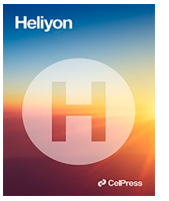
Heliyon 9 (2023) e16285

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Estimate of population density and diagnosis of main factors of spatial heterogeneity in the metropolitan scale, western China

Guangjie Wang [a,b], Wenfu Peng [a,b,*,1], Lindan Zhang [a,b]

[a] The Institute of Geography and Resources Science, Sichuan Normal University, Chengdu, 610068, PR China
[b] Key Lab of Land Resources Evaluation and Monitoring in Southwest, Ministry of Education, Sichuan Normal University, Chengdu, 610068, PR China

ARTICLE INFO

Keywords:
Population density
Night time lights imagery
Influencing factors
Geographic detector
Remote sensing (RS)
Geographic information system (GIS)

ABSTRACT

We estimated the population density and quantified its characteristics using remote sensing, census data, and Geographic Information System (GIS). The interactive influence of these factors on population density was quantified based on geographic detectors to identify the differentiation mechanisms in the Chengdu metropolitan area of China. We identified the key factors that contribute to population density growth. The models used to simulate population density had the highest R^2 values (>0.899). Population density tended to increase with time, with a multicentre spatial agglomeration pattern; the centre of gravity of the spatial distribution tended to move from the southeast to the northwest. Industry proportions, Normalised Difference Vegetation Index (NDVI), land use, distance to urban centers or construction land, and GDP per capita can satisfactorily explain population density changes. The combined impact of these elements on population density variation exhibited mutual and non-linear strengthening, with the mutual effect of the two elements intensifying the impact of each individual element. Our study identified the key driving forces that contribute to the differentiation of population density, which can provide valuable support for the development of effective regional and targeted population planning guidelines.

1. Introduction

The 2020 Report on the World Population Situation highlights that the global population reached 7.75 billion in that year, reflecting a highly diverse state of affairs [1]. With reform and transparency, the spatial structure of China's economic development has experienced profound changes [2], regional economic development is substantially differentiated, population and agglomeration space are markedly different, and central cities and urban agglomerations are becoming the main spatial forms of populations [3]. Population and spatial distributions provide important support for socioeconomic development, as the clustering of people provides human resources and drives demand in regional markets [4]. For the present and future development of a country, population density must be accurately analyzed, and corresponding population policies must be formulated [5,6]. Although censuses are an important source of administrative data for population distributions [7], they reflect only capture the total population within a given administrative regions and do not offer location information that can accurately reflect temporal and spatial concentration and distribution

* Corresponding author. The Institute of Geography and Resources Science, Sichuan Normal University, Chengdu, 610068, PR China.
E-mail addresses: wanggj@sicnu.edu.cn (G. Wang), pwfzh@126.com (W. Peng), yjx133505@gmail.com (L. Zhang).
[1] Equally contributed to this work.

https://doi.org/10.1016/j.heliyon.2023.e16285
Received 3 October 2022; Received in revised form 7 May 2023; Accepted 11 May 2023
Available online 27 May 2023

patterns of population [5,6]. This does not offer more precise information for developing effective regional and targeted population planning guidelines [5,7,8]. Population data spatialization is a method for inverting the distribution of a population in a certain time and space by using parameters and models that can clearly expressing the number and patterns of population [9]. Using demographic data or applying the kernel density method, some researchers have obtained a continuous change map of regional population density or have displayed demographic data by county administrative district [10–13]. In recent years, researchers have used methods such as area weighting [14], geographically weighted regression [15,16], and zonal density mapping [17,18] to decompose census data into spatial grid units and obtain gridded population data covering different regions.

The progress of geospatial information technique has enabled remote sensing (RS) data to be extensively used in population drawing study due to its quick and wide range cover [5]. Nightlight data and satellite image are at present go to calculate population [5,6]. Researchers have utilized the Defense Meteorological Program Operational Line-Scan System (DMSP/OLS) and the National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) along with population statistics to calculate the population of prefecture-level cities of three northeastern provinces in China [19]. The population density of the Su–Xi–Chang region in China was calculated by using Luojia-1A Nighttime Light data [20]. The population density of Wuhan, China, was estimated using Landsat Operational Land Imager (OLI) images and NPP–VIIRS light data [21]. Based on experiments in African cities, researchers have make sure that NPP–VIIRS data can provide more accurate population estimates than DMSP/OLS nightlight data [22]. However, the employ of nightlight data merely to calculate population information is undependable [5]. The establishment of a new population calculation method was showed with integrating land use and night light data [6]. Some researchers have utilized satellite and open-source GIS data, such as the WorldPop program, to simulate the global population density [23]. A random forest model was used to multi-source data and points of interest to calculate population patterns in impervious areas with a spatial resolution of 30 m in the Bohai region [9]. However, population modeling is still subject to the organization work restriction in census [24].

Both foreign and domestic studies have analyzed the drivers of population change [5]. By utilizing the centre of gravity analysis [11], spatial autocorrelation analysis, methods of centralisation index and migration [12], or multiple linear regression models [13], researchers have studied the evolution of spatial patterns and its influence elements of population in China from 2000 to 2020, or other regions like the Yangtze River Delta. Using the partial least squares (PLS) method [25,26], a spatial simultaneous equation model [27], or a correlation analysis method [28], the elements impacting population density of different areas in China have been quantitatively analyzed. The immediate and spatial overflow impact of the influencing factors on population distribution have been revealed using a general nesting spatial econometric model [29]. Researchers have employed spatial analysis, statistical analysis of GIS, and regression methods to evaluate the spatial-temporal variation features of population distributions and its impact elements in mountainous counties, urban agglomerations, and the Beijing–Tianjin–Hebei region [4,6,30,31].

Although the findings of these studies are important and have contributed substantially to the knowledge and understanding of changes in population distribution and its complex relationship of impacting factors, the study ways are mostly based on census and statistical yearbook data with the application of linear, trend, and correlation analyses [32]. Spatial heterogeneity is a presentation of no human influence and socio-economic procedure. Although the issue of spatial stratification difference is emphasized by both micro areas or huge regions research and spatial big data, statistical ways for examining spatial difference are yet deficiencies [33–35].

A geodetector is a new statistical ways for investigating spatial difference and uncovering its key impacting elements. An advantage of geographic detectors can investigate both quantitative data and type data [36]. Another unique advantage of a geographical detector can investigate the mutual effect between two elements [36]. We can decide whether the two elements mutual effect, such as intensity, orientation, linear or nonlinear of mutual effect through computing and contrasting the q values of every element and the q values of the two overlay elements [36–38]. This method has been widespread applied in a lot of sciences fields [36]. Therefore, this study used geodetector to investigate the spatial difference and patterns of population density.

The main goals of the study were to: 1) develop a model based on DMSP-OLS, NPP-VIIRS data, and NDVI for simulating population density at the county scales, as well as the map population density to uncover spatial-temporal schema and features; 2) identify the main factors in the spatial differentiation of population density; 3) detect mutual effect of elements in the spacial difference of population density; and 4) identify the features and range of main elements in the spatial difference in population density change. Our research consequence and calculation methods can supply significance demographic information and method help of regions that lack population data.

2. Study area

Chengdu metropolitan area is located in the Chengdu Plain Economic Zone at 29°23′49″N–31°41′50″N and 102°23′49″E−105°45′7″E. This region has the most thriving economic development, strongest innovation capacity, and highest degree of openness in China. Chengdu metropolitan area has a resident population of about 2966×10^4 people. The city of Chengdu comprises three cities (Deyang, Meishan, and Ziyang), as well as 35 districts, county-level cities, and counties. It covers an area of 33,100 km^2. Additionally, the total economic volume of the Chengdu metropolitan area accounted for more than 10% of China in 2020. Longmen Mountains in northwestern Chengdu metropolitan area belong to the Minshan-Qionglai Ecological Function Protection Zone. The topography is dominated by low-altitude hills, plains, terraces, and rolling mountains. The soil primarily soil types include acidic purple, paddy, and bleached yellow. Plant life is composed of cultivated vegetation (paddy and upland field rotations, evergreen orchards, and subtropical economic forests), subtropical coniferous forests, and subtropical evergreen and deciduous broadleaf mixed forests. The land use is mainly arable (non-irrigated and paddy), forestland, and construction land. Chengdu metropolitan area belongs to subtropical monsoon climate, and its climate features is early spring, hot summer, cool autumn and warm winter, with an AAT of 16 °C and AAP of about 1000 mm. Owing to vertical height differences, it has vertical climate zones with

obvious variations in heat; thus, it has rich, diverse, and concentrated biological resources and a varied natural ecology.

3. Data sources and methods

3.1. Data sources

The data used in this study included RS images and auxiliary data. RS images included DMSP-OLS, NPP-VIIRS, and the NDVI. Auxiliary data covered DEM, land use types, statistical yearbook data, and administrative district boundary (Table 1). The study data are unified into WGS1984 projection coordinates and GCS_WGS_1984 geographic coordinates.

(1) DMSP-OLS and NPP-VIIRS data were obtained from Earth Observation Group (http://ngdc.noaa.gov/eog/). The pixel gray values of DMSP-OLS data and NPP-VIIRS data is range from 0 to 63, with a spatial resolution of about 1000 m and 500 m, respectively [39].
(2) DEM, Landsat 5/8 images, weather data, and geomorphological data were obtained from the Resource and Environmental Sciences Data Center of the Chinese Academy of Sciences (http://www.resdc.cn). Land use data were extracted from Landsat 8 remote sensing images.
(3) The NDVI data were obtained from MOD13Q1 in NASA (https://www.nasa.gov/), with a spatial resolution of 250 m. The MODIS Reprojection Tool was applied to convert the sinusoidal projection of the MODIS MOD13Q1 product into a Universal Transverse Mercator projection, with the projection coordinate system set to WGS_84. The annual NDVI was synthesized according to the maximum synthesis method.
(4) Distance, road, farmland, forestland, grassland, construction land, and water, as well as TOR and land use, were obtained using GIS technology. The GDP density, population density, PPI,PSI,PTI, and IEE data were obtained from the Sichuan Statistical Yearbook (2001–2021), spatialized using GIS.
(5) Administrative district vector map include 1:1 million Chinese counties, townships, and town-level vector boundary, obtained from the National Geomatics Center of China (http://www.ngcc.cn).

