer spatial variations in population distribution between rural and urban areas. The second purpose is to suppress the extremely high values, especially those values

TABLE V  
CORRELATION OF THE INTL AND NTL WITH THE POPULATION

Category	NTL	INTL
Rural	0.77	0.85
Urban	0.48	0.64
All subdistricts	0.66	0.69

in urban commercial centers. If the original Luojia1-01 NTL data are directly used to estimate the spatial distribution of population, it tends to distribute the population excessively into the area with extremely high NTL values. Therefore, INTL can alleviate the overestimating problem in some regions.

We used skewness to analyze the validity of the INTL, and a positive bias means that most values are to the right of the average. The skewness of the NTL and INTL were 10.99 and 0.52, respectively. The INTL greatly alleviated the deviation of the NTL, indicating that the INTL enlarged the brightness variation degree of the NTL and enriched the spatial variation information. To verify the superiority of the INTL, we analyzed the correlation of the INTL and NTL with the population. As shown in Table V, the correlation between the INTL and the population was higher than that of the NTL in all regions, and the increments in urban and rural areas were 0.16 and 0.08, respectively.

The excellent performance of the INTL was verified in this study. It can be extended to other social and economic research, including the extraction of built-up areas and GDP assessment. It certainly provides a new data processing method, which can be used as a reference for other NTL satellite data, such as Jilin1-07 [49], EROS-B [18], and SDGSAT-1 [59].

### B. Importance of the Zone Strategy

At present, the POI is the most widely used geospatial Big Data for population spatialization [32]. Previous studies have pointed out that POI contributes significantly to fine-scale population distribution [35]. However, few studies have noted that the regression relationship between the population and the POI may be different due to the variance in urban structures in different regions, although this phenomenon has been put forward in the study of population spatialization using NTL data [19], [51], [60], [61]. In rural and urban areas with different spatial structures, the same NTL data represent different populations. The TBNII used in our study reflected regional differences according to the average NTL data of the residential ISA. It is assumed that the higher the TBNII, the greater intensity of social and economic activities. Table IV shows that the accuracy of the zonal classification model was significantly improved, indicating that adopting a zone strategy was necessary. Considering the strong spatial heterogeneity in urban and rural areas, building the relationship between independent variables and dependent variables in one regression model is not suitable. Therefore, when the NTL data and POI are used as representative variables, the areas with different structures should be modeled separately.



