

Research on characteristics in the spatial distribution of the whole county population of Sichuan province, China via ARCGIS analysis

Guiyao Zhou¹, Yanyou Wu^{1*}, Xianjian Xie², Deke Xing¹, Fang Fan³, Rui Yu¹

¹Key laboratory of Modern Agricultural Equipment and Technology, Ministry of Education & Jiangsu province, Jiangsu University, Zhenjiang, China 212013; ²Geography and Resource Science College, Neijiang Normal University, Neijiang, China 641110; ³Software College, Chongqing University, Chongqing, China 401331

¹jdzhouguiyao@163.com, ^{1*}yanyowu@ujs.edu.cn, * Corresponding author

Abstract:

2000 and 2010 year census population data of each county of Sichuan Province, China has been chosen as research objects in our research. In this study, different spatial fitting model of geostatistics modules which are based on Arcgis Software has been used to discuss the fitting effects of population density data. The best spatial distribution fitting model was obtained which has been used to make optimal Kriging interpolation. The factors affecting Sichuan Province Population change have also been analyzed. Results showed that the overall distribution of Sichuan province population has uneven spatial distribution and circular style. Counties population in space formed two macroscopic distribution patterns: One is the Western Sichuan-intensive area and the other one is basins gathering area. Time distribution of different density region has changed over time, especially the population of Chengdu economic region and the Southern Sichuan economic region which are growing rapidly. The whole population in Sichuan province shows an increasing trend from west to east. Natural conditions, policies and economic conditions are the most important factors that influence the Spatial and Temporal distribution regularity of population of Sichuan Province.

Keywords: Geostatistics; Sichuan Province; Distribution of population; Spatial interpolation model

1.Introduction

Population, Resources, Environment, Energy and Food problems are five major issues that the world is facing nowadays, with the population problem being the core issue (CAS, 2000). The spatial distribution of population is characterized by the combination of results of geographical conditions, climate suitability, degree of district economy and social policy orientation and others. The traditional sense of demographic data comes from national census every 10 years. It takes administrative areas such as counties and cities as statistics unit. However, in practice there often exists some problems such as low spatial resolution data and administrative boundaries not coinciding with population boundaries. Regional population data can usually only provide discrete and limited data sample. In recent years, comparative studies on spatial distribution of population density were carried out by many authors (Feng et al., 2002; Christopher et al., 1999). Suggesting that people's research methods on this subject of spatial distribution of population data have gradually changed from traditional statistics method to diverse space data analysis method. The emergence of various geostatistical analysis software such as Arcgis Series, GS +, etc has created a broad application of conditions for spatial distribution statistics of population. There are various geostatistical models such as Gaussian, exponential, spherical models etc.

However, on processing data spatialization, the selection of best space model directly relate to the quality of spatial distribution of statistical data. In this paper, the geostatistics methods have been made use of and county has been used as administrative units to analyze and discuss the fifth and sixth national census population density data of all 181 counties and cities in Sichuan Province. The best models were chosen to simulate the population densitiys' spatial distribution in Sichuan Province, and kriging interpolation methods have been used to simulate the best trend surface of spatial distribution of population density. Finally, the dynamic characteristics of population distribution of each region have been explored within two censuses period.

2. Materials and methods

2.1 Study area

Sichuan Province is situated in the south west of China, upstream of Yangtze River and surrounded by Chongqing, Yunnan, Tibet, Shaanxi and other provinces. Its geographical coordinates are $97^{\circ}21'—108^{\circ}31'E, 26^{\circ}03'—34^{\circ}19' N$. Sichuan covers an area of $48.5 \times 10^4 km^2$, ranking as the fifth largest Province in Western China. It mainly constitutes of plateaus, mountains, hills and other topographies. Western topography is high, while Eastern topography is lower. The Study area is shown in Figure 1. It is dominated by a subtropical climate. In year 2000, the population of Sichuan Province was nearly 87 million, accounting to 23.7% of Western China Population. Sichuan Province consists of Chengdu, Mianyang, Ya'an, Wenchuan and other 181 counties and cities. It is one of the richest regions with important national resources, and also an important ecological barrier to the upstream of Yangtze River. The economy in this area is flourishing, often ranked first in Western China. Sichuan has a provincial situation with "large population, weak economic foundation, and under development." Due to the differences in natural, historical economic conditions, policy

differences and economic disparities, the economic spatial differences are still apparent. Differences between economic regions and differences in city and outside coexist at the same time (Zhang and Ren, 2011).