3.2. Methodology

3.2.1. Processing of DMSP-OLS and NPP-VIIRS data

RS image processing includes the correction of DMSP-OLS and NPP-VIIRS, and mutual correction of two types of nighttime light data [5]. This study used various methods, such as correction for continuity, oversaturation correction, and regression fitting, to correct the 2000 and 2010 DMSP-OLS and 2020 NPP-VIIRS [40–43]. To minimize projection distortion, the projection was converted to an Albers projection for China; the spatial resolution was uniformly resampled to 1 km by bilinear interpolation. Due to the saturation problems in core cities and bright areas, *DN* anomalies in the DMSP/OLS and NPP/VIIRS were eliminated based on band math in ENVI 5.3 [44,45]: (*DN lt* − 63) × 0 + (*DN gt* 63) × 0 + (*DN ge* − 63 and *DN le* 63) × *DN*. The maximum image value of Chengdu area was used to remove the *DN* value of extremely bright isolated pixels in the night RS data [45]. Therefore, a data set expressing the natural surface light strength was established.

Table 1
Factors indicators.

Category	Code	Index	Abbreviation	Unit	Resolution
Economics	x_1	GDP density	–	RMB yuan/km^2	30 m
	x_2	GDP per capita	GDPC	RMB yuan/person	30 m
	x_3	Proportion of primary industry	PPI	%	30 m
	x_4	Proportion of secondary industry	PSI	%	30 m
	x_5	Proportion of tertiary industry	PTI	%	30 m
	x_6	Investment in real estate	IEE	10^4 RMB yuan	30 m
Location	x_7	Distance to urban centers	DUC	m	30 m
	x_8	Distance to road	DRO	m	30 m
	x_9	Distance to farmland	DFR	m	30 m
	x_{10}	Distance to forestland	DFO	m	30 m
	x_{11}	Distance to grassland	DGR	m	30 m
	x_{12}	Distance to construction land	DCOL	m	30 m
	x_{13}	Distance to water	DWA	m	30 m
Environment	x_{14}	Elevation	–	m	30 m
	x_{15}	Topographic relief	TOR	°	30 m
	x_{16}	Geomorphy	–	types	1 km
	x_{17}	NDVI	–	–	250 m
	x_{18}	Annual average temperature	AAT	°C	1 km
	x_{19}	Average annual precipitation	AAP	mm	1 km
	x_{20}	Land use degree index	LUDI	–	30 m

3.2.2. Population density estimation model

There are many methods to simulate spatialization of socioeconomic data through night-time lights and NDVI, include linear regression, logarithmic regression, and second-order regression [46]. In these equations, linear regression equations are more precise and easier to come true [47]. An precise regression equations for simulating population density was constructed by examining the correspondence between population density and three indicators: mean nighttime light data, mean NDVI, and their ratio (m). The model is suitable for calculating population density based on high R^2 values and good *F*-test values; this population density dataset is referred to as the statistical population density. Pixel-level population density was calculated using raster calculators in GIS or Band math in ENVI 5.3 based on the calculated model and spatial raster layer data (mean nighttime light data and mean NDVI). The regression equations was utilized in this research to estimate population density as follows:

$$pd = a \times x_1 + b \times x_2 + c \times x_3 + d \tag{1}$$

$$x_3 = \frac{n_1}{n_2} \tag{2}$$

where *pd* is the population density of the statistical counties; x_1 is the mean value of nighttime lights; x_2 is the mean NDVI; x_3 is the ratio of mean nighttime lights to the mean NDVI; n_1 is the night-time lights value of the statistical county; n_2 is the NDVI value of the statistical counties; *a*, *b*, *c*, and *d* are regression coefficients and intercepts, respectively.

3.2.3. Spatial autocorrelation method

Spatial correlational analysis is applied to examine the interdependent degree in data at different location [47]. Global *Moran's I* is a synthetically detection of spatial autocorrelation across the research region. It was calculated to detection if the spatial patterns of the population density was collective, scatter, or stochastic [48–50]. The calculation of the global *Moran's I* (Moran, 1950) is as follows [45]:

$$I = \frac{n}{s_0} \times \frac{\sum_i^n \sum_j^n w_{ij}(x_i -)(s_j -)}{\sum_i^n (x_i -)^2} \tag{3}$$

where *n* is the number of townships in the research region, x_i, x_j are the surveyed values of the spatial vector boundary units, respectively, $\bar{x}$ is the mean of the investigations, w_{ij} is the spatial weight matrices, and S_0 is the sum of the spatial weight matrix.

Global *Moran's I* was applied to check spatial correspondence, with values is ranging of −1 and 1. If *Moran's I* > 0, it showed a positive spatial relativity in population agglomeration between regions. The higher the value, the greater the magnitude of positive relativity. If *Moran's I* < 0, there was a negative relativity with a diffuse spatial patterns. The lower the value, the greater the degree of negative correspondence.

Local *Moran's I* was used to discern the locations of the spatial collective and spatial scatter at the township level [48]. Local *Moran's I* of the spatial relativity is as follows:

$$I_i = \frac{x_i - }{s^2} \sum\nolimits_{j=1, j \neq i}^{n} w_{i,j}(x_j -) \tag{4}$$

where x_i is the population density of the spatial vector boundary units *i,* i.e., the population density, $\overline{X}$ is the mean of the population density, $w_{i,j}$ is the spatial weight between the vector boundary units *i* and *j*, and there is:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j -)^2}{n - 1} \tag{5}$$

where *n* is the number of spatial vector boundary units in the research region.

3.2.4. Hotspot analysis

Global spatial autocorrelation analysis was applied to discern whether the population density was spatially collective; but it did not reflect the specific location of the collective area. The Getis-Ord *G*I* statistic was calculated applying for hot pints analysis method to assess where the spatial collective of high and low values of population density elements occurrences [48], population density element values [45,48]. The *G*I* statistics are expressed as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}} \tag{6}$$

where x_j is the population density of the spatial vector boundary units *i,* i.e., the population density, $\overline{X}$ is the mean of the population density, $w_{i,j}$ is the spatial weight between vector boundary units *i* and *j*, and *n* is the number of spatial vector boundary units in the research region, and there is:

$$S=\sqrt{\frac{\sum_{j=1}^{n}x_j^2}{n}-(\overline{X})^2} \tag{7}$$

where the *Gi** statistic is the *Z* score.

3.2.5. Geographical detector

Geographic detectors are a spatial statistical model used to discern spatial non uniformity and recognize the key impact elements applying for risk, ecological factors, and spatial mutual effect [33].

3.2.5.1. Detection of spatial heterogeneity and factors. This is applied to discern the spatial non uniformity of the population density and extent to which the detection elements make clear the spatial non uniformity of the attribute population density. First, a spatial overlay analysis of the population density layer with the factor layer was conducted. Second, the different spatial classes of the factors were partitioned or classified. Third, significance tests of the differences in factor means were verified to discern the comparatively significance of the factors. The method used to calculate the interpretation power (p) of the factors [33] was as follows:

$$p=1-\frac{\sum_{h=1}^{L}N_h\sigma_h^2}{N\sigma^2}=1-\frac{SSW}{SST} \tag{8}$$

where the range of p-values is [0, 1], with larger values make clear higher interpretation power for the elements on population density; $h=1,\ldots,L$ is the types or division of the variable (Y) or elements (x); N_h and N are layer h and the number of regional vector boundary units, respectively; σ_h^2 and σ^2 are the variance of Y values for layer h and the study area, respectively; SSW and SST are the sum of variance within the layer and total regional variance, respectively [45]. The variance of the regional Y values is calculated as follows:

$$\sigma^2=\frac{1}{N-1}\sum_{i=1}^{N}(Y_i-)^2 \tag{9}$$

where Y_j and are the value of sample j and the average value of region Y, respectively.

3.2.5.2. Factor interaction detection. To identify interactions between different factors, interaction detection was employed [33]. This method allowed us to assess how different factors, both individually and in combination, influenced changes in population density. First, we calculated p-values for each of the two population density factors being analyzed. Second, the p-values of the factor interactions were calculated; p (x_i) and p (x_j) were compared with p ($x_i \cap x_j$).

3.2.5.3. Detection of risk zones. Risk detection is applied to determine whether there is a remarkable difference in attribute average values between two elements' sub-regions [33]. It is used to search for areas with high population density. The t-statistic is applied to test risk:

$$t=\frac{\overline{Y}_{h=1}-\overline{Y}_{h=2}}{\left[\frac{Var(Y_{h=1})}{n_h=1}+\frac{Var(Y_{h=2})}{n_h=2}\right]^{1/2}} \tag{10}$$

where $\overline{Y}_h$ denotes the average values of population density in sub-region h, n_h denotes the number of samples in sub-region h, and Var denotes the variance (Wang et al., 2010).

3.2.5.4. Ecological detection. Ecological detection is applied to determine whether there is a remarkable difference in the effect of two elements, x_i and x_j, on the spatial patterns of the population density [33]. This can determine whether x_i has a more remarkable effect on the spatial patterns of the population density than that of x_j. Whether there is a remarkable diversity in the effect of factor x_i over x_j on the spatial patterns of population density is measured by the F-statistic [33].

$$F=\frac{N_{xi}\times\left(N_{xj}-1\right)\times SSW_{xi}}{N_{xj}\times(N_{xi}-1)\times SSW_{xj}} \tag{11}$$

$$SSW_{xi}=\sum_{h=1}^{L_i}N_h\sigma_h^2 \tag{12}$$

where N_{xi} and N_{xj} are the specimen size of the two elements [33]; SSW_{xi} and SSW_{xj} are the total square deviation within layers formed by the two elements, respectively [24]; and L_i and L_j are the number of layers for variables x_i and x_j, respectively.

4. Results

4.1. Estimation model analysis

We examined the correspondence between population density and three indicators: the mean nighttime light data, mean NDVI, and ratio (*m*). The results showed that the correlation coefficients were 0.954, −0.933, and 0.972 in 2000; 0.885, −0.955, and 0.949 in 2010; and 0.779, −0.907, and 0.876 in 2020, respectively (Table 2). The level of correspondence between the variables was found to be quite high and passed a remarkable test at the 0.001 level. This indicates that the data can be used effectively for spatial modeling of population density. We established the equation between the population density and mean nighttime light data, mean NDVI, and ratio (m) (Table 2). Table 2 indicates that the linear regression model had high R^2 values and good *F*-test values, indicating that it could model population density based on statistical analysis.

We created a scatter plot of statistical versus simulated population densities (Fig. 1). According to Fig. 2, the population density estimated integrating the DMSP-OLS and NPP-VIIRS data and the statistic population density had a high value ($R^2 > 0.899$), showing that DMSP-OLS and NPP-VIIRS data are relatively effective at simulating population density. These consequence show the possibility of using DMSP-OLS and NPP-VIIRS data to estimate the population density.

4.2. Spatial patterns of population density

The region with the highest population density in the Chengdu metropolitan area was determined to be mainly in the central districts of the city and adjacent districts. The density of population was low in northwest mountain region, far from the county-level administrative centers (Fig. 3). The spatial pattern showed polycentric spatial agglomeration dominated by the provincial capital of Chengdu, with population density decreasing from the center of Chengdu to adjacent districts, as well as from the core centers of cities in peripheral areas to their surrounding areas (Fig. 3).