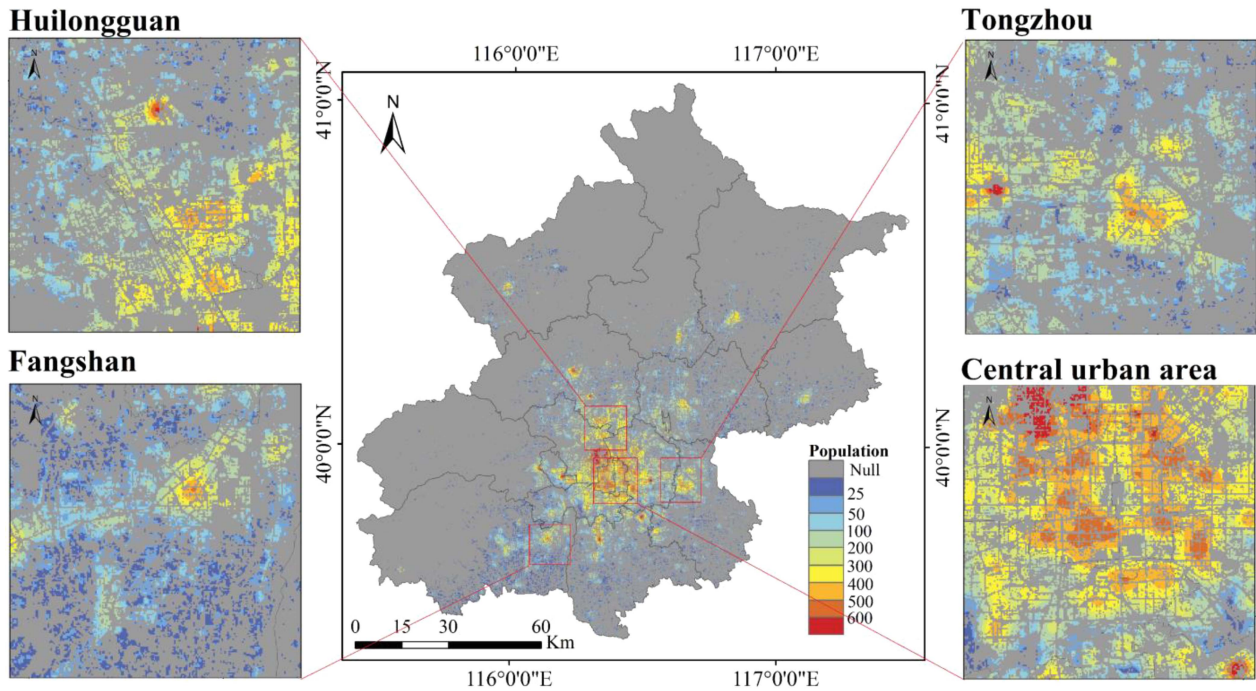


Fig. 9. Spatial distribution result of the gridded population (grid size is 100 m × 100 m) in Beijing in 2020. Four different population densities are shown.

TABLE VI  
STATISTICAL INFORMATION ON THE DISTRIBUTION OF THE LOCAL  $R^2$

Model \ Range	<0.65	0.65–0.70	0.70–0.75	0.75–0.80	0.80–0.85	>0.85
INTL+POI	19.64%	19.33%	19.64%	33.23%	8.15%	0
Zonal classification	0.91%	11.78%	16.31%	11.18%	13.29%	46.53%

As shown in Table VI, before the implementation of the zonal classification strategy, 41.38% of the subdistricts had an  $R^2$  higher than 0.75. After the zone strategy was implemented, 71% of the subdistricts had an  $R^2$  higher than 0.75. As shown in Fig. 10, the accuracy of all areas was greatly improved, especially in rural areas. In general, the implementation of the zonal classification strategy significantly improved the regional population spatialization accuracy, which is also meaningful for other spatialization research using POI and NTL data.

We can also see from Fig. 10 that there are differences in the spatial distribution of Local R-squared. Local R-squared varies from region to region and had higher values in the surrounding area of the city center. In the future study, we will introduce GW framework to explore and analyze the influence of each variable on population to explore the spatial heterogeneity [62].

C. Limitations and Recommendations for Future Work

We developed an INTL index and used the zonal classification method to build a model between the POIs, INTL, and subdistrict population, which greatly improved the accuracy of the population spatial distribution. However, our research had certain limitations. The grid-level error may be higher, as it was

difficult to obtain the grid-unit census data for Beijing; hence, we were unable to conduct more in-depth accuracy verification. Although the global population dataset can provide grid-level population data, it is still difficult to use as an absolute reference value. There are many derived properties of POIs, such as the number of types and density. The selection of the POI category also needs to be considered. When we used the kernel density method to process the POIs, the search radius for all types of POIs was set as 1.5 km. In fact, the impact radius of POIs varies greatly depending on the type, for example, large hospitals and small clinics. At the same time, some errors and omissions were observed in the ISA extracted from the satellite data, which also led to inevitable errors. The LuoJia1-01 satellite stopped operating in 2019, which limits the use of this data in the long term.

The spatial distribution of the population is a complex process influenced by the natural environment, economic conditions, and social development [63]. Therefore, more variables should be considered when constructing the population estimation model. In future research, we need to consider further geospatial Big Data, such as mobile communication positioning data, POIs, and check-in data, to distinguish commercial, industrial, and residential areas and villages. Building height information is also

