International Proceedings of Chemical, Biological and Environmental Engineering, Vol. 91 (2016)
DOI: 10.7763/IPCBEE. 2016. V91. 10

Geological Statistics Analysis of Population Distribution at Township Level in Henan Province, China

Haixia Zhang, Wei Qu [+], Shuwen Niu, Jinghui Qi, Liqiong Ye, Guimei Zhang

The College of Earth and Environment Sciences, Lanzhou University, Lanzhou 730000, China

Abstract. Based on the sixth population census data at township level, this article analyzes the population distribution of Henan province, China by the geological statistics method. The result shows that population distribution of Henan province could be divided into three types: low density in mountain areas, medium density in plain areas, and high density in urban regions. The variation functions have similar trends in the four directions of E-W, N-S, NE-SW, and NW-SE. When the distance is over 80km, the anisotropy enhances. The exponential model has the best fitting effect for the variation function. The interpolation results represent the gradient change process of population density intuitively. Terrain condition is the basic factor influencing on the population spatial pattern. High population density in urban regions are the outcomes of mutual effects between the superior geographical condition and socioeconomic development.

Keywords: population distribution, township level, geological statistics, variation function, Henan Province.

1. Introduction

Population distribution is a reflection of the human-earth relationship in a special space-time background. Understanding the population distribution and what determines this distribution is fundamental to understanding the relationships between humans and the environment [1]. With the advancement of modern space technology and geographic information processing technology, the study on Chinese population distribution has experienced from qualitative analysis and simple quantitative to spatial-temporal modeling [2]-[4]. Many research analyzed the population distribution at the scope of provincial level in China [5], [6]. Cities, as populated densely areas, their population distribution has been getting more attention [7], [8]. Other some researches are dedicated to explaining the factors influencing population distribution [9], [10].

Spatial scale is a basic parameter to measure population density. In the large spatial scale, we can only get the macro pattern of population distribution. The characteristics of population distribution will be represented more details in the small spatial scale. However, it is meaningless if the spatial scale is too small, for the spatial variation in population density will disappear. Most of the current researches take county-level region as the basic unit to study population distribution on national or provincial scope [2], [6], [11]. There are also some works to discuss this issue in 1km×1km grid [4], [12], [13]. Due to the dense population in urban areas, most of the related research take the street, block or building as analytic unit [14]-[16]. Based on the daily living space of the major urban-rural people, it is appropriate to take township and block as the basic unit to estimate population density.

Population is dense in Henan province where the human-earth relationship is in the status of tension. Based on the method of geological statistics and the sixth population census data, the article analyzes the features and trends of population distribution at township level in Henan province and reveal what factors determines this distribution.

[+] Corresponding author. Tel.: +86-18919847995; fax: +86-8914027
E-mail address: quwei@lzu.edu.cn; zhanghx13@lzu.edu.cn; shuwenn@lzu.edu.cn

2. Overview of Research Region

Henan province is located in the middle and lower reaches of the Yellow River region. Most of the regions belong to the North China Plain. The mountains lie around its western, southern, and northern edge regions. Nanyang Basin sites in southwestern mountains. Here belongs to continental monsoon climate, with good match of light, heat, water and land resources, which provides favorable conditions for people's survival and development. Hence, here becomes the important birthplace of the Chinese agricultural civilization. Historically, Henan province has been the population gravity center of China and played an important role in ethnic integration, people movement and cultural exchange [17]. Now this province includes 17 prefecture-level cities and 159 counties (including counties, the cities and districts of county level). Until the sixth census, there was 94.03 million population, it is a densely populated province in China. Its total area was approximately 16.7×10^4 km^2 and population density was 568 people/km^2 in 2010, which is approximately 4 times higher than national average level. In 2010, the demographic urbanization level of Henan province was 38.8%, and per capita GDP in current rate was 24446 yuan [18]. Its economic development level ranks the medium in the whole country.

3. Research Method

3.1. Data source and processing

Population data used in this article are derived from the sixth national census. The vector shape file in GIS format for census tracts is available from 2013 edition of the *"Atlas of Henan Province"* [19]. Detail methods are described as follows:

1)Townships in rural areas and municipal districts have clearly defined boundaries, we take these townships as separate units. And so do some streets in municipal districts.

2)Part of the city streets interlace with each other, and it is difficult to dissect their boundaries. We merge these streets of which boundaries are intertwined into one analytic unit.

3)Some industrial parks, farms, forest farms, *et al.* have separate population data, but they lack clear boundaries. In this case, we merge them into the nearest units.

4)Lastly, we get 1955 analytic units.