4.3. Spatial autocorrelation characteristics

The global *Moran's I* was 0.842534 in 2000, 0.929473 in 2010, and 0.939337 in 2020 (Fig. 4). The make clear that the spatial patterns of the population density had a positive spatial self-correlation. The *Moran's I* values indicated an increasing trend, indicating that the degree of spatial correlation of the population density tended to increase and that cities with similar population densities tended to be more spatially concentrated. According to Fig. 4, the *z*-scores of 59.462, 65.343, and 65.864 at the same time intervals were significantly greater than 1.65. All of the *p*-values were 0.000; the confidence levels were greater than 95%, indicating that the spatial patterns of the population density observed using the spatial autocorrelation tool had notable clustering features and a spatially positive correlation pattern.

The Chengdu metropolitan area showed local spatial clustering of high observed values for the population density from 2000 to 2020 (Fig. 5), with small spatial differences in the population density and high population density in the regions themselves and surrounding towns and townships with a high population density. According to Fig. 5, in 2000, high population density agglomeration was distributed in most of central city regions in Chengdu成都的中心城区the central urban area of Chengdu. In 2010, it was distributed in the central urban area of Chengdu and neighboring Pidu, the southeast of Wenjiang, and the northern region of Shuangliu in 2020. The agglomeration and distribution of high density of population further expanded in the central city of Chengdu and the southeastern neighboring Pidu; southeast and the western Wenjiang; the northern area and the southwestern Shuangliu; the west, southeast, and northeast of Xindu; and west and southwest of Longquanyi.

According to Fig. 6, Chengdu center and adjacent regions in 2000 formed a significant hotspot distribution region; the population density of these areas showed a high population aggregation distribution at a remarkable level of 99%. The hot spot map of the

Table 2
Population density estimation model parameters.

	Coefficient	95% Confidence interval						
2000	B	Lower limit	Upper limit	Correlation	R^2	Adjusted R^2	*F*	Sig. *F*
(constant)	2271.213	−1025.642	5568.069		0.954	0.950	216.771	0
x_1	−75.384	−150.531	−0.238	0.954				
x_2	−2586.317	−7472.346	2299.711	−0.933				
x_3	63.04	33.656	92.424	0.972				
2010	B	Lower limit	Upper limit	Correlation	R^2	Adjusted R^2	*F*	Sig. *F*
(constant)	3062.861	−4080.925	10206.647		0.952	0.947	203.131	0
x_1	−121.058	−173.589	−68.527	0.885				
x_2	−3571.42	−13318.872	6176.032	−0.955				
x_3	86.666	51.387	121.945	0.949				
2020	B	Lower limit	Upper limit	Correlation	R^2	Adjusted R^2	*F*	Sig. *F*
(constant)	2597.537	−15304.71	20499.784		0.899	0.889	91.937	0
x_1	−321.555	−456.618	−186.491	0.779				
x_2	−2883.213	−25479.339	19712.913	−0.907				
x_3	223.29	119.695	326.885	0.876				

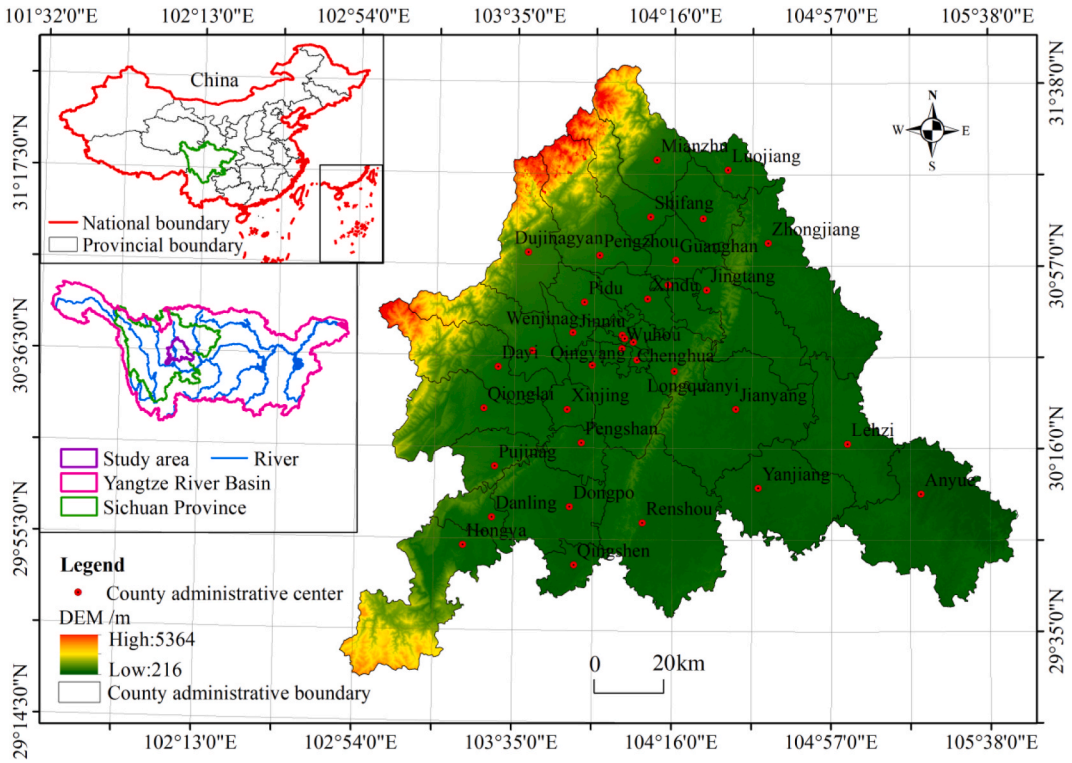


Fig. 1. Location of the study area.

population density in 2010 showed that high and low values of population density were mainly concentrated in one large hot spot and four large cold spots. The central and adjacent areas of Chengdu City formed a significant hot spot distribution, indicating that the population density of these areas showed a high population aggregation distribution at a remarkable level of 95–99% (Fig. 6). The population densities of Anyue in the southeast, Renshou in the south, and Qionglai in the northwest showed a low population aggregation distribution at a significance level of 99% while the population densities of the northeast, east, southwest, and west showed a low population aggregation distribution at a remarkable level of 95%. In 2020, the high and low population density were majorly centralized in one large hot point area, six large cold spot areas. The center, adjacent areas, and peripheral areas of Chengdu formed a significant hotspot distribution; the range was further expanded, indicating that the population density of these areas showed a high population aggregation distribution at a remarkable level of 95–99% (Fig. 6). The population densities of Shifang and Mianzhu-jingyang in the north; Anyue Jianyang, Zhongjiang, and Renshou in the southeast and south; and Qionglai and Dayi in the west showed a low population aggregation distribution at a significance level of 99%. Compared to 2000 and 2010, the range of high population density showed a rapid diffusion trend. Compared with 2010, the density of population in the area of the northeast, east, south, southwest, and west in study region showed a low population aggregation distribution at a significance level of 90%, but the range showed a rapid decreasing trend (Fig. 6).

4.4. Distribution ellipse of population density

Following the calculation method [48], we plotted and calculated the population density standard deviation ellipse (Fig. 7) and its parameters (Table 3). According to Fig. 7 and Table 3, the range of the population density standard deviation ellipse was larger in 2000 than that in 2010 and 2020. The flattening indicates the degree of definiteness and centripetal force of population density changes. The flattening of the ellipse in 2020 was lower than that in 2010 and 2000, indicating that the development of population density in 2020 had a more pronounced trend than those in 2000 and 2010. The change in the spatial rotation angle showed that the rotation angle increased from 125°6′6″ in 2000 to 128°29′52″ in 2020 (Table 3); the generated ellipse direction was consistent with the direction of population aggregation in Chengdu. The gravity centre of for the spatial patterns of population density in the research period generally showed a trend of moving from the southeast to the northwest between 2000 (30°36′4″N, 104°9′21″E) and 2020 (30°38′8″N, 104°4′52″E). The semi-major axis decreased from 53.236 km in 2000 to 33.306 km in 2020 (Fig. 6 and Table 3). Both decreases were relatively large, indicating that the spatial patterns of the population density was increasingly clustered in main direction; the degree of clustering was not large. The semi-minor axis decreased from 42.094 km in 2000 to 29.618 km in 2020. The shortening of long axis and short axis of the ellipse indicated that spatial distribution of the population density was more clustered during this period, dominated by a southeast-to-northwest spatial distribution pattern (Fig. 7).

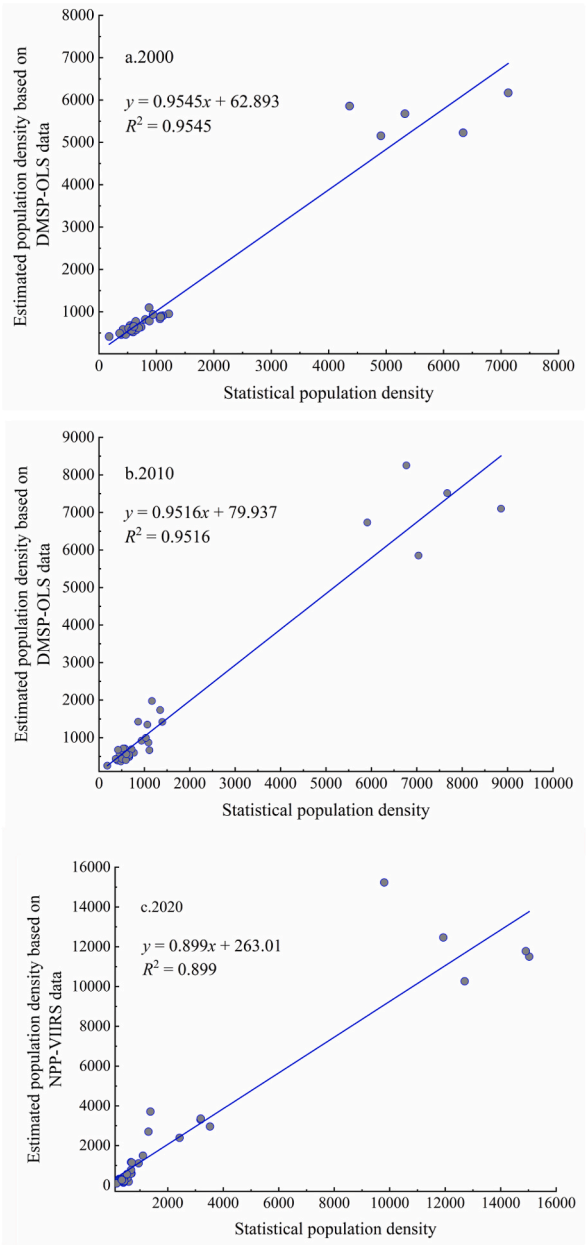


Fig. 2. Accuracy evaluation between the calculated population density and statistics population density at the county scale by linear regression analysis equation (person/km^2).