2.2 Data sources and data analysis

Research data include following parts: (1) Sichuan digitized maps, which were digital scan on Sichuan 1:100000 base map by Arcgis Software. (2) Administrative Region boundary data. (3) The total counties, cities population and administrative area data in Sichuan in year 2000 and 2010. These data comes from 《Sichuan Statistical Yearbook 2000, 2010》. Due to the accelerating urbanization, the original administrative division's settings do not suit the pace of development. Therefore, the local government made some adjustments which lead to changes in administrative boundaries of many counties during this decade. For example, the old three districts in Pengxi County have been changed into Daying County. In order to facilitate comparison and analysis process, the fifth census data has been used to do the appropriate conversion, which is based on the sixth census of the administrative unit divisions. In the first place, the 2000 and 2010 year population data was fed into the county's population density database and the equation which is $Population\ density = total\ population / total\ area$ was used to calculate out 2000 and 2010 year population density of all the 181 counties and cities.

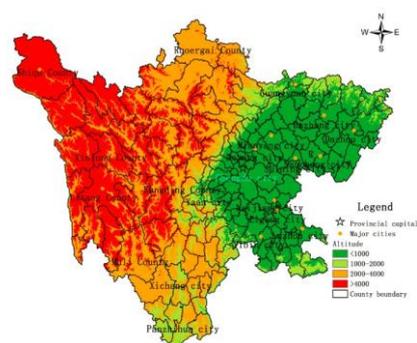


Figure 1. Study area

2.3 Research methods

Geostatistics, also called the geostatistical method, has been put forward by the famous French statistician, G. Matheron after doing a large number of theoretical studies on the basis of the gradual formation of a new branch of statistics. It is based on regionalized variable and use of variograms to research natural phenomenon that not only has randomness but also structural or spatial correlation and dependence. Geostatistics is generally based on the semivariogram and kriging interpolation methods to simulate and fit the spatial modes.

Geostatistical semivariogram mainly include Circular, spherical, Gaussian, Exponential and other models. Arcgis (9.3vesion) includes 11 kinds of models. Each model primarily through nugget (CO), partial sill value (C), sill (CO+C), and nugget coefficient to describe the regular pattern of spatial distribution data. Kriging interpolation method includes simple Kriging, ordinary Kriging, universal Kriging and other methods. In practical work, ordinary Kriging is often used to explore the spatial differential features of statistical data.

3. Geostatistical analysis

According to the county-level administrative units, population database of Sichuan Province, which was previously established, a spatial analysis of the database was done that included 2000 and 2010 year population density data of 181 counties and cities of Sichuan province and simulated out the Spatial distribution trends characteristics. Finally, based on the parameters under different models, the best forecasting model was chosen.

3.1 Exploratory data analysis

Purpose of exploratory data analysis is to find out the internal law of data-setting the interpolation parameters and choose theoretical variogram model rationally. The best spatial interpolation can be made using this analysis. The exploratory analysis tools of Arcgis software was used to make exploratory analysis

on spatial distribution, variation characteristics of the population density database of all 181 Counties and cities of Sichuan province.

3.1.1 Trend analysis

With Arcgis software, Trend Analysis Tools provide 3D perspective view that convert sampling sites in study areas into the view that treat the interested property value as height, and allow users to analyze the whole trend of the data sets of sampling point from a different perspective (Tang, 2006). This tool on a certain extent can reflect the spatial phenomenon change on spatial region. Each vertical bar in trend analysis graph represents a data point's value (Height) and position. These points are projected onto an east-west and a north-south direction orthogonal surface, thus Arcgis software can make a best-fit line through the projection point, and use it to impersonate the existent trend on a specific direction. If the lines are straight, this shows no trend existence on the space (Tang, 2006). Therefore, trend surface analysis module of the GIS software was used to analyze the trend of the year 2000 and 2010 Sichuan county population density, respectively. Results are listed in figure 2.