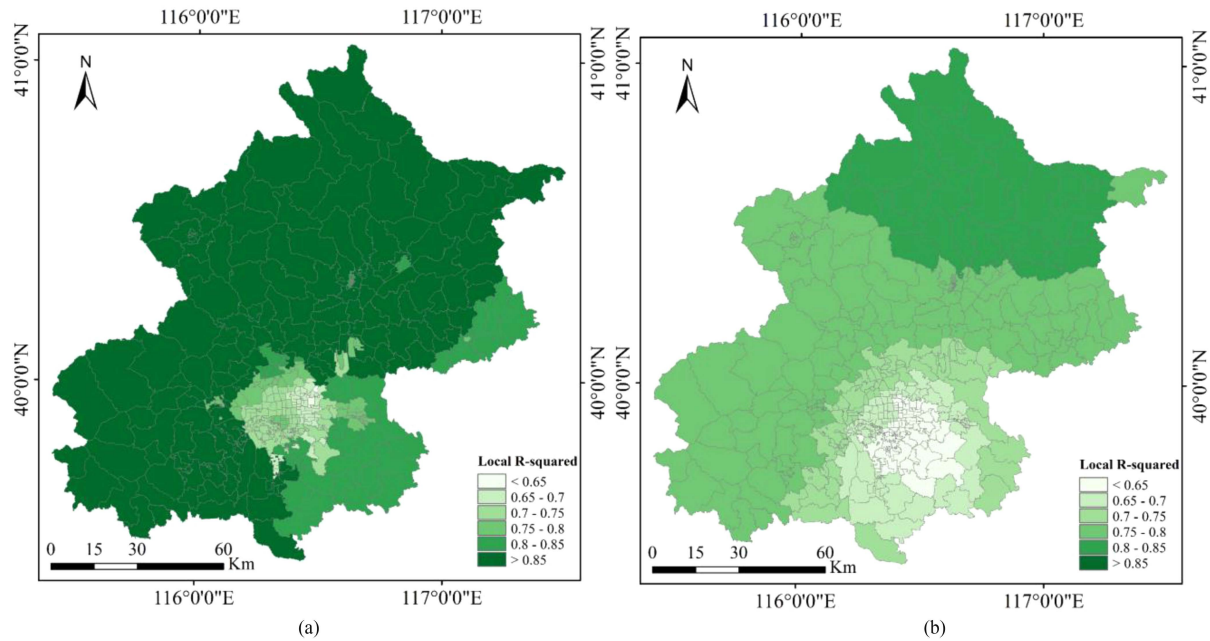


Fig. 10. Spatial information of the Local  $R^2$ . (a) Local  $R^2$  of the zonal classification model. (b) Local  $R^2$  of the model combining the INTL and POI.

important in population estimation but was not used in this study [3]. High-resolution stereo images, which can be used to obtain large-scale and high-precision building height, would have more potential to estimate the population density. Furthermore, the spatial structures in different cities cause differing results for population distribution. Thus, more studies using a combination of remote sensing and geospatial Big Data should be conducted in the future.

## VI. CONCLUSION

In this study, the grid level population spatial distribution of Beijing in 2020 was mapped based on the POIs, Luojia1-01 NTL data, and census data at the subdistrict level. The fine-scale population spatialization dataset provides rich information for urban planning, resource allocation, and environmental protection. According to the results of this study, the main conclusions can be summarized as follows:

- 1) The INTL suppressed high values and enhanced low values in the Luojia1-01 NTL image through using square roots and logarithm transformation. The transformed image better represented the spatial distribution of the population.
- 2) After zoning the study area into rural and urban areas based on the TBNII, different GWR models were established, which decreased the confusion between the dependent variables and the independent variable. The zonal classification model reduced the overestimation of the population in rural areas and the underestimation in urban areas.
- 3) Compared with the WorldPop, LandScan, and previous models based on the NTL and POI, the zonal classification model had the highest  $R^2$  and the lowest MAE and RMSE, and the accuracy of the model was significantly improved.

In future research, we will consider high spatial and multi-spectral satellite imagery in the model as auxiliary data. Meanwhile, we will extend the population spatialization to a larger scale such as for the whole of China.