4.5. Detection factor influence

The impact of various factors on variation in population density was found to be statistically significant ($P < 0.01$). Among the variables examined, tertiary industry, NDVI, and LUDI had the highest *q*-values of 0.5059, 0.4807, and 0.4421, respectively, indicating that they were the most influential factors. Specifically, these variables were able to explain 50%, 48%, and 44% of the changes in population density, respectively, as shown in Table 4. The *q*-values of the PPI, PSI, DCOL, GDPC, and DUC were 0.3966, 0.3895, 0.3334, 0.2849, and 0.2328, respectively; their interpretation level were all above 23%.The *q*-values of the GDP density, landform, AAT, and IEE were 0.1596, 0.1486, 0.1277, and 0.1196, respectively, with explanatory powers above 11%.

Although individual factors, such as elevation, DFO, AAP, DFR, DRO, TOR, DGR, and DWA, had low explanatory power, all less than 0.08, they synergize with other important factors and display a non-linear or mutually reinforcing effect, which has a greater impact on population density changes.

Therefore, the PTI, NDVI, LUDI, PPI, PSI, DCOL, GDPC, and DUC were identified as important factors affecting changes in the

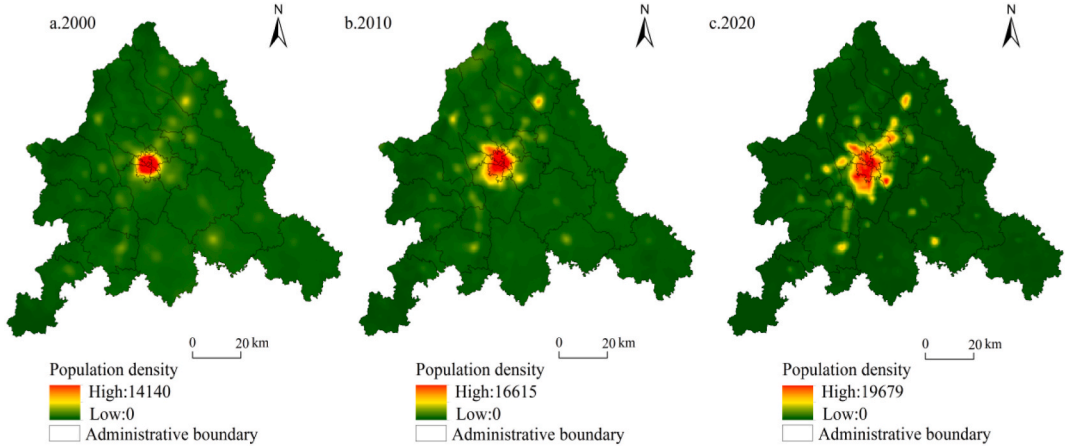


Fig. 3. Spatial patterns of population density using the DMSP-OLS and NPP-VIIRS data (unit: person/km^2).

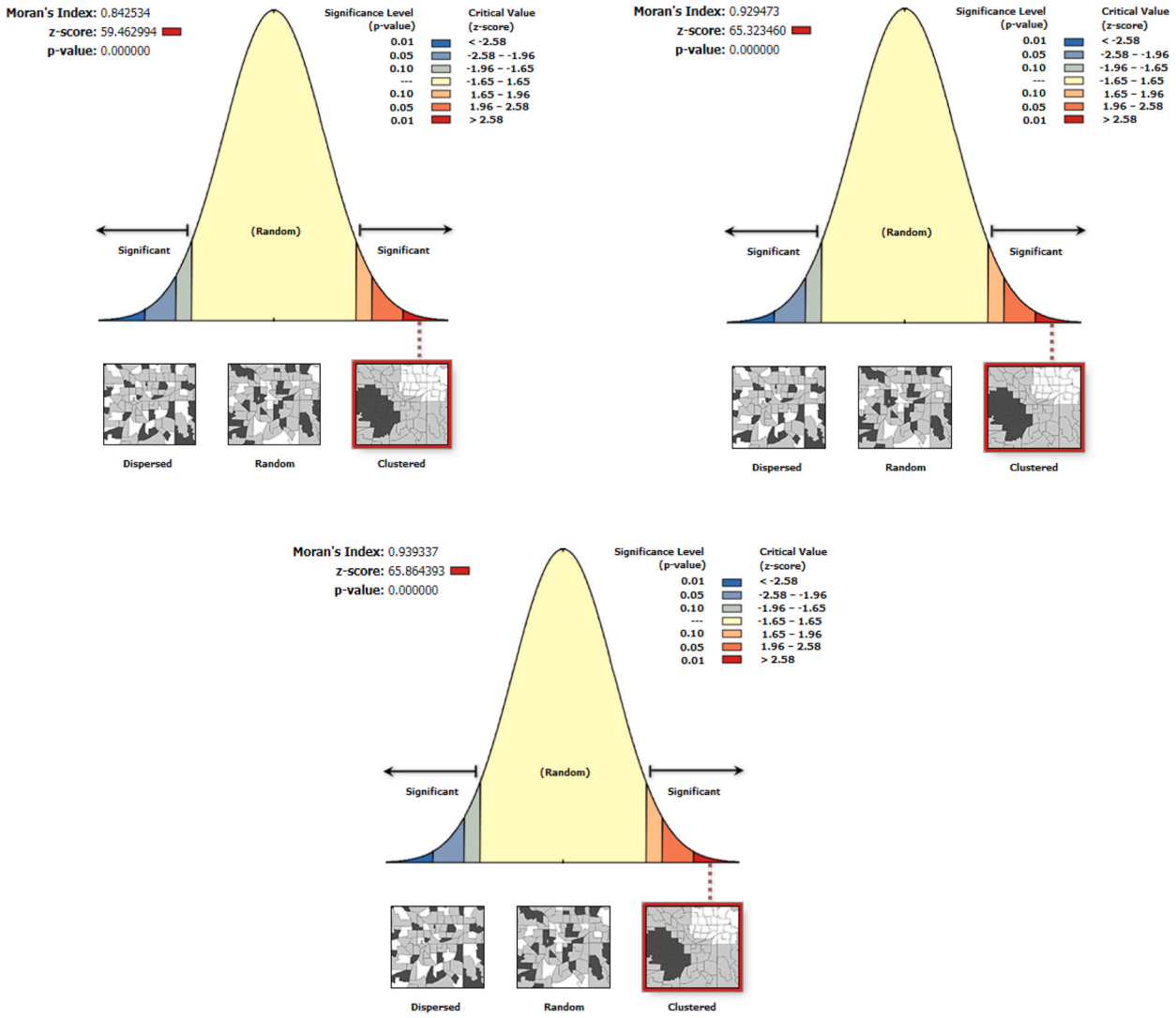


Fig. 4. Results of the spatial auto-correlation analysis of population density (left: 2000; right: 2010; bottom: 2020).

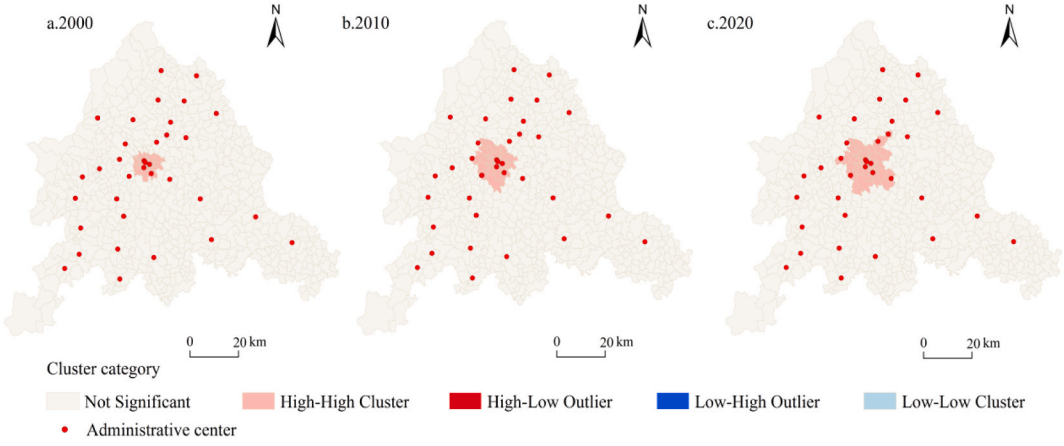


Fig. 5. Spatial clustering of density of population at the township level.

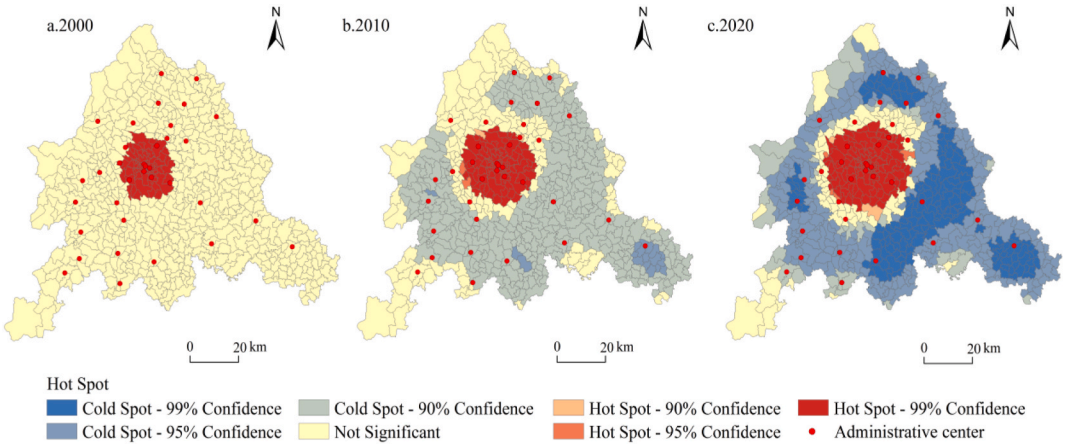


Fig. 6. Hotspots of the population density at the town and township scale.

population density.