3.2. Geostatistics method in brief

With regionalized variable as its theory base, Geostatistics is a kind of mathematical geological method using variation functions and spatial interpolation as its essential tools, researching those spatial phenomena with structured and stochastic characters [20].

Under the hypothesis of meeting the second order stationarity, $Z(x)$ is a regionalized variable, $Z(x_i)$ and $Z(x_i+h)$ are respectively the attribute value in spatial position of x_i and $(x_i + h)$, and the variation function $r(h)$ is defined as below.

$$r(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(x_i)-Z(x_i+h)]^2 \quad [i=1,2,\ldots,N(h)] \tag{1}$$

In the equation (1), h is the lag distance, which stands for the separation between samples in both distance and direction. $N(h)$ is the number of paired comparisons at lag h. Sill, Nugget, Range and Fractal Dimension are the four basic parameters of a variation function. The Sill represents the largest variation of system properties. The Nugget represents the variability and measuring errors of variables when h is less then the minimum sampling scale. Spatial autocorrelation of regionalized variable is reflected by the Range. When h exceeds the Range, the spatial autocorrelation will disappear. The Fractal Dimension stands for the curvature of the variation function, higher the Fractal Dimension, stronger the degree of spatial autocorrelation.

As the theoretical model of variogram is unknown, it need to be estimated based on effective spatial samples. Spherical, Exponential, Gaussian and Hole-effective models and so on, are the common variation function models. Based on these models, spatial interpolation can achieve the continuous spatial distribution of population.

4. Results and Analysis

4.1. Population density in different spatial scales

It is a usual phenomenon that population is unevenly distributed in space. Larger the spatial scale, smaller the extreme ratio (the ratio of the maximum and the minimum) of attribute values and stronger the average trend. Taking the 17 prefecture-level cities as basic spatial units, the most dense city is Zhengzhou, and the most sparse one is Sanmenxia, their extreme ratio is 5.4. Taking the counties as basic spatial units, the highest population density county is Weidu District of Xuchang city, and the lowest one is Lushi county, their extreme ratio is 83.5. Taking the townships as basic spatial units, the most dense unit is 83.5 times of the most sparse one. This indicates that spatial scale shrinking can effectively decrease the homogenizing trend in population density and more accurately show the actual distribution status. As shown in Table 1.

Table 1: The extreme ratios on three different space scales

spatial scale	Population density （people/km^2）		
	Max	Min	$\frac{Max}{Min}$
prefecture-level city	1145	213	5.4
County level	7434	88	84.5
township level	17590	21	837.6

Based on the cumulative percentage of population and land area of the 1955 township level units, we draw a Lorenz curve (Fig. 1). It displays the imbalanced phenomenon of population distribution intuitively. In the bottom left corner of the Fig. 1, about 10% of population reside in 35.16% of total area, where population is sparse. These units mainly distribute in the western, southern, and northern mountain regions of Henan province. In the upper right corner of Fig. 1, about 36.20% of population occupy 10% of area, where population is dense. These units scatter across all over province, they mainly are the urban land areas and some organizational system towns with superior geographical position. The middle part of the Lorenz curve represents the plain and basin units where agricultural production is in dominant.

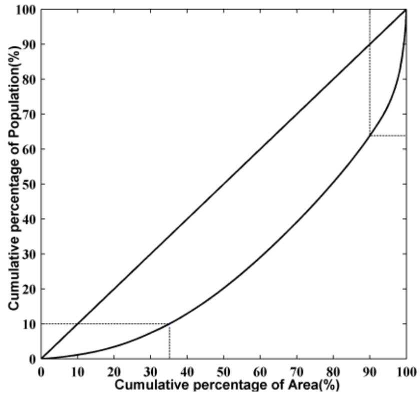

Fig. 1. Lorenz curve of population distribution of Henan Province

4.2. Distribution pattern

Frequency histogram is formed by density interval with 100 people/km^2, it is left-skewed apparently. If interval data is transferred to logarithm, the frequency histogram is approximate to normal distribution, its kurtosis value is 2.28 and skewness value is 0.09. See Fig. 2.