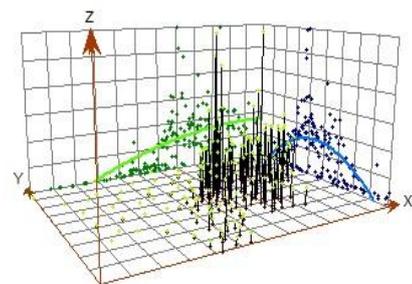


Figure 2 (a): 2000 year trend chart

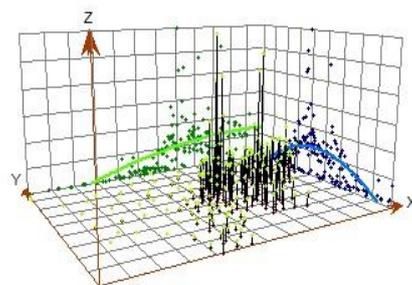


Figure 2 (b): 2010 year trend chart

As can be seen from Figure2 (a) and (b), both 2000 and 2010 year population densities of Sichuan Province were displayed as a gradually increasing straight line from west to east on east-west direction (X-axis direction of extension), indicating that the pattern of population distribution on East-west direction did not change, and eastern population density is still far greater than the western population density in this decade. On the north-south direction (Y-axis direction of extension), the population density of Sichuan on the south to north direction is expressed as an inverted "U"-shaped curve. The pattern of population distribution that centralized from south, north to the middle direction did not change, and central Sichuan regions still have high population density areas. It can also be deduced from the figure that although the 2000 and 2010 on the east-west direction are approximately an inclined line, the population density data on north-south direction are complex function curves. Therefore, when the Arcgis software is used to carry out kriging Interpolation, trend with linear or quadratic functions cannot be removed separately.

3.1.2 Semivariogram cloud analysis

The Semivariogram cloud represent the semivariogram and covariance values of all dataset sample points, and express them as distance function between two points (Tang, 2006). Its abscissa denote the distance between geometric center of two administrative units, while ordinate denote the logarithmic population density's variogram value. Generally, closer point in space has more similarities and, the furthest has greater difference. Therefore, a semivariogram cloud analysis was made on 2000 and 2010 year county population density, Sichuan. Results are listed in figure3.

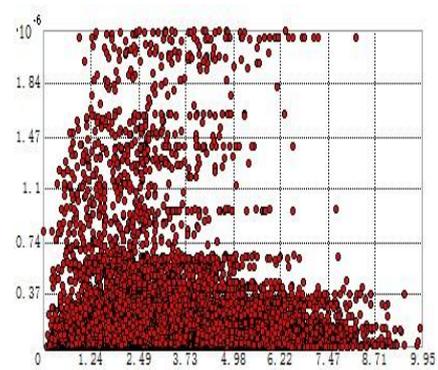


Figure 3 (a): 2000 year semivariogram

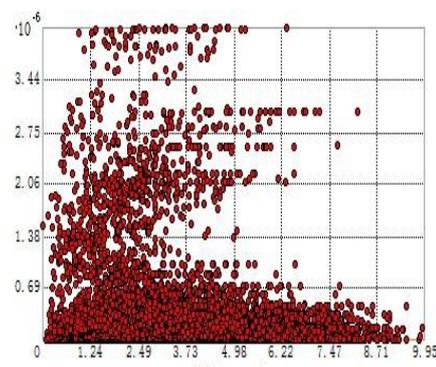


Figure 3 (b): 2010 year semivariogram

From figure3 (a) and (b), it can be seen that in both two years, most data are distributed in bottom left or left partial to middle region and few points are distributed in top region where high population density intersect zones with low population density. Therefore, it can be concluded that the spatial distribution of two years points conform to the law that closer point in space has more similarities, furthest has greater difference.

3.2 Model select

From the above analysis, Circular, Spherical, Gaussian, Exponential and Stable models were initially identified to carry out Kriging spatial interpolation. When this is done, the quality of Model parameters directly determines the quality degree of Fitting effect. The model parameters mainly include Nugget (CO), Sill Value (C+CO), Partial Sill(C), (CO/(CO+C)), Average error (AE) and Root Mean Square Standardized (RMSS). If the nugget coefficient is less than 25%, it reflects that the system has a strong spatial correlation; and more than 75%, reflects that the system has a weak spatial correlation. If RMSS value is near

one and AE is near zero, the model fitting effect is better. Therefore, the above five models and 6 parameters were used to make choice analysis on best fitting effects model of 2000 and 2010 year county population density of Sichuan Province, respectively. Selection analysis process is listed in table 1.