4.6. Analysis of interaction between factors

The interplay between the factors demonstrated non-linear enhancing associations and mutually reinforced effects, with the interaction effects primarily boosting each other; no factors acted independently of each other (Fig. 8, Table 5). The diagonal line in Fig. 8 signifies the impact of a single factor, while the *q*-values for 52% and 48% of the interaction factors were either greater or lesser than the *q* values of the single factors, respectively. According to Table 5, the interaction of most factors impacting the population density indicated an enhancement effect on each other, whereas the interaction of a few factors indicated a nonlinear enhancement effect. First, $x_2 \cap x_{19}$ (0.620) > $x_2 \cap x_2$ (0.608) > $x_2 \cap x_5$ (0.570) > $x_2 \cap x_{20}$ (0.569) > $x_2 \cap x_{13}$ (0.563) > $x_2 \cap x_7$ (0.445) > $x_2 \cap x_3$ (0.453), indicating that the interactions between the GDPC and major factors such as the PPI, PSI, PTI, DUC, DCOL, NDVI, and LUDI demonstrated mutually reinforcing effects; Second, $x_3 \cap x_{19}$ (0.692) > $x_3 \cap x_{20}$ (0.656) > $x_3 \cap x_5$ (0.618) > $x_3 \cap x_7$ (0.563) > $x_3 \cap x_4$ (0.560), the PPI demonstrated mutually enhancing effects with major factors such as NDVI, LUDI, PTI, DCOL, DUC, and PSI, indicating strong interplay between these factors. Third, $x_4 \cap x_{20}$ (0.641) > $x_4 \cap x_{13}$ (0.547) > $x_4 \cap x_7$ (0.438), indicating that the interactions between the PSI and the major factors of the NDVI, LUDI, DCOL, and DUC showed mutually reinforcing effects; Fourth, $x_5 \cap x_{19}$ (0.704) > $x_5 \cap x_{20}$ (0.674) > $x_5 \cap x_{19}$ (0.627) > $x_5 \cap x_{13}$ (0.606), showing that the interactions between the PTI and major factors such as the LUDI, NDVI, and DCOL show mutually reinforcing effects. Finally, $x_4 \cap x_{19}$ (0.680) > $x_3 \cap x_{13}$ (0.561), showing that the interactions between the PSI and major factors such as the NDVI, PPI, and DCOL generated a nonlinear increase impact.

4.7. Analysis of significant differences in detection factors

The impact of the PTI on the spatial distribution of population density differed significantly from that of GDP density, GDPC, PPI,

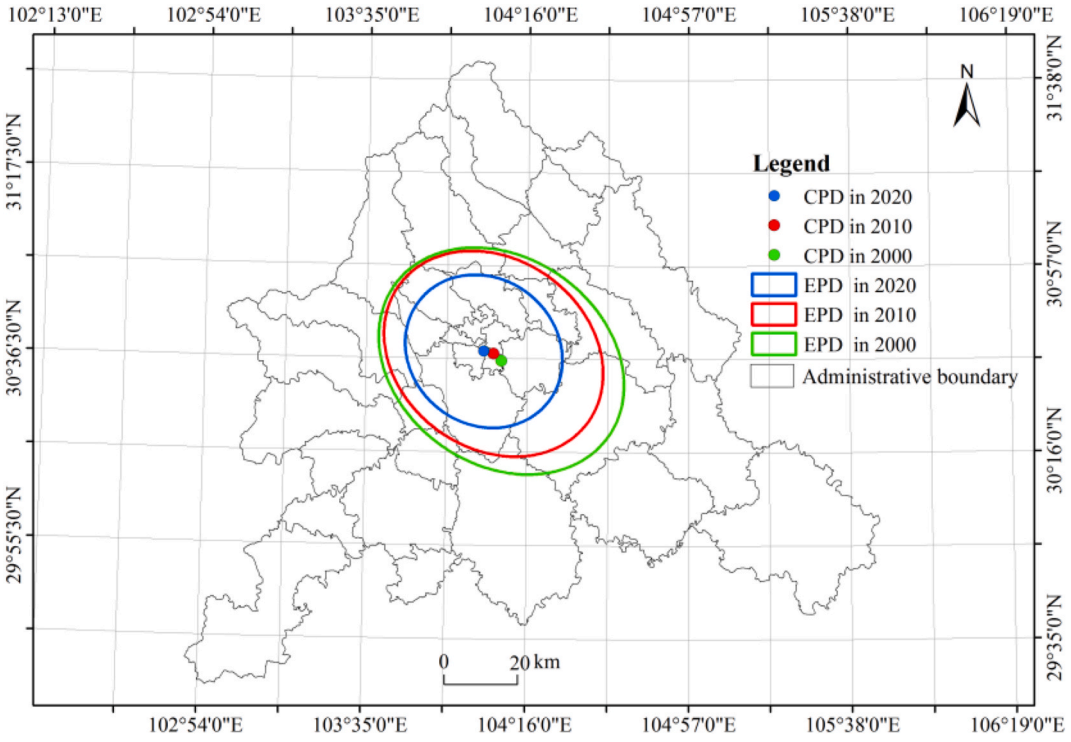


Fig. 7. Center and distribution ellipse of the population density distribution during 2000–2020 (CPD is center of the population density distribution; EPD is ellipse of the population density distribution).

Table 3
Standard deviation ellipse parameters of the population density from 2000 to 2020.

Year	Center *x*	Center *y*	Perimeter/km	Area/km^2	*x* StdDist/km	*y* StdDist/km	Rotation
2000	104°9′21″E	30°36′4″N	300.508	7039.723	53.236	42.094	125°6′6″
2010	104°4′20″E	30°37′38″N	270.460	5726.529	47.401	38.457	125°55′34″
2020	104°4′52″E	30°38′8″N	197.850	3098.937	33.306	29.618	128°29′52″

Table 4
The *q*-values for different impact factors.

Factor	*q* Statistic	*p*-Value	Factor	*q* Statistic	*p*-Value
x_1	0.15957	0.000	x_{11}	0.01322	0.000
x_2	0.28486	0.000	x_{12}	0.01125	0.000
x_3	0.39660	0.000	x_{13}	0.33338	0.000
x_4	0.38948	0.000	x_{14}	0.08007	0.000
x_5	0.50593	0.000	x_{15}	0.02640	0.000
x_6	0.11958	0.000	x_{16}	0.14863	0.000
x_7	0.23280	0.000	x_{17}	0.12766	0.000
x_8	0.03715	0.000	x_{18}	0.05913	0.000
x_9	0.04057	0.000	x_{19}	0.48072	0.000
x_{10}	0.06987	0.000	x_{20}	0.44208	0.000

Note: the *p*-value represents the level of significance of a statistical test. A smaller *p*-value indicates stronger evidence against the null hypothesis, and therefore a higher degree of confidence in the inference that a certain factor or variable has an impact on the dependent variable. *x* has an impact on the dependent variable *y*. $*p < 0.05$ is considered significant, $**p < 0.01$ is considered highly significant.

and PSI. However, no significant difference was evident in the effect on the population density distribution with the remaining factors, including DUC, DCOL, NDVI, and LUDI (Table 6). Table 4 indicates that the influence of NDVI and LUDI on the patterns of population density differs significantly from that of all other factors, except for the PTI and GDP density. Additionally, Table 6 demonstrates that the effect of the LUDI on population density patterns varies notably from that of all other factors, except for the PTI and NDVI. Table 6 reveals that the impact of the PPI and PSI on population density patterns did not differ significantly from that of all other factors, except

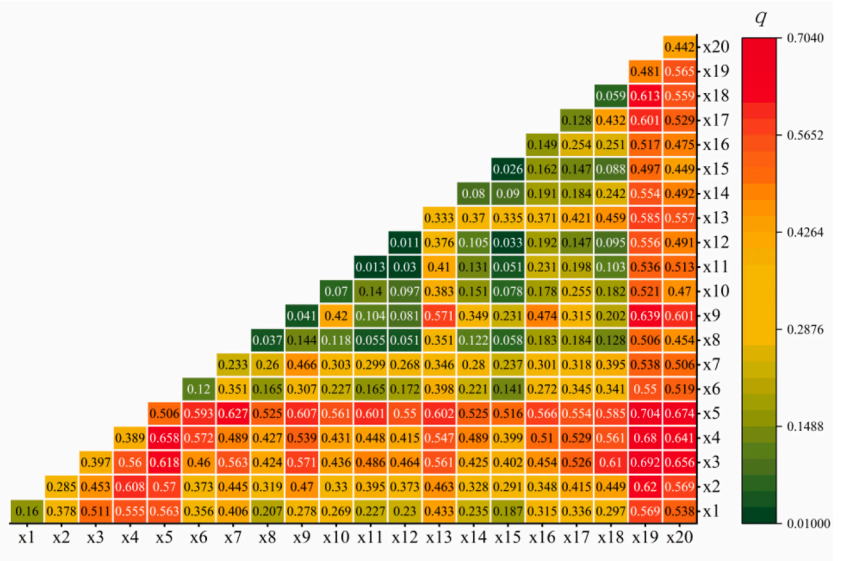


Fig. 8. Thermal diagram of the interaction of influencing factors.

Table 5
Interactions between the major elements that influence changes of the population density.

C	A + B	Result	Interpretation	C	A + B	Result	Interpretation
$x_2 \cap x_3 = 0.453$	$<0.681 = x_2 + x_3$	C > A + B	↑	$x_4 \cap x_7 = 0.438$	$<0.622 = x_4 + x_7$	C < A + B	↑
$x_2 \cap x_4 = 0.608$	$<0.674 = x_2 + x_4$	C > A + B	↑	$x_4 \cap x_{13} = 0.547$	$<0.723 = x_4 + x_{13}$	C < A + B	↑
$x_2 \cap x_5 = 0.570$	$<0.791 = x_2 + x_5$	C > A + B	↑	$x_4 \cap x_{19} = 0.680$	$>0.449 = x_4 + x_{13}$	C > A + B	↑↑
$x_2 \cap x_7 = 0.445$	$<0.518 = x_2 + x_6$	C < A + B	↑	$x_4 \cap x_{20} = 0.641$	$<0.832 = x_4 + x_{20}$	C < A + B	↑
$x_2 \cap x_{13} = 0.563$	$<0.618 = x_2 + x_7$	C > A + B	↑	$x_5 \cap x_7 = 0.627$	$<0.739 = x_5 + x_7$	C < A + B	↑
$x_2 \cap x_{19} = 0.620$	$<0.766 = x_2 + x_8$	C < A + B	↑	$x_5 \cap x_{13} = 0.606$	$<0.839 = x_5 + x_{13}$	C < A + B	↑
$x_2 \cap x_{20} = 0.569$	$<0.727 = x_2 + x_3$	C > A + B	↑	$x_5 \cap x_{19} = 0.704$	$<0.987 = x_5 + x_9$	C < A + B	↑
$x_3 \cap x_4 = 0.560$	$<0.786 = x_3 + x_4$	C > A + B	↑	$x_5 \cap x_{20} = 0.674$	$<0.948 = x_5 + x_{10}$	C < A + B	↑
$x_3 \cap x_5 = 0.618$	$<0.903 = x_3 + x_5$	C < A + B	↑	$x_7 \cap x_{13} = 0.233$	$<0.566 = x_7 + x_{13}$	C < A + B	↑
$x_3 \cap x_7 = 0.563$	$<0.629 = x_3 + x_6$	C < A + B	↑	$x_7 \cap x_{19} = 0.538$	$<0.714 = x_7 + x_{14}$	C < A + B	↑
$x_3 \cap x_{13} = 0.561$	$>0.434 = x_3 + x_7$	C > A + B	↑↑	$x_7 \cap x_{20} = 0.506$	$<0.675 = x_7 + x_{20}$	C < A + B	↑
$x_3 \cap x_{19} = 0.692$	$<0.977 = x_3 + x_8$	C < A + B	↑	$x_{13} \cap x_{19} = 0.585$	$<0.814 = x_{13} + x_{19}$	C < A + B	↑
$x_3 \cap x_{20} = 0.656$	$<0.839 = x_3 + x_9$	C < A + B	↑	$x_{13} \cap x_{20} = 0.557$	$<0.776 = x_{13} + x_{20}$	C < A + B	↑
$x_4 \cap x_5 = 0.618$	$<0.903 = x_4 + x_5$	C < A + B	↑	$x_{19} \cap x_{20} = 0.565$	$<0.923 = x_{19} + x_{20}$	C < A + B	↑