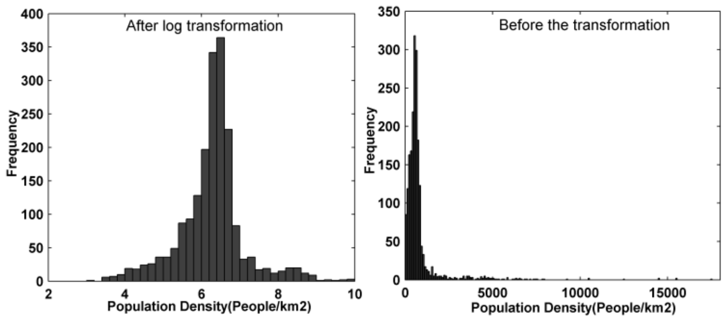

Fig. 2. Histogram for population density of Henan Province

4.3. Variation analysis

Since the average area of the 1955 analysis units is 84km^2 and the side length is about 9km if it is a square, the variation functions in the two directions of E-W and N-S are calculated based on the initial step of 9 km and the two directions of NE-SW and NW-SE are calculated based on the initial step of 12 km, as shown in Fig. 3. The Nuggets in four directions are between 0.27 and 0.38, this indicates that the difference in geographical condition leads to the different spatial distribution of population.

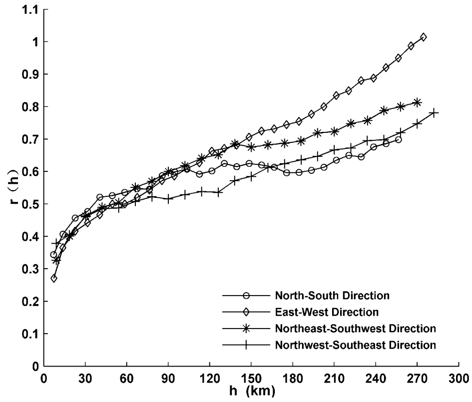

Fig. 3. Variance functions of the four different directions

The Ranges in the four directions range from 262-286km and are very close to half the distance of the provincial boundaries. The maximum variation is in the E-W direction, its *r(h)* increase from 0.271 to 1.041 and its Partial Sill is 0.742. The Partial Sill is 0.354, the smallest in the N-S direction. This index are 0.486 and 0.402 in the NE-SW and NW-SE direction, respectively. Along with the increase of Lag distance, the variation functions of the four directions escalate, which displays the spatial autocorrelation of population density. In other words, within the Range, the greater the distance is, the more significant the population density differences. The variation functions trend similarly within 80km, indicating that the variations are similar in different directions. When *h* is over 80km, the anisotropy enhances. That is, the anisotropic characteristics is expressed fully mainly when h is over 80km.

As shown in Table 2, Fractal Dimensions of the four different directions decrease in the order of N-S, NW-SE, NE-SW and E-W. It means that the population density in N-S direction has the strongest homogeneity and the smallest variation, and the spatial variation of population density in E-W direction is more obviously then other three directions.

Table 2: Fractal dimensions of variance function in the four different directions

Direction	Fractal dimension	SE	R^2
N-S	1.917	0.090	0.944
NE-SW	1.872	0.034	0.993
E-W	1.831	0.063	0.967
NW-SE	1.899	0.120	0.919

4.4. Variation model and interpolation

Exponential model is the best fitted result among the variation function models. Parameters of exponential model are given in Table 3. Based on this model, the interpolation results are shown in Fig. 4. It represents intuitively the gradient change of population density. Sparsely populated areas mainly distribute on the edge of northwestern and southwestern mountains. The plain and basin are the densely populated areas, in which the population density in rural areas is generally lower than that in urban areas. All large, medium, and small cities and most of the county towns form high-value spots. The population spatial pattern can be basically divided into three types, sparsely populated area in mountains, less densely populated area in rural plain, densely populated area in urban plain.

Table 3: Parameters of the exponential variation function model

Lag	R^2	Nugget	Sill	Range（km）
8km	0.974	0.3700	1.0630	292.70
8.5km	0.974	0.3700	1.0720	298.00
9km	0.972	0.3660	1.0460	279.50
9.5km	0.971	0.3610	0.9880	243.00
10km	0.970	0.3620	1.0260	264.50

4.5. Factors affecting population distribution in Henan province

Although there are no extreme natural conditions limiting human existence in Henan province, the variant combinations between natural and human factors influence the population distribution on different levels, which results in uneven spatial distribution.

4.6. Terrain condition is the basic factor impacting population spatial pattern

The difference of hydrothermal condition in Henan province, which is controlled by continental monsoon climate, has no big effect on the population distribution. Terrain condition is the basic factor influencing population distribution. Mountain areas are located in the northern, western and southern sides with forest, in which earth surface is quite steep and crisscrossed with ridge and valley. As there only is a small amount of farm land distributing in valleys and piedmont benchlands, population and villages are very sparse. The plain has superior natural conditions and is traditional agricultural production area. Hence, here has the dense population and villages. Under the influence of topographic factors, there is a dividing line of population density from Luoyang, Pingdingshan, Zhumadian to Huaibin in NW-SE direction, which shows the spatial pattern of population distribution, see Fig. 4. The northeastern part of the line is densely populated whereas the southwestern part is sparsely populated, except Nanyang Basin. Obviously, population distribution in Henan province is affected by the level of land reclamation that is controlled by terrain conditions. In general, the topography conditions of mountains and plains are the essential factors affecting population density.