Table1: Population density variogram model selection

Theoretical model	CO		CO+C		CO/CO+C		AE		RMSS	
	2000	2010	2000	2010	2000	2010	2000	2010	2000	2010
Circular	0.00	0.00	7.04	6.95	0	0.00	3.08×10^{-4}	3.09×10^{-4}	3.35×10^{-3}	3.53×10^{-3}
Spherical	0.00	0.00	6.13	6.05	0	0.00	7.85×10^{-4}	7.81×10^{-4}	8.17×10^{-3}	8.56×10^{-3}
Gaussian	0.39	0.42	7.66	7.38	5.05×10^{-2}	0.06	1.54×10^{-5}	1.22×10^{-4}	1.91×10^{-3}	1.74×10^{-3}
Exponential	0.00	0.00	4.70	4.62	0	0.00	4.90×10^{-3}	5.01×10^{-3}	2.79×10^{-2}	2.99×10^{-2}
Stable	0.31	0.34	8.42	8.09	3.64×10^{-2}	0.04	2.82×10^{-5}	6.48×10^{-5}	7.08×10^{-4}	7.85×10^{-5}

From table1, it can be concluded that both nugget coefficient (CO/CO+C) of 2000 and 2010 year population density data that were logarithmically transformed were 0 under circular, spherical and exponential models. Relatively speaking, nugget coefficient under Gaussian model is the largest, so it could not be used as the best model. For average error test indicator, the average error of circular model has the minimum nugget coefficient of '0'. For RMSS test indicator, the circular model was the largest. Therefore, the model is the best Interpolation model for both 2000 and 2010 year spatial distribution of population density of each county, Sichuan province.

3.3 Interpolation map create

The 3.1, 3.2 analysis results showed that the circular model enable 2000 and 2010 year spatial distribution of population density of each county of Sichuan Province to achieve the best effect of interpolation. In order to make the data more visually expressive, the population density of Sichuan Province was divided into 6 levels on the basis of grading principle and standard of population density and combined with special spatial distribution pattern of population history of this province, that is, less than 40 people/km² is extremely sparse population areas; 40~100 people/km² is sparsely populated areas; 100~

250people/km² is relatively sparse population regions; 250~500people/km² is relatively dense population areas; 500 ~ 1000people/km² is densely populated areas; more than 1000 people/km² is extremely dense population areas. The interpolation results are shown in figure4.

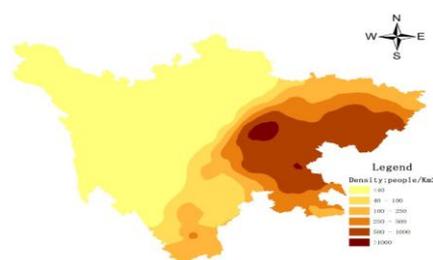


Figure4.(a): 2000 year interpolation map

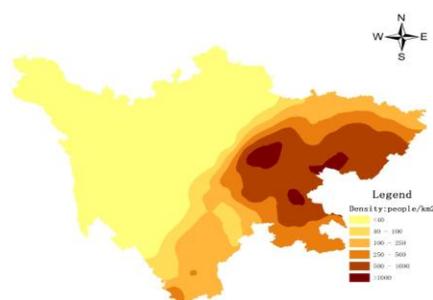


Figure 4(b): 2010 year interpolation map

Summarized from figure 4 (a) and (b), Sichuan population density has the following spatial distribution patterns:

- The overall pattern of county population density spatial distribution has significant difference. Uneven distribution of overall space and circle structure on East-West, South-north direction exist at the same time.

①Axial variation: From the population density trend analysis in Figure 2.1, it can be noted that both the population densities on

East-West and South-North direction had significant differences. The population density tended to increase from west to east and Chengdu has the highest population density value. In the south-north direction, population density concentrate from south and north to central direction, showing an inverted "U" shape curve feature. Chengdu-Deyang area has the most population concentration regions. The moderate gathering area distributed outside Neijiang, Zigong city and other level of density zones with outward expansion having descending order.

②Circle style pattern: Sichuan Province population formed a typical circle structure; Chengdu area with high population density value showed outward diffusion characteristics towards East and West direction. Western Sichuan (except Panzhihua region) is an obviously low population dense region, and Sichuan basin (Chengdu to the east area) with high population density. On the whole, it formed Sichuan basin gathering area and Western Sichuan sparse area District which were two macro patterns of population distribution of geographical units.

- Having compared the charts of 2000 and 2010 year spatial distribution of population density, It is not difficult to find that there is larger time variation characteristics of counties population density in Sichuan province.