Note: "C" denotes the interplay between two factors, $x_i \cap x_j$; "A + B" denotes the addition of two factor q-values ($q(x_i)+q(x_j)$); "↑" denotes that x_i and x_j enhance each other; "↑↑" denotes a nonlinear enhancement of x_i and x_j.

for the PTI, NDVI, and LUDI. The influence of DCOL on population density distribution differed significantly from that of the PTI, NDVI, GDPD, GDPC, LUDI, DUC, DFR, DRO, DFO, DGR, and DWA. The remaining factors did not exhibit any significant differences in their impact on population density. However, the effects of GDPC, PPI, PSI, and PTI, DCOL, and NDVI were notably distinct on the population density patterns. Nonetheless, there were no significant differences in the effects of the remaining factors. The effects of DUC on population density distribution differed significantly from those of GDP density, LUDI, DCOL, PTI, and NDVI. However, there were no significant differences compared to the effects of the remaining factors.

4.8. Risk test of factors

The risk test outcomes indicate how population density responds to variations in a specific factor. The study found that changes in population density were notably distinct across different levels of the graded factor. Additionally, the results were statistically significant with a confidence level of 95% (Fig. 9a–c).

Based on Fig. 3a, the population density exhibited a declining trend as the PPI and PSI increased, with both factors reaching their maximum values at the first level (0–3.58% and 0–19.07%) of 8133 and 15,058 persons/km^2, respectively. With the PTI, GDP density, and GDPC increased, the population density showed a U-shaped trend, reaching a maximum at the 9th level (75.87–88.56%, 1879.124 $\times$ 10^8–2792.375 $\times$ 10^8 yuan, and 12.745 $\times$ 10^4 to 15.761 $\times$ 10^4 yuan) for 15,437, 12,126, and 7487 persons/km^2, respectively. The population density tended to increase with an increase in real estate investment, reaching a maximum value at the 8th level (148.6872 $\times$ 10^8 to 190.6531 $\times$ 10^8 yuan) for 5547 persons/km^2.

Table 6
Statistical significance of the detection factors at the 95% confidence level.

Factors	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
x_1																				
x_2	*Y*																			
x_3	*Y*	*Y*																		
x_4	*Y*	*Y*	*N*																	
x_5	*Y*	*Y*	*Y*	*Y*																
x_6	*N*	*N*	*N*	*N*	*N*															
x_7	*Y*	*N*	*N*	*N*	*N*	*Y*														
x_8	*N*	*N*	*N*	*N*	*N*	*N*	*N*													
x_9	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*												
x_{10}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*Y*	*Y*											
x_{11}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*										
x_{12}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*									
x_{13}	*Y*	*Y*	*N*	*N*	*N*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*								
x_{14}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*Y*	*Y*	*N*	*Y*	*Y*	*N*							
x_{15}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*						
x_{16}	*N*	*N*	*N*	*N*	*N*	*Y*	*N*	*Y*	*Y*	*Y*	*Y*	*Y*	*N*	*Y*	*Y*					
x_{17}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*Y*	*Y*	*Y*	*Y*	*Y*	*N*	*Y*	*Y*	*N*				
x_{18}	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*Y*	*Y*	*N*	*Y*	*Y*	*N*	*N*	*Y*	*N*	*N*			
x_{19}	*Y*	*Y*	*Y*	*Y*	*N*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*		
x_{20}	*Y*	*Y*	*Y*	*Y*	*N*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*N*	

Note: *Y* implies that there are noteworthy differences between the two factors' NDVI, with a 95% confidence interval. *N* indicates no significant difference between them.

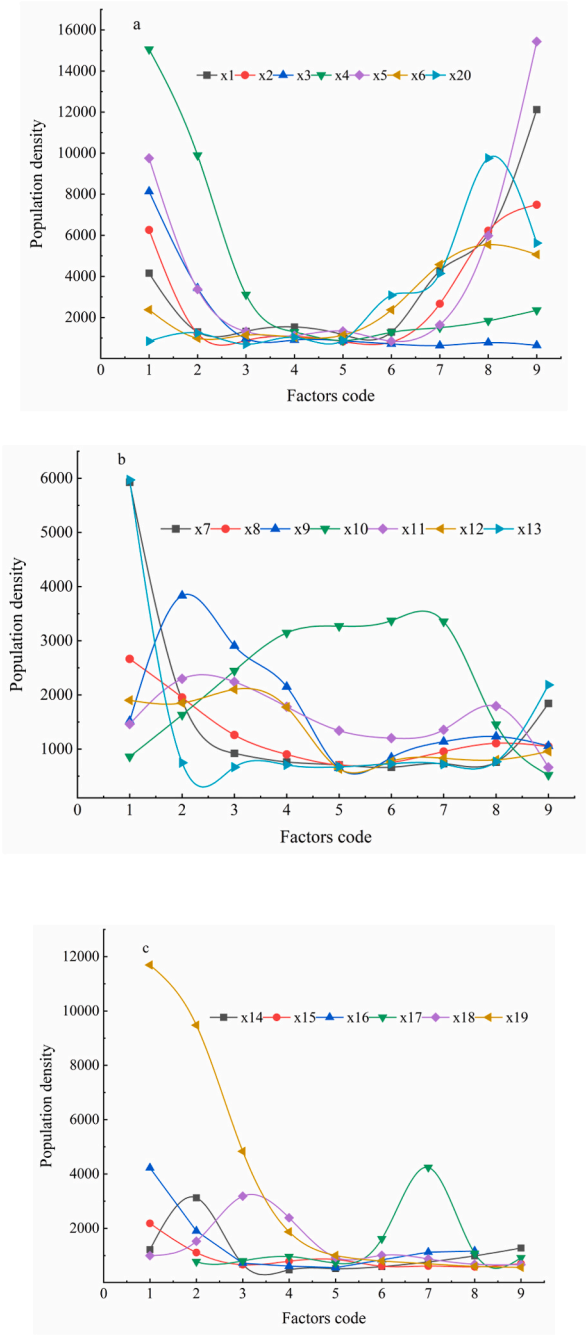


Fig. 9. Population density according to the influence of factors and their levels.

Fig. 3b illustrates that there is a negative correlation between population density and DUC, DRO, and areas of construction land, all of which reached their maximum values at the 1st level (0–6,737, 0–726, and 0–3842 m) for 5,927, 5,971, and 2663 persons/km^2, respectively. The population density showed a changing trend from increasing to decreasing as the DFR, DGR, and DWA increased, reaching maximum values at the 2nd (806–2555 and 1345–3027 m) and third (2823–4720 m) levels for 3,833, 2,298, and 2101 persons/km^2, respectively. The population density tended to increase with an increasing DFO, reaching a maximum of 3370 persons/km^2 at the 6th level (3722–4925 m).

According to Fig. 3c, the population density tended to decrease with changes in geomorphic and TOR and an increasing NDVI, reaching maximum values at the 1st level (plain, 0–0.08, and 0.42) for 4,227, 2,183, and 11,691 persons/km^2, respectively. The population density tended to increase and then notably decrease with changes in the AAT and LUDI, reaching maxima at the 7th (16.72–17.12 °C) and 8th (0–114.29) levels for 4242 and 9757 persons/km^2, respectively. As elevation and AAP change, the

population density showed an increasing to decreasing trend, with factors reaching their maximum values at the 2nd (465–618 m) and 3rd (564.79–602.44 mm) levels for 3122 and 3183 persons/km^2, respectively.

5. Discussion

Our simulation of population density features was based on established comparative DMSP/OLS and NPP-VIIRS nightlight data. This study has particular reference significance for population density estimation via remote sensing at the metropolitan area scale. However, certain problems require further investigation [40]. First, the two types of nightlight data used in this study were constructed at the county and raster scales. The ability to fit two types of nightlight and explore the optimal scale remains a problem [40,51]. Therefore, coordinating the advantages of the two types of data while accounting for population growth and increases in nighttime light brightness is a significant challenge in simulating population densities using nighttime light data [41]. Secondly, the rise in nighttime light intensity follows a roughly linear pattern, but since nighttime light data do not capture the economic activity of human society during daylight hours, they are incomplete [40,52]. Accurate measurement of changes in population activity using nightlight data requires more in-depth research [41].

The coupling of multiple data sources and technologies, such as remote sensing and GIS, has transformed the manner in which the distribution of human populations is examined in space and time [24]. At present, available statistical data use 'year' as the unit of time, which yields a low temporal resolution for spatial population data based on statistical data simulations. In this study, processing nighttime light data based on NDVI data helped to address the issue of spatial spillover of light [53], eliminates the interference of non-residential light in population simulations, and improves the population difference between grids with the same or similar light values. This study overcomes the limitations of traditional statistical data, which may suffer from inconsistencies in calibre, missing data, and low update frequency [54,55], by providing data sources that allow for the analysis of the spatial patterns of various economic indicators and enable fine-grained assessments of population activity levels at different scales [56].