## REFERENCES

- [1] Z. Cheng, J. Wang, and Y. Ge, "Mapping monthly population distribution and variation at 1-km resolution across China," *Int. J. Geographical Inf. Sci.*, vol. 36, pp. 1–19, Dec. 2020, doi: [10.1080/13658816.2020.1854767](https://doi.org/10.1080/13658816.2020.1854767).
- [2] X. Huang, C. Wang, Z. Li, and H. Ning, "A 100 m population grid in the CONUS by disaggregating census data with open-source Microsoft building footprints," *Big Earth Data*, vol. 5, no. 1, pp. 112–133, Jan. 2021, doi: [10.1080/20964471.2020.1776200](https://doi.org/10.1080/20964471.2020.1776200).
- [3] S. Shang, S. Du, S. Du, and S. Zhu, "Estimating building-scale population using multi-source spatial data," *Cities*, vol. 111, 2021, Art. no. 103002, doi: [10.1016/j.cities.2020.103002](https://doi.org/10.1016/j.cities.2020.103002).
- [4] T. Ye et al., "Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model," *Sci. Total Environ.*, vol. 658, pp. 936–946, 2019, doi: [10.1016/j.scitotenv.2018.12.276](https://doi.org/10.1016/j.scitotenv.2018.12.276).
- [5] D. Azar, R. Engstrom, J. Graesser, and J. Comenetz, "Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data," *Remote Sens. Environ.*, vol. 130, pp. 219–232, 2013, doi: [10.1016/j.rse.2012.11.022](https://doi.org/10.1016/j.rse.2012.11.022).
- [6] N. A. Wardrop et al., "Spatially disaggregated population estimates in the absence of national population and housing census data," *Proc. Nat. Acad. Sci.*, vol. 115, no. 14, pp. 3529–3537, Apr. 2018, doi: [10.1073/pnas.1715305115](https://doi.org/10.1073/pnas.1715305115).
- [7] E. Daniele, B. Deborah, and R. Sliuzas, "Measuring and understanding global human settlements patterns and processes: Innovation, progress and application," *Int. J. Digit. Earth*, vol. 13, no. 1, pp. 2–8, Jan. 2020, doi: [10.1080/17538947.2019.1630072](https://doi.org/10.1080/17538947.2019.1630072).
- [8] J. M. Fitton, M. Lehmann, and D. C. Major, "Identifying coastal towns and small cities in Denmark using global population data to support climate change adaptation," *Int. J. Digit. Earth*, vol. 13, no. 9, pp. 1040–1054, Sep. 2020, doi: [10.1080/17538947.2019.1642403](https://doi.org/10.1080/17538947.2019.1642403).
- [9] S. Freire et al., "Enhanced data and methods for improving open and free global population grids: Putting 'leaving no one behind' into practice," *Int. J. Digit. Earth*, vol. 13, no. 1, pp. 61–77, Jan. 2020, doi: [10.1080/17538947.2018.1548656](https://doi.org/10.1080/17538947.2018.1548656).