①Area of region with high population density (more than 1000people/km²) increased significantly during the last decade. These regions were found in Chengdu surroundings and small region in Southern Sichuan in the year 2000. However, the high value region gradually expanded outward with high speed in the year 2010. Distribution area of Chengdu high-density region increased significantly, and the Suining-Nanchong regions population density grew faster during this decade. Southeast Sichuan (including Neijiang, Zigong) also

showed that distribution of population density start to expand clearly. This phenomenon is mainly caused by rapid social development.

②The overall distribution is of 500 ~ 1000people/km² population density with relatively little change, but has an increasing trend. Compared with the year 2000, population of Dazhou, Leshan, Luzhou, Panzhihua Cities have grown rapidly, maybe because of the implementation of "Chengdu-Chongqing Economic Zone" policy enacted by the Chinese government.

③The areas with 250 ~ 500 people/km² population density expand slowly during this time, mainly growth in Bazhong, Mianyang, Ya'an, Yibin etc., and city.

④The sparsely populated region with 100~ 250people/km² showed relatively small change, main focus being on Liangshan autonomous region, internally growing towards to Chengdu.

4. Spatiotemporal variation of population analysis

4.1Natural factors

Natural conditions are direct influencing factors for population spatial distribution, and are also the most fundamental factors for the formation and development of regional structure of Sichuan Province. Geological and topographical characteristics have obvious influence on population spatial distribution in Sichuan Province. Macro-regional landscape pattern laid the foundation for the development of Sichuan Province. Western Sichuan is dominated by plateau terrain, while the Central, Southern and Eastern Sichuan are mainly dominated by basins and hilly topography. Western Sichuan and Eastern Sichuan have large altitude differences with horizontal development such as low-lying basins and hill terrains. Regional climate also influences the population spatial distribution. Eastern and Western Sichuan have large climatic differences. In recent years, a variety of meteorological and geological disasters occurred in Western Sichuan

regions, especially landslides and mudslides. Geological disasters frequently take place in Western Sichuan regions in the rainy season since the 5.12 Wenchuan earthquakes and 4.20 Ya'an earthquakes. These played important roles in modifying the spatial and temporal distribution of population of Western Sichuan to some extent.

4.2 Policy factors

The evolution of the political situation and the government's policy on the population distribution is a significant impact factor, and also a "growth agent" for the development process of the population. In order to bring a change in the pattern of western regions, Chinese government implemented the "western development" macro strategy as from the year 2000. Sichuan is one of the main beneficiaries of this policy. Since then, a large number of foreign-funded enterprises have invested and established a large number of factories, economic development zones, etc. in eastern and southern Sichuan. This attracted a large number of workers from outside the province, which made the population of these regions grow in mechanical form, and formed a high density of population groups. It changed the regularity of the temporal and spatial distribution of the population in Sichuan province to a certain extent.

4.3 Economic factors

Development of the population and economic development are often synchronized. Economic prosperity will promote the growth of population. "Chengdu circle" is the highest population density areas in Sichuan Province, Chengdu being the capital of Sichuan Province and also the largest city in Southwest china. It is the center for financial, commercial, administrative, cultural, educational, and medical, sports and insurance transactions. Its economic development or urbanization attracted people from outside to come to this province, which resulted in regional population growth in

mechanical form over time. Western Sichuan regions are sparsely populated with poor natural resources, weak economic foundation and characteristics. Despite the increased regional population density in this decade, growth rate is slow.

5. Discussion

This article is only based on Arcgis geostatistical analysis module discussing the best model selection problem of GIS spatial analysis applications, and the use of 2000 and 2010 year counties population density in Sichuan Province as an example to verify this analysis. Finally, the corresponding spatial and temporal distribution with the economic development of this province during this decade was achieved. However, only the relative optimal fitting model was used in this paper to solve the problem. In order to make the spatial expression of ecology, economy, climate science and other fields with GIS geostatistical method more intuitive, a variety of statistical methods for in-depth discussion and continuous improvement and exploration should be integrated.

Reference

- [1] Sustainable Development Research Group of Chinese Academy of Sciences,(2000).Chinese Sustainable Development Strategy Report, Science Press, Beijing.
- [2] Feng J,(2002).Research on Spatial distribution of population density and its Evolutionary model in Hangzhoucity,GeographicalResearch21(5): 635-646.
- [3] Christophers,CohenJE,(1999).Continental physiography, climate and the global distribution of human population.Science Press Beijing:24-136.
- [4] Tang G A,(2006) .ARCGIS GIS spatial analysis experimental Course, Science Press, Beijing.