A substantial correlation exists between regional population changes and the economic development level [57]. As the economy advances, wages and living conditions improve, making the region more appealing, resulting in increased population. The findings align with the outcomes documented by Qi et al. [57], Liu et al. [58], and Wang et al. [59]. The PTI, and PSI, as well as GDPC, were found to be highly correlated with population density, with correlation coefficients of 0.968, 0.894, 0.918, and 0.813, respectively. All fitted equations were quadratic polynomial equations. With an increase in the PTI, good infrastructure and comfortable living conditions also affect the investment environment of the city, leading to population aggregation and an increase in population density. According to the fitted equation, when the PTI is below 55.06%, population density exhibits a decreasing trend at first, but once the proportion exceeds 63.62, it sharply increases. When the PTI, and PSI increases, the population density decreases. With GDPC exceeding 84,834 yuan per person, population density continues to increase. Therefore, economic development brings about further population aggregation, resulting in economic aggregation [57].

The extent of land use measures the influence of human activities on the land system. It quantitatively describes the overall level and changes in land use, serving as a critical indicator for assessing the land use status in a specified region. Population density is highly correlated with the LUDI, with a correlation coefficient of 0.627. The fitted equation is a power function equation. As land use practices continue to evolve and cultivated land is transformed into construction areas (such as industrial, mining, transportation, and residential areas), the population density on construction land substantially increased. These findings align with the outcomes documented by Guo et al. [60].

The city serves as a hub for population concentration in the region, which has an economic agglomeration effect [61]. The development of urban economic agglomeration is accompanied by a large-scale influx of population [62]. The DUC reflects the radiation and driving effect of the economic and transportation infrastructure, which attracts migrants and influences the population density. Population density has a high correlation (−0.667) with DUC. With an increase in DUC, the population density shows a downward trend. These findings align with those researched by Zhao et al. [61]. An increase in human mouth density has a negative effect on NDVI [63]. As the population concentration increases, there is a notable reduction in NDVI. The correlation coefficient between population density and NDVI was −0.821, indicating a strong negative correlation. This suggests that human activities have a constraining impact on NDVI. Areas with high population densities show strong human activity and serious damage to vegetation resources [64]. Therefore, in areas with high population density and intense human activities, NDVI showed a marked decrease, which was negatively correlated, in line with findings by Xu et al. [63] and Zhao et al. [64].

The proximity of construction land has a significant influence on urban land use as it represents the cost of land development and impacts urban infrastructure and real estate, thereby influencing population concentration. Population density had a strong correlation (−0.635) with the DCOL. As the DCOL increased, a decrease in population density was observed. However, the relationship between the size of construction land and population patterns is not a simple linear correlation, as found by Dong et al. [65].

The results of this study indicate that geographic detectors can effectively identify drivers and capture spatial heterogeneity. While there are numerous classification algorithms, such as *K*-means and SOM, available for partitioning, statistical methods for spatial differentiation remain limited [36]. Currently, the main methods for spatial differentiation include spatial anisotropy measurements and factor analyses of geographic detector *q* statistics [36]. When compared to conventional approaches such as principal component analysis and classical regression models [11,26,55], the geographic detector approach has no linearity assumptions or clear physical implications and is better able to reveal the spatial heterogeneity of driving factors.

However, this study has some limitations. Due to the varied features of population distribution across different areas considered in this study, the applicability and accuracy for rural areas and the urban–rural fringe populations requires further exploration [21,23]. Using the same spatial model for the study area resulted in estimation errors [53]. Nighttime light data also have limitations;

background noise must be accounted for more effectively [66].

The simulation of population distribution using nighttime light data can be significantly impacted by non-residential lights, as well as the illumination from commercial and industrial areas, and roads, all of which affect the extraction of population distribution information [56]. It is worth mentioning that the accuracy of population distribution simulations using nighttime light data may be affected by the presence of non-residential lights, commercial and industrial areas, and road networks. In large-and medium-sized cities with dense populations, nighttime light data are saturated and do not reflect the differences in population distributions within the city. In contrast, in rural and mountainous areas with sparse populations, nighttime lights are dim, which makes it difficult to accurately express the population distribution [56]. Future studies on population spatialization should focus on improving the accuracy of simulation and the spatial-temporal resolution of population spatial data, by exploring new data sources and utilizing advanced techniques [56].

6. Conclusions

We employed DMSP/OLS/NPP–VIIRS and NDVI data to estimate and map population density, while also using geographic detectors to analyze the individual and interactive effects of various factors on population density. This facilitated identification of the most suitable features for each major factor. Understanding the changes and driving mechanisms of population density has a positive influence in the rational planning of population spatial layouts.

The model developed in this study demonstrated optimal performance in predicting population density, with a high R^2 value (>0.889, $P = 0.01$) indicating a strong correlation with the estimated population density. This suggests that NDVI, DMSP/OLS/NPP, and VIIRS are reliable predictors for the spatial patterns of urban population density.

Our analysis showed that the population density exhibited clear aggregating features and positive spatial correlations across the study area. From 2000 to 2020, the Global *Moran's I*-index and *Z*-scores ranged from 0.843 to 0.939 and from 59.462 to 65.864, respectively, indicating a significant increase. The high population density values were found to cluster in central Chengdu, adjacent and peripheral districts, forming hotspots of local spatial clustering. Moreover, the areas with high population density expanded rapidly over time.

Our analysis revealed that several factors were influential in the changes observed in population density, including the PPI, PTI, and PTI, NDVI, LUDI, DCOL, GDPC, and DUC. These factors exhibited interactive and nonlinear effects on population density, with mutually enhancing synergies.

Author contribution statement

Guangjie Wang: conceived and designed the experiments; Performed the experiments; analyzed and interpreted the data; Wrote the paper.

Wenfu Peng: conceived and designed the experiments; Analyzed and interpreted the data; contributed reagents, materials, analysis tools or data; Wrote the paper.

Lindan Zhang: analyzed and interpreted the data.

Funding statement

This study was supported by the Humanities and Social Science Research Foundation of the Ministry of Education, China under grant number (17YJA850007).

Data availability statement

Data will be made available on request. Landsat data were acquired from the USGS (https://earthexplorer.usgs.gov) and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn).

Additional information

No additional information is available for this paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors express their appreciation to the editors and anonymous referees for their helpful feedback, which contributed to the improvement of the manuscript. The authors would like to acknowledge the USGS and RESDC for providing the Landsat data used in this study. It should be noted that the funding sources did not participate in the collection, analysis, interpretation, or writing of the

data, nor in the decision to submit this study for publication.