- [10] L. Han, W. Zhou, W. Li, and Y. Qian, "Urbanization strategy and environmental changes: An insight with relationship between population change and fine particulate pollution," *Sci. Total Environ.*, vol. 642, pp. 789–799, 2018, doi: [10.1016/j.scitotenv.2018.06.094](https://doi.org/10.1016/j.scitotenv.2018.06.094).
- [11] S. Leyk et al., "The spatial allocation of population: A review of large-scale gridded population data products and their fitness for use," *Earth Syst. Sci. Data*, vol. 11, no. 3, pp. 1385–1409, Sep. 2019, doi: [10.5194/essd-11-1385-2019](https://doi.org/10.5194/essd-11-1385-2019).
- [12] Y. Qiu, X. Zhao, D. Fan, S. Li, and Y. Zhao, "Disaggregating population data for assessing progress of SDGs: Methods and applications," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 2–29, Dec. 2022, doi: [10.1080/17538947.2021.2013553](https://doi.org/10.1080/17538947.2021.2013553).
- [13] J. Mennis, "Dasymetric mapping for estimating population in small areas: Dasymetric mapping for estimating population in small areas," *Geography Compass*, vol. 3, no. 2, pp. 727–745, Mar. 2009, doi: [10.1111/j.1749-8198.2009.00220.x](https://doi.org/10.1111/j.1749-8198.2009.00220.x).
- [14] J. Mennis and T. Hultgren, "Intelligent dasymetric mapping and its application to areal interpolation," *Cartography Geographic Inf. Sci.*, vol. 33, no. 3, pp. 179–194, Jan. 2006, doi: [10.1559/152304006779077309](https://doi.org/10.1559/152304006779077309).
- [15] F. J. Goerlich and I. Cantarino, "A population density grid for Spain," *Int. J. Geographical Inf. Sci.*, vol. 27, no. 12, pp. 2247–2263, Dec. 2013, doi: [10.1080/13658816.2013.799283](https://doi.org/10.1080/13658816.2013.799283).
- [16] X. Li and W. Zhou, "Dasymetric mapping of urban population in China based on radiance corrected DMSP-OLS nighttime light and land cover data," *Sci. Total Environ.*, vol. 643, pp. 1248–1256, 2018, doi: [10.1016/j.scitotenv.2018.06.244](https://doi.org/10.1016/j.scitotenv.2018.06.244).
- [17] F. R. Stevens et al., "Comparisons of two global built area land cover datasets in methods to disaggregate human population in eleven countries from the global South," *Int. J. Digit. Earth*, vol. 13, no. 1, pp. 78–100, Jan. 2020, doi: [10.1080/17538947.2019.1633424](https://doi.org/10.1080/17538947.2019.1633424).
- [18] N. Levin et al., "Remote sensing of night lights: A review and an outlook for the future," *Remote Sens. Environ.*, vol. 237, 2020, Art. no. 111443, doi: [10.1016/j.rse.2019.111443](https://doi.org/10.1016/j.rse.2019.111443).
- [19] L. Wang, H. Fan, and Y. Wang, "Improving population mapping using Luojia 1-01 nighttime light image and location-based social media data," *Sci. Total Environ.*, vol. 730, 2020, Art. no. 139148, doi: [10.1016/j.scitotenv.2020.139148](https://doi.org/10.1016/j.scitotenv.2020.139148).
- [20] M. Alahmadi, S. Mansour, D. Martin, and P. M. Atkinson, "An improved index for urban population distribution mapping based on nighttime lights (DMSP-OLS) data: An experiment in Riyadh Province, Saudi Arabia," *Remote Sens.*, vol. 13, no. 6, Mar. 2021, Art. no. 1171, doi: [10.3390/rs13061171](https://doi.org/10.3390/rs13061171).
- [21] L. Sun, J. Wang, and S. Chang, "Population spatial distribution based on Luojia 1-01 nighttime light image: A case study of Beijing," *Chin. Geographical Sci.*, vol. 31, no. 6, pp. 966–978, Dec. 2021, doi: [10.1007/s11769-021-1240-6](https://doi.org/10.1007/s11769-021-1240-6).
- [22] X. Yang, W. Yue, and D. Gao, "Spatial improvement of human population distribution based on multi-sensor remote-sensing data: An input for exposure assessment," *Int. J. Remote Sens.*, vol. 34, no. 15, pp. 5569–5583, Aug. 2013, doi: [10.1080/01431161.2013.792970](https://doi.org/10.1080/01431161.2013.792970).
- [23] Q. Zhou, Y. Zheng, J. Shao, Y. Lin, and H. Wang, "An improved method of determining human population distribution based on Luojia 1-01 nighttime light imagery and road network data—A case study of the city of Shenzhen," *Sensors*, vol. 20, no. 18, Sep. 2020, Art. no. 5032, doi: [10.3390/s20185032](https://doi.org/10.3390/s20185032).
- [24] C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh, "VIIRS nighttime lights," *Int. J. Remote Sens.*, vol. 38, no. 21, pp. 5860–5879, Nov. 2017, doi: [10.1080/01431161.2017.1342050](https://doi.org/10.1080/01431161.2017.1342050).
- [25] C. Elvidge, M. Zhizhin, K. Baugh, F. Hsu, and T. Ghosh, "Extending nighttime combustion source detection limits with short wavelength VIIRS data," *Remote Sens.*, vol. 11, no. 4, Feb. 2019, Art. no. 395, doi: [10.3390/rs11040395](https://doi.org/10.3390/rs11040395).
- [26] X. Li, Z. Liu, X. Chen, and J. Sun, "Assessing the ability of Luojia 1-01 imagery to detect feeble nighttime lights," *Sensors*, vol. 19, no. 17, Aug. 2019, Art. no. 3708, doi: [10.3390/s19173708](https://doi.org/10.3390/s19173708).
- [27] X. Li, X. Li, D. Li, X. He, and M. Jendryke, "A preliminary investigation of Luojia-1 night-time light imagery," *Remote Sens. Lett.*, vol. 10, no. 6, pp. 526–535, Jun. 2019, doi: [10.1080/2150704X.2019.1577573](https://doi.org/10.1080/2150704X.2019.1577573).
- [28] N. Gao, F. Li, H. Zeng, D. van Bilsen, and M. De Jong, "Can more accurate night-time remote sensing data simulate a more detailed population distribution?," *Sustainability*, vol. 11, no. 16, Aug. 2019, Art. no. 4488, doi: [10.3390/su11164488](https://doi.org/10.3390/su11164488).