4.7. Social-economic development and locational conditions are the factors impacting urban population aggregation

The high level of land productivity, large population and central location conditions are considered as the basic factors for controlling and managing the country well. Here plain is one of the cradles of Chinese agriculture civilization. There had been many cities, such as Anyang, Luoyang, Xuchang and Kaifeng cities *et al.* taken as the capital of the Central Plains Dynasty in history. The demand for fresh water made some places near the rivers and lakes become the preferred choices for towns or cities, such as Kaifeng, Luohe cities and so on. Some cities were built up by favorable position of mountains and rivers for enhancing military defense capabilities, such as Luoyang, Xinyang cities and so on. Some cities were formed in the centers of regions and population distribution in order to facilitate commercial trade, for example, Xuchang city.

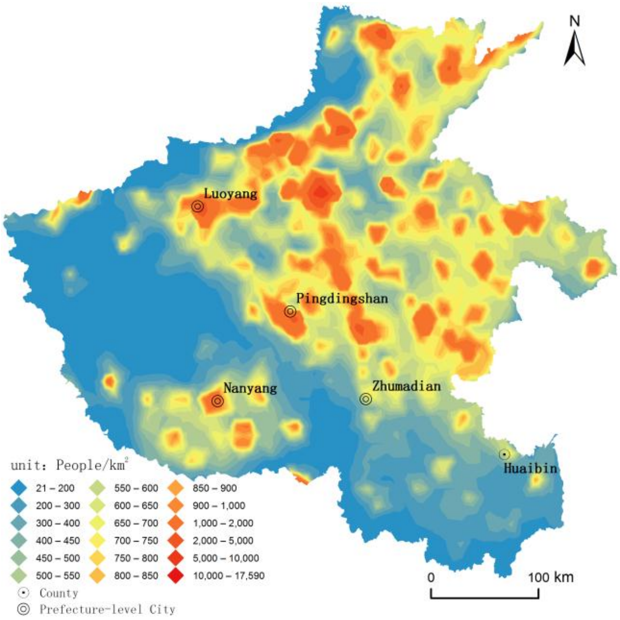

Fig. 4. Interpolation of population density of Henan Province

Since modern times, the demand for energy and mineral resources has increased sharply in the process of industrialization. The exploitation of natural resources, such as coal in Pingdingshan, Jiaozuo and Hebi cities, oil and gas resources in Puyang city, and non-metallic minerals in Zhumadian city, have become an important condition for population gathering and urban development. Moreover, some superior agricultural products also become the raw materials for industrial production, such as tobacco in Xuchang, cotton in Anyang, *et al.* In general, the diversification of economic activities, the specialization of social division of labor and the variety of lifestyle have led to commercial trade and various service industries growing prosperity, and employment has become the key factor to attract population gathering.

Over a long period of time, there is a good interaction relationship between land and water communication and urban development. Especially, Zhengzhou city with the intersection of the two main railways of Beijing-Guangzhou and Lanzhou-Lianyungang become a central hub in whole China. Its geographical advantages highlight increasingly. With the acceleration of industrialization and urbanization, a number of public infrastructures, such as aviation, railways, highways, oil and gas pipelines, water delivery projects and power grids, *et al.* have become the essential conditions for modern people's life in urban areas. The high reliance and close interaction between the cities and the infrastructure produced the Matthew Effect, which has become the driving force of urbanization and population agglomeration.

To sum up, we think that the issue on population distribution may be understood from three levels: 1) Upon a global scale, there are any feasible conditions for human's survival demands, including air, suitable temperature, water and food *et al.*; 2) In the region with survival conditions, the land capacity, which depends on the combination of topography, soil and hydrothermal conditions, constrains population density; 3)The regional advantage and socioeconomic development level become the essential factors in densely populated urban areas. Therefore, population distribution is the result of the multiple factors, including natural and human respects.

5. Conclusion

Spatial scale is a basic parameter to study population distribution, and the actual population distribution will be expressed more accurately in a small spatial level. The population distribution can be divided basically into three categories in Henan province, sparsely populated area in mountain areas, medium density area in rural areas of plains and basins, and densely populated area in urban areas. There are 10% population in 35.1% land areas with low density, and 36.2% population in 10% land with high density.

The *r(h)* is the highest in the E-W direction. The Ranges in four directions range from 262-286km, and the change of *r(h)* in four different directions are similar within 80km. Its anisotropy enhances when the distance is over 80km.