References

[1] United nations, annual report state of World population. https://www.unfpa.org/sites/default/files/pub-pdf/SoWP2021, 2021. Report-EN web.3.21 0.pdf.
[2] Q.Z. Mao, Y. Long, K. Wu, Spatio-temporal evolution of population density and spatial pattern of urbanization in China from 2000 to 2010, Urban Plann. 39 (2) (2015) 38–43.
[3] H. Guo, Z. Fang, Study on spatial characteristics,changing trend and influencing factors of population agglomeration, Rev. Econ. Res. 2973 (5) (2021) 86–101.
[4] M.C. Zhu, J.R. Gu, Y.D. Zhao, Spatial heterogeneity analysis of population and economic factors, Environ. Sustain. Dev. 43 (3) (2018) 18–22.
[5] Y. Xu, Y.M. Song, J.X. Cai, H. Zhu, Population mapping in China with Tencent social user and remote sensing data, Appl. Geogr. 130 (2021), 102450, https://doi.org/10.1016/j.apgeog.2021.102450.
[6] L.T. Wang, S.X. Wang, Y. Zhou, W.L. Liu, Y.F. Hou, J.F. Zhu, F.T. Wang, Mapping population density in China between 1990 and 2010 using remote sensing, Remote Sens. Environ. 210 (2018) 269–281.
[7] C. Zeng, Y. Zhou, S. Wang, F. Yan, Q. Zhao, Population spatialization in China based on night-time imagery and land use data, Int. J. Rem. Sens. 32 (24) (2011) 9599–9620.
[8] F. Checchi, B.T. Stewart, J.J. Palmer, C. Grundy, Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations, Int. J. Health Geogr. 12 (4) (2013) 1–12, https://doi.org/10.1186/1476-072X-12-4.
[9] X.M. Gao, X.H. Yang, B.R. Chen, L. Lin, Spatial simulation of population in Bohai Rim based on stochastic forest model, J. Geoinf. Sci. 24 (6) (2022) 1150–1162.
[10] Y.J. Zhang, X.X. Wang, Population spatial distribution in China from the perspective of urban agglomeration, Popul. Econ. 240 (3) (2020) 1–13.
[11] D.S. Yan, W. Sun, Y. Wang, S.Y. Xu, Change in distribution and growth shifts of population in the Yangtze River Delta and influencing factors, Prog. Geogr. 39 (12) (2020) 2068–2082.
[12] T. Liu, R.X. Peng, Y. Zhuo, G.Z. Cao, China's changing population distribution and influencing factors: insights from the 2020 census data, Acta Geograph. Sin. 77 (2) (2022) 381–394.
[13] J. Liu, X.L. Zhang, H.B. Chen, The impact of urban population agglomeration on the high-quality economic development in the Yangtze River Delta, East China Econ. Manage. 36 (2) (2022) 12–20.
[14] M. Bakillah, S. Liang, A. Mobasheri, J. Arsanjani, F. Zip, Fine-resolution population mapping using OpenStreetMap points-of-interest, Int. J. Geogr. Inf. Sci. (9) (2014) 1940–1963.
[15] L. Wang, S. Wang, Z. Yi, W. Liu, Y. Hou, J. Zhu, F. Wang, Mapping population density in China between 1990 and 2010 using remote sensing, Remote Sens. Environ. 210 (2018) 269–281.
[16] J.B. Yuan, Y.W. Cao, F.Z. Ni, H.X. Qiu, C.S. Zhou, A study on the spatial pattern of county population agglomeration and the spatial heterogeneity of its influencing factors in China, Geogr. Geo-Inf. Sci. 36 (3) (2020) 25–33.
[17] J. Lin, R.G. Cromley, Evaluating geo-located Twitter data as a control layer for areal interpolation of population, Appl. Geogr. 58 (2015) 41–47.
[18] F. Qiu, R. Cromley, Areal interpolation and dasymetric modeling, Geogr. Anal. 45 (3) (2013) 213–215.
[19] H.L. You, J. Yang, B. Xue, X.M. Xiao, J.H. Xia, J. Cui, X.M. Li, Spatial evolution of population change in Northeast China during 1992–2018, Sci. Total Environ. 776 (2021), 146023, https://doi.org/10.1016/j.scitotenv.2021.146023.
[20] Y.J. Zou, Q.W. Yan, J. Huang, F. Li, Modeling the population density of Su-xi-Chang region based on Luojia-1A nighttime light image, Resour. Environ. Yangtze River Basin 29 (5) (2020) 1086–1094.
[21] J.J. Wang, Z.Y. Li, Modeling population density of Wuhan city using landsat OLI and NPP-VIIRS, Areal Res. Dev. 38 (6) (2019) 146–151.
[22] X. Chen, W. Nordhaus, A test of the new VIIRS lights data set: population and economic output in Africa, Rem. Sens. 7 (4) (2015) 4937–4947.
[23] Y. Xu, Y.M. Song, J.X. Cai, H. Zhu, Population mapping in China with Tencent social user and remote sensing data, Appl. Geogr. 130 (2021), 102450, https://doi.org/10.1016/j.apgeog.2021.102450.
[24] P. Deville, C. Linardc, S. Martin, M. Gilbert, F.R. Stevens, A.E. Gaughan, V.D. Blondel, A.J. Tatem, Dynamic population mapping using mobile phone data, Proc. Natl. Acad. Sci. U.S.A. 111 (45) (2014) 15888–15893.
[25] L. Wang, Z.M. Feng, Y.Z. Yang, Z. You, The change of population density and its influencing factors from 2000 to 2010 in China on county scale, Acta Geograph. Sin. 69 (12) (2014) 1790–1798.
[26] T.T. Yu, Y.X. Song, F.L. Hao, R. A, Space pattern evolution of population distribution and the driving factors in Northeast China, Sci. Geogr. Sin. 37 (5) (2017) 709–717.
[27] Y. Feng, Spatial effects of population and economic agglomeration-A case study of Chengdu-Chongqing Shuangcheng Economic Circle, J. Southwest Univ. Nat. (Human. Social Sci. Ed.) (2) (2022) 127–135.
[28] C.M. Du, L.M. Jiao, G. Xu, Spatio-temporal patterns and driving factors of urban population density change of prefecture-level cities in China during 2006-2016, Trop. Geogr. 38 (6) (2018) 791–798.
[29] R.H. Mi, X.D. Gao, Spatial Econometric analysis of the influencing factors of population distribution in northwest China, Popul. Econ. 235 (4) (2019) 65–78.
[30] Q. Luo, B.B. Wang, X.S. Fan, X.J. Li, Spatio-temporal change of population distribution and its influencing factors in a poor mountainous county: a case study of Song County, Henan Province, Prog. Geogr. 39 (7) (2020) 1073–1084.
[31] J. Luo, P.J. Shi, X.B. Zhang, Spatiotemporal characteristics of population distribution in the upper reaches of the Yellow River: a multi-dimensional approach in the Lanzhou-Xining urban agglomeration, Resour. Sci. 42 (3) (2020) 474–485.
[32] W.F. Peng, T.T. Kuang, S. Tao, Quantifying influences of natural factors on vegetation NDVI changes based on geographical detector in Sichuan, western China, J. Clean. Prod. 233 (2019) 353–367.
[33] J.F. Wang, C.D. Xu, Geodetector:Principle and prospective, Acta Geograph. Sin. 72 (1) (2017) 116–134.
[34] J.F. Wang, X.H. Li, G. Christakos, Y.L. Liao, T. Zhang, X. Gu, X.Y. Zhang, Geographical detectors- based health risk assessment and its application in the neural tube defects study of the Heshun region, China, Int. J. Geogr. Inf. Sci. 24 (1) (2010) 107–127.
[35] J.F. Wang, T.L. Zhang, B.J. Fu, A measure of spatial stratified heterogeneity, Ecol. Indicat. 67 (2016) 250–256.
[36] J.F. Wang, C.D. Xu, Geodetector:Principle and prospective, Acta Geograph. Sin. 72 (1) (2017) 116–134.
[37] J.F. Wang, X.H. Li, G. Christakos, Y.L. Liao, T. Zhang, X. Gu, X.Y. Zhang, Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China, Int. J. Geogr. Inf. Sci. 24 (1) (2010) 107–127.
[38] J.F. Wang, Y. Ge, L.F. Li, B. Meng, J.L. Wu, Y.C. Bo, S.H. Du, Y.L. Liao, M.G. Hu, C.D. Xu, Spatio-temporal data analysis methods of geography, Acta Geograph. Sin. 69 (9) (2014) 1326–1345.
[39] N. Levin, C.C.M. Kyba, Q.L. Zhang, A. Sánchez de Miguel, M.O. Román, X. Li, B.A. Portnov, A.L. Molthan, A. Jechow, S.D. Miller, Z.S. Wang, R.M. Shrestha, C. D. Elvidge, Remote sensing of night lights: a review and an outlook for the future, Remote Sens. Environ. 237 (1) (2020), 111443, https://doi.org/10.1016/j.rse.2019.111443.
[40] J.C. Zhao, G.X. Ji, Y.L. Yue, Z. Lai, Z.Y. Chen, D. Yang, X. Yang, Z. Wang, Spatio-temporal dynamics of urban residential CO^2 emissions and their driving forces in China using the integrated two nighttime light datasets, Appl. Energy 235 (2019) 612–624.
[41] H.B. Du, W. Wei, X.Y. Zhang, X.P. Ji, Spatio-temporal evolution and influencing factors of energy-related carbon emissions in the Yellow River Basin: based on the DMSP/OLS and NPP/VIIRS nighttime light data, Geogr. Res. 40 (7) (2021) 2051–2065.
[42] Z.S. Xu, Y.M. Xu, Study on the spatio-temporal evolution of the Yangtze River Delta urban agglomeration by integrating DMSP/OLS AND NPP/VIIRS nighttime light data, J. Geo-inf. Sci. 23 (5) (2021) 837–849.
[43] R. Fensholt, K. Rasmussen, T.T. Nielsen, C. Mbowd, Evaluation of earth observation based long term vegetation trends: intercomparing NDVI time series trend analysis consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT data, Remote Sens. Environ. 113 (9) (2009) 1886–1898.

[44] W.F. Peng, G.J. Wang, J.M. Zhou, X.L. Xu, H.L. Luo, J.F. Zhao, C.J. Yang, Dynamic monitoring of fractional vegetation cover along Minjiang River from Wenchuan County to Dujiangyan City using multi-temporal landsat 5 and 8 images, Acta Ecol. Sin. 36 (7) (2016) 1975–1988.
[45] G.J. Wang, W.F. Peng, J.Y. Xiang, L.N. Ning, Y.N. Yu, Modelling spatiotemporal carbon dioxide emission at the urban scale based on DMSP-OLS and NPP-VIIRS data: a case study in China, Urban Clim. 46 (2022), 101326, https://doi.org/10.1016/j.uclim.2022.101326.
[46] K. Shi, B. Yu, Y. Huang, Y. Hu, B. Yin, Z. Chen, L. Chen, J. Wu, Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: a comparison with DMSP-OLS data, Rem. Sens. 6 (2014) 1705–1724.
[47] J.C. Zhao, Y.L. Chen, G.G. Ji, Z. Wang, Residential carbon dioxide emissions at the urban scale for countylevel cities in China: a comparative study of nighttime light data, J. Clean. Prod. 180 (2018) 198–209.
[48] L. Anselin, Local indicators of spatial association, Geogr. Anal. (1995) 93–101.
[49] Y.G. Chen, Development and method Improvement of spatial autocorrelation theory based on Moran statistics, Geogr. Rese. 28 (6) (2009) 1449–1463.
[50] P. Moran, The interpretation of statistical maps, J. Roy. Stat. Soc., Ser. B (Methodol.), London Ser. B 10 (1948) 243–634.
[51] J. Yang, B. Xue, X.M. Xiao, J.H. Xia, C. Jin, X.M. Li, Spatial evolution of population change in Northeast China during 1992–2018, Sci. Total Environ. 776 (2021), 146023, https://doi.org/10.1016/j.scitotenv.2021.146023.
[52] Z.W. Huang, S.Y. Li, F. Gao, F. Wang, J.Y. Lin, Z.L. Tan, Evaluating the performance of LBSM data to estimate the gross domestic product of China at multiple scales: a comparison with NPP-VIIRS nighttime light data, J. Clean. Prod. 328 (2021), 129558.
[53] L. Zhao, Z.Q. Zhao, Projecting the spatial variation of economic based on the specific ellipses in China, Sci. Geogr. Sin. 34 (8) (2014) 979–986.
[54] X. Chen, W.D. Nordhaus, Using luminosity data as a proxy for economic statistics, Proc. Natl. Acad. Sci. U.S.A. 108 (21) (2011) 8589–8594.
[55] J.V. Henderson, A. Storeygard, D.N. Weil, Measuring economic growth from outer space, Am. Econ. Rev. 102 (2) (2012) 994–1028.
[56] C.N.H. Doll, J.P. Muller, C.D. Elvidge, Night-time imagery as a tool for global mapping of socioeconomic parameters and greenhouse gas emissions, AMBIO A J. Hum. Environ. 29 (3) (2000) 157–162.
[57] W. Qi, S. Liu, Z. Liu, The novel pattern and driving factors of population spatial distribution on both sides of the "Hu Line" based on seventh census in China, Acta Geograph. Sin. 77 (12) (2022) 3023–3040.
[58] T. Liu, R.X. Peng, Y.X. Zhuo, G.Z. Cao, The evolution and influencing factors of population distribution pattern in China from 2000 to 2020, Acta Geograph. Sin. 77 (2) (2022) 381–394.
[59] L. Wang, Z.M. Feng, Y.Z. Yang, Z. You, Population density changes and its influencing factors in different regions of China from 2000 to 2010, Acta Geograph. Sin. 69 (12) (2014) 1790–1798.
[60] Z.H. Guo, F. Chen, X.L. Liu, Z.X. Dai, Method of improving spatial accuracy of population grid data based on land use type, Bull. Surv. Mapp. 66 (12) (2021) 66–70.
[61] D.L. Yang, Z.C. Ren, P.G. Li, Study on the influence of population density on Talent Agglomeration in provincial cities of China, Popul. J. 42 (4) (2020) 82–92.
[62] N. Baum Snow, R. Pavan, Understanding the city size wage Gap, Rev. Econ. Studies79 (1) (2012) 88–127.
[63] Y. Xu, C. Zhao, S.Q. Dou, W.Q. Hao, Z.W. Zheng, J.L. Jing, Temporal and spatial variation of vegetation NDVI and its correlation with population density in Bohai Rim from 2000 to 2020, Bull. Soil Water Conserv. 42 (2) (2022) 264–274.
[64] Y.B. Long, X.D. Li, D.Y. Cheng, Changes of NDVI and its relationship with population in karst canyon in Guizhou, Ecol. Sci. 41 (5) (2022) 134–143.
[65] Z.Y. Dong, D.Q. Xue, Evolution of rural construction land and its relation with population in China, Geogr. Geo-Inf. Sci. 38 (5) (2022) 96–103.
[66] X. Chen, L. Wang, L. Tong, S. Sun, X. Yue, S. Yin, L. Zheng, Mode selection of China's urban heating and its potential for reducing energy consumption and CO^2 emission, Energy Pol. 67 (2014) 756–764.