- [29] C. Kyba et al., "High-resolution imagery of earth at night: New sources, opportunities and challenges," *Remote Sens.*, vol. 7, no. 1, pp. 1–23, Dec. 2014, doi: [10.3390/rs70100001](https://doi.org/10.3390/rs70100001).
- [30] N. Levin and Y. Duke, "High spatial resolution night-time light images for demographic and socio-economic studies," *Remote Sens. Environ.*, vol. 119, pp. 1–10, 2012, doi: [10.1016/j.rse.2011.12.005](https://doi.org/10.1016/j.rse.2011.12.005).
- [31] P. Gao, T. Wu, Y. Ge, and Z. Li, "Improving the accuracy of extant gridded population maps using multisource map fusion," *GIScience Remote Sens.*, vol. 59, no. 1, pp. 54–70, Dec. 2022, doi: [10.1080/15481603.2021.2012371](https://doi.org/10.1080/15481603.2021.2012371).
- [32] Y. Yao et al., "Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data," *Int. J. Geographical Inf. Sci.*, vol. 31, pp. 1–25, Feb. 2017, doi: [10.1080/13658816.2017.1290252](https://doi.org/10.1080/13658816.2017.1290252).
- [33] M. Bakillah, S. Liang, A. Mobasheri, J. J. Arsanjani, and A. Zipf, "Fine-resolution population mapping using OpenStreetMap points-of-interest," *Int. J. Geographical Inf. Sci.*, vol. 28, no. 9, pp. 1940–1963, Sep. 2014, doi: [10.1080/13658816.2014.909045](https://doi.org/10.1080/13658816.2014.909045).
- [34] L. Li et al., "Mapping changing population distribution on the Qinghai-Tibet Plateau since 2000 with multi-temporal remote sensing and point-of-interest data," *Remote Sens.*, vol. 12, no. 24, Dec. 2020, Art. no. 4059, doi: [10.3390/rs12244059](https://doi.org/10.3390/rs12244059).
- [35] Y. Wang, C. Huang, M. Zhao, J. Hou, Y. Zhang, and J. Gu, "Mapping the population density in Mainland China using NPP/VIIRS and points-of-interest data based on a random forests model," *Remote Sens.*, vol. 12, no. 21, Nov. 2020, Art. no. 3645, doi: [10.3390/rs12213645](https://doi.org/10.3390/rs12213645).
- [36] X. Yang et al., "Population mapping with multisensor remote sensing images and point-of-interest data," *Remote Sens.*, vol. 11, no. 5, Mar. 2019, Art. no. 574, doi: [10.3390/rs11050574](https://doi.org/10.3390/rs11050574).
- [37] Z. Yang, C. Fang, G. Li, and X. Mu, "Integrating multiple semantics data to assess the dynamic change of urban green space in Beijing, China," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 103, 2021, Art. no. 102479, doi: [10.1016/j.jag.2021.102479](https://doi.org/10.1016/j.jag.2021.102479).
- [38] M. A. Wulder et al., "Current status of Landsat program, science, and applications," *Remote Sens. Environ.*, vol. 225, pp. 127–147, 2019, doi: [10.1016/j.rse.2019.02.015](https://doi.org/10.1016/j.rse.2019.02.015).
- [39] W. Guo, G. Li, W. Ni, Y. Zhang, and D. Lu, "Exploring improvement of impervious surface estimation at national scale through integration of nighttime light and Proba-V data," *GI-Science Remote Sens.*, vol. 55, no. 5, pp. 699–717, Sep. 2018, doi: [10.1080/15481603.2018.1436425](https://doi.org/10.1080/15481603.2018.1436425).
- [40] W. Guo, Y. Zhang, and L. Gao, "Using VIIRS-DNB and Landsat data for impervious surface area mapping in an arid/semiarid region," *Remote Sens. Lett.*, vol. 9, no. 6, pp. 587–596, Jun. 2018, doi: [10.1080/2150704X.2018.1455234](https://doi.org/10.1080/2150704X.2018.1455234).
- [41] D. Palacios-Lopez et al., "New perspectives for mapping global population distribution using world settlement footprint products," *Sustainability*, vol. 11, no. 21, Oct. 2019, Art. no. 6056, doi: [10.3390/su11216056](https://doi.org/10.3390/su11216056).
- [42] L. Li and D. Lu, "Mapping population density distribution at multiple scales in Zhejiang Province using Landsat thematic mapper and census data," *Int. J. Remote Sens.*, vol. 37, no. 18, pp. 4243–4260, Sep. 2016, doi: [10.1080/01431161.2016.1212422](https://doi.org/10.1080/01431161.2016.1212422).
- [43] S. K. Mcfeeters, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1425–1432, May 1996, doi: [10.1080/01431169608948714](https://doi.org/10.1080/01431169608948714).
- [44] W. Guo, C. Zhao, Y. Zhang, and S. Gao, "Mapping impervious surface distribution and dynamics in an arid/semiarid area—A case study in Ordos, China," *IEEE Access*, vol. 9, pp. 19659–19673, 2021, doi: [10.1109/ACCESS.2021.3054963](https://doi.org/10.1109/ACCESS.2021.3054963).
- [45] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, Jul. 1991, doi: [10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).
- [46] G. Foody, "Assessing the accuracy of remotely sensed data: Principles and practices: Book reviews," *Photogrammetric Rec.*, vol. 25, no. 130, pp. 204–205, Jun. 2010, doi: [10.1111/j.1477-9730.2010.00574\\_2.x](https://doi.org/10.1111/j.1477-9730.2010.00574_2.x).
- [47] C. Zhang, Y. Pei, J. Li, Q. Qin, and J. Yue, "Application of Luojia 1-01 nighttime images for detecting the light changes for the 2019 spring festival in western cities, China," *Remote Sens.*, vol. 12, no. 9, Apr. 2020, Art. no. 1416, doi: [10.3390/rs12091416](https://doi.org/10.3390/rs12091416).
- [48] X. Huang, J. Yang, J. Li, and D. Wen, "Urban functional zone mapping by integrating high spatial resolution nighttime light and daytime multi-view imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 175, pp. 403–415, 2021, doi: [10.1016/j.isprsjprs.2021.03.019](https://doi.org/10.1016/j.isprsjprs.2021.03.019).
- [49] Q. Zheng et al., "A new source of multi-spectral high spatial resolution night-time light imagery—JLI-3B," *Remote Sens. Environ.*, vol. 215, pp. 300–312, 2018, doi: [10.1016/j.rse.2018.06.016](https://doi.org/10.1016/j.rse.2018.06.016).