The exponential model has the highest fitting precision, and the spatial interpolation results intuitively show the gradient variation of population density. In the interpolation figure, there is a dividing line of population density from Luoyang, Pingdingshan, Zhumadian to Huaibin in NW-SE direction. The northeastern part of the line is densely populated region whereas the southwestern part is sparsely populated, which indicates the macro pattern of population distribution in Henan province.

The factors impacting population distribution could be divided into three levels. Firstly, there are any suitable environment for human living upon a global scale. Secondly, population density depends on the land capacity which mainly is effected by terrain conditions with the suitable survival environment. Thirdly, the regional advantage and socioeconomic development level are the main factors for urban areas with dense population.

6. References

[1] C. Small, J. E. Cohen. Continental physiography, climate, and the global distribution of human population. *Current anthropology*. 2004, 45(2):269-277.

[2] Z. M. Feng, Y. Z. Yang, Z. You Zhen, J. H. Zhang. Research on the suitability of population distribution at the county level in China. *Acta Geographica Sinica*. 2014, 69(06):723-737.

[3] Q. Pan, X. B. Jin, Y. K. Zhou. Population change and spatiotemporal distribution of China in recent 300 years. *Geographical Research*. 2013, 32(7):1291-1302.

[4] T. X. Yue, Y. A. Wang, J. Y. Liu, S. P. Chen, D. S. Qiu, X. Z. Deng, et al.Surface modelling of human distribution in China. *Ecological Modelling*. 2004, 181 (4):461－478.

[5] Y. Y. Ma, Y. Q. Lu, L. Zhang. Evolvement of Spatial Pattern of Population with Data at County Level in Jiangsu Province. *Progress in Geography*. 2012, 31(02): 167-175.

[6] S. W. Niu, X. D. Guo, Z. G. Liu, M. M. Zeng. Time and Spatial Analysis on the Distribution of Population in Gansu Province over 100 Years. *Chinese Journal of Population Science*. 2006, 6:49-57.

[7] W. Qi, Y. Li, S. H. Liu, X. L. Gao, M. F. Zhao. Estimation of urban population at daytime and nighttime and analyses of their spatial pattern: A case study of Haidian District, Beijing. *Acta Geographica Sinica.* 2013, 68(10): 1344-1356.

[8] X. Mao, R. R. Xu, X. S. Li, Y. Wang, C. Li, B. Zeng. Fine Grid Dynamic Features of Population Distribution in Shenzhen. *Acta Geographica Sinica*. 2010, 65(04): 443-453.

[9] Z. M. Feng, Y. Tang, Y. Z. Yang, D. Zhang. The Relief Degree of Land Surface in China and Its Correlation with Population Distribution. *Acta Geographica Sinica*. 2007, 62(10):1073-1082.

[10] Z. B. Zhang, J. Pan, X. H. Li. The Spatial Evolution and Formation Mechanism of Population Density in Lanzhou City over the Past 30 Years. *Scientia Geographica Sinica*. 2013, 33(01):36-44.

[11] L. Wang, Y. Z. Yang, Z. M. Feng, Z. You. Prediction of China's population in 2020 and 2030 on county scale. *Geographical Research*. 2014, 33(2):310-322.

[12] J. G. Francisco, C. Isidro. A population density grid for Spain. *International Journal of Geographical Information Science*. 2013, 27(12):2247-2263.

[13] F. J. Gallego. A population density grid of the European Union. *Population and Environment*. 2010, 31(6):460-473.

[14] X. H. Qin, Y. H. Wei, W. Chen, X. J. Duan. Population expansion and polycentric development of Nanjing city in a period of hyper-growth. *Geographical Research*. 2013, 32(4):711-719.

[15] F. N. Lin, J. W. Zhao, P. Zhang. GIS based urban population spatial distribution model and application in Changchun. *Science of Surveying and Mapping*. 2008, 33(4):163-165.

[16] L. Zhuo, X. R. Huang, H. Y. Tao, F. Wang, Y. H. Xie. High spatial resolution population distribution simulation based on building information and multi-agent. *Geographical Research*. 2014, 33(3):520-531.

[17] S. Y. Zhang. *Population Geography of China.* Science Press, 2003.

[18] Y. Q. Liu, S. D. Liu, et al. *Henan statistical yearbook-2011*, Available: http://www.ha.stats.gov.cn/hntj/lib/tjnj/2011/indexch.htm.

[19] Map Institute of Henan Province. *Henan province atlas.* China Atlas Press, 2013.

[20] M. A. Oliver, R. Webster. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *CATENA*. 2014, 113: 56-69.