- [50] H. Bagan and Y. Yamagata, "Analysis of urban growth and estimating population density using satellite images of nighttime lights and land-use and population data," *GIScience Remote Sens.*, vol. 52, no. 6, pp. 765–780, Nov. 2015, doi: [10.1080/15481603.2015.1072400](https://doi.org/10.1080/15481603.2015.1072400).
- [51] L. Wang et al., "Mapping population density in China between 1990 and 2010 using remote sensing," *Remote Sens. Environ.*, vol. 210, pp. 269–281, 2018, doi: [10.1016/j.rse.2018.03.007](https://doi.org/10.1016/j.rse.2018.03.007).
- [52] W. Song, H. Jia, J. Huang, and Y. Zhang, "A satellite-based geographically weighted regression model for regional PM<sub>2.5</sub> estimation over the Pearl River Delta region in China," *Remote Sens. Environ.*, vol. 154, pp. 1–7, 2014, doi: [10.1016/j.rse.2014.08.008](https://doi.org/10.1016/j.rse.2014.08.008).
- [53] S. Wang, C. Shi, C. Fang, and K. Feng, "Examining the spatial variations of determinants of energy-related CO<sub>2</sub> emissions in China at the city level using geographically weighted regression model," *Appl. Energy*, vol. 235, pp. 95–105, 2019, doi: [10.1016/j.apenergy.2018.10.083](https://doi.org/10.1016/j.apenergy.2018.10.083).
- [54] A. Stewart Fotheringham, M. Charlton, and C. Brunsdon, "The geography of parameter space: An investigation of spatial non-stationarity," *Int. J. Geographical Inf. Syst.*, vol. 10, no. 5, pp. 605–627, Jul. 1996, doi: [10.1080/02693799608902100](https://doi.org/10.1080/02693799608902100).
- [55] A. R. da Silva and F. F. Mendes, "On comparing some algorithms for finding the optimal bandwidth in geographically weighted regression," *Appl. Soft Comput.*, vol. 73, pp. 943–957, Dec. 2018, doi: [10.1016/j.asoc.2018.09.033](https://doi.org/10.1016/j.asoc.2018.09.033).
- [56] G. Zhang, X. Guo, D. Li, and B. Jiang, "Evaluating the potential of LJI-01 nighttime light data for modeling socio-economic parameters," *Sensors*, vol. 19, no. 6, Mar. 2019, Art. no. 1465, doi: [10.3390/s19061465](https://doi.org/10.3390/s19061465).
- [57] J. Lin, S. Luo, and Y. Huang, "Poverty estimation at the county level by combining LuoJia1-01 nighttime light data and points of interest," *Geocarto Int.*, vol. 37, pp. 1–17, Jan. 2021, doi: [10.1080/10106049.2020.1870166](https://doi.org/10.1080/10106049.2020.1870166).
- [58] M. Morshed, "Measuring Dhaka's urban transformation using nighttime light data," *J. Geovisualization Spatial Anal.*, vol. 6, pp. 1–13, 2022.
- [59] J. Chen et al., "A TIR-visible automatic registration and geometric correction method for SDGSAT-1 thermal infrared image based on modified RIFT," *Remote Sens.*, vol. 14, no. 6, Mar. 2022, Art. no. 1393, doi: [10.3390/rs14061393](https://doi.org/10.3390/rs14061393).
- [60] B. Yu et al., "Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement," *Int. J. Geographical Inf. Sci.*, vol. 33, no. 4, pp. 687–706, Apr. 2019, doi: [10.1080/13658816.2018.1555642](https://doi.org/10.1080/13658816.2018.1555642).
- [61] C. Zeng, Y. Zhou, S. Wang, F. Yan, and Q. Zhao, "Population spatialization in China based on night-time imagery and land use data," *Int. J. Remote Sens.*, vol. 32, no. 24, pp. 9599–9620, Dec. 2011, doi: [10.1080/01431161.2011.569581](https://doi.org/10.1080/01431161.2011.569581).
- [62] J. E. H. Percival, "Exploratory spatial data analysis with gwpcorMapper: An interactive mapping tool for geographically weighted correlation and partial correlation," *J. Geovisualization Spatial Anal.*, vol. 6, pp. 1–16, 2022.
- [63] J. J. Nieves et al., "Examining the correlates and drivers of human population distributions across low- and middle-income countries," *J. Roy. Soc. Interface*, vol. 14, no. 137, Dec. 2017, Art. no. 20170401, doi: [10.1098/rsif.2017.0401](https://doi.org/10.1098/rsif.2017.0401).



**Wei Guo** received the B.Eng. and M.Eng. degrees from Shandong University of Science and Technology, Qingdao, China, in 2008 and 2011, respectively, and the Ph.D. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2015.

He worked with Michigan State University from 2013 to 2015, and the State Key Laboratory of Remote Sensing Science, China, from 2016 to 2019. He is currently an Assistant Professor with the College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include

urban remote sensing and ecological environment monitoring.



**Jinyu Zhang** was born in Weifang, China, in 1999. She is currently working toward the doctor's degree with the School of College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China.

Her research interests include spatialization of socio-economic statistics, urban remote sensing, and regional development.



**Xuesheng Zhao** received the B.Eng. degree from Shandong University of Science and Technology, Tai'an, China, in 1988, and the M.Eng. and Ph.D. degrees from China University of Mining and Technology, Beijing, China, in 1995 and 2002, respectively.

He is currently a Professor with the College of Geoscience and Surveying Engineering, China University of Mining and Technology. His research interests include geographic grid computing and SDGs.



**Yongxing Li** was born in Zaozhuang, China, in 1999. He is currently working toward the master's degree with the School of College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China.

His research interests include remote sensing of nighttime light, urban remote sensing, and spatiotemporal analysis of CO<sub>2</sub> emissions.



**Jinke Liu** was born in Xinzhou, China, in 2002. She is currently working toward the undergraduate degree with the School of College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China.

Her research interests include urban remote sensing and regional development.



**Wenbin Sun** received the B.Eng. and M.Eng. degrees from Southwest Jiaotong University, Chengdu, China, in 2000 and 2003, respectively, and the Ph.D. degree in cartography and geographic information system from China University of Mining and Technology, Beijing, China, in 2007.

He is currently a Professor with the College of Geoscience and Surveying Engineering, China University of Mining and Technology. His research interests include geographic grid computing and ecological environment assessment.



**Deqin Fan** received the B.Eng. degree from Northeast Forestry University, Harbin, China, in 2008, the M.Eng. degree from China University of Mining and Technology, Beijing, China, in 2011, and the Ph.D. degree in cartography and geographic information system from Beijing Normal University, Beijing, China, in 2014.

She is currently an Associate Professor with the College of Geoscience and Surveying Engineering, China University of Mining and Technology. Her research interests include remote sensing of resources

and environment and ecological environment monitoring.