Chinese Geographical Science 2006 16(3) 217–222
DOI 10.1007/s11769-006-0217-9
www.springerlink.com

Character and Causes of Population Distribution in Shenyang City, China

DU Guoming[1, 2], ZHANG Shuwen[1], ZHANG Youquan[1, 2]
(*1. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130012, China; 2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Character of population distribution is one of the focuses studied by urban geography. Using the fifth national census data as basic data and using areal interpolation method, this paper analyzes character of urban population distribution of Shenyang City, Northeast China, in terms of three aspects of statistical character, spatial auto-correlation and spatial structure. Furthermore, this research analyzes the factors affecting the population distribution of the city. The main conclusions include: 1) There is an obvious structure character of population distribution in the grid with a grain of 300m, which is appropriate scale when researching population distribution in Shenyang City. 2) Urban population distribution has the character of assembling while population density distribution takes on variability in Shenyang City. 3) Population density distribution shows spatial auto-correlation within 7.36km. Spatial heterogeneity of population density is low. 4) Urban center, population distribution barycenter and population density maximum points separate each other. Population density distribution has multi-cores character. 5) Layout of governments, primary schools, middle schools, colleges, hospitals and marketplaces affects population distribution directly. With the increase of distance to these factors, population density decreases as logarithm.
Keywords: population density model; population distribution; scale; Shenyang City

1 Introduction

Research on the character of urban population distribution is one of the focuses of western urban geography. From the 1950s, the research has attracted the attention of many western researchers, and a series of classical theoretical models were put forward. For example, Clark (1951) found that with the increase of distance from the city center, the urban population density tends to decrease in exponential. Through the statistical analysis of more than 20 cities, he put forward the famous Clark Model. Naroll and Bertalanffy (1956) and Stewart and Warntz (1958) developed the allometric growth model between urban population and urban area. Sherratt (1960) and Tanner (1961) raised the normal density model (Sherratt Model). Smeed (1961) put forward the negative power-exponential model (Smeed Model). Newling (1969) developed conic exponential model (Newling Model). After the 1980s, with the developing of western research on the structure of multi-cores cities, some scholars have put forward the city population density multi-cores model, and then it became a hot research field (Berry and Kim, 1993; Small and Song, 1994). With the adoption of the later modern mathematic tool, such as fractal and the bionic theory, the quantitative expression ability and the simulation experiment methods of urban geography were strengthened. So some classical mathematics models, such as Clark Model, were given the new explanation, which makes these models to keep developing (Batty and Longley, 1994). Recently, a lot of the theoretical discussion and case studies have been done in China (Chen and Xu, 1999; Chen, 2000; Feng, 2002; 2004).

This paper uses the fifth national census data as basic data, the data of first floor area of the residential houses as auxiliary data to spatialize population data, and uses regular grid to resample population distribution and methodology of landscape ecology, and geo-statistics to research character of urban population distribution, analyze the appropriate scale of population density, and simulate the spatial distribution of population density of Shenyang City, Northeast China. In the last, this research analyzes factors affecting population distribution, which include governments, primary schools, middle schools, colleges, hospitals and marketplaces.

2 Study Area and Data Sources

Shenyang City, capital of Liaoning Province, lies in the south of Northeast China. The total area of Shenyang is 12,980km^2. Until the end of the year 2000, the built-up area was 217km^2. The terrain is mainly plain, with an average elevation of 50m. The Hunhe River flows through the south of the city. Shenyang City is the center

---

Received date: 2006-01-20; accepted date: 2006-04-21
Foundation item: Under the auspices of Knowledge Innovation Program of Chines Academy of Sciences (No. KZCX2-SW-210-1)
Biography: DU Guoming (1978–), male, a native of Ningcheng of Inner Mongolia, Ph.D. candidate, specialized in GIS and RS application. E-mail: nmgdgm@126.com

of polity, economy and culture of Liaoning Province. It is also one of the most important heavy industry cities in China. Until the fifth census, the urban population of Shenyang was 5.3×10^6. The research range is a rectangles region that includes urban area and suburb area. The total study area is $336km^2$, involving 981 census sections.

The source data of this research include the fifth national census and GIS spatial data. The census data was supplied by population statistic department. The spatial data, including the census sections data, the data of first floor of the residential houses and the basic urban geographical information data, are mapped and supplied by the related departments and are up to precision of 1∶5000 mapping. The census data and spatial data, which are integrated together, make up of GIS database of this research.

3 Methods

3.1 Spatializing census data

In the research, the process of spatializing census data and calculating population density includes three steps as follows: 1) To establish relation between the census data and census sections by the census section code. 2) To spatialize population census data using areal interpolation (Pan and Liu, 2002). This step includes overlaying the vector data of the first floor of the residential houses with the census sections data, calculating the area of the first floor in every census section and the population density of first floor in every census section. The calculation formula is as follows:

$$D_i=P_i/A_i \tag{1}$$

where D_i is population density of the first floor of the residential houses in the ith census section; P_i, the quantity of population in the ith census section; A_i, the total area of the first floor of the houses in the ith census section. 3) To use the regular grid to resample, and to calculate the quantity and density of population for every grid. In Arc/Info environment, we overlay the grid data with the first floor data and calculate the quantity and density of population for every grid. The formula is as follows:

$$P_j=\sum_i A_{ij} * D_{ij} \tag{2}$$

$$D_j=P_j/A_j \tag{3}$$

where P_j represents the quantity of population in the jth grid; A_{ij}, the area of the first floor of the residential houses that belongs to the ith census section and is located in the jth grid; D_{ij}, the density of the first floor of the residential houses that belongs to the ith census section and is located in the jth grid; D_j, the density of population of the jth grid; A_j, the area of the jth grid.

According to landscape ecology, we call the side of grid as grain (G). In order to choosing appropriate scale, grains in the research include ten kinds from 100m to 1000m.

3.2 Spatial statistical analysis methods

3.2.1 Statistics

Statistics can be used for analyzing population density. The statistics indexes used in this research include mean, maximum (max), minimum (min), and standard deviation (SD).

3.2.2 Landscape index method

Landscape index is a basic tool of landscape ecology for researching spatial pattern. The population density is different from the land use, and can not be classified, but it can be grouped according to the density range, and the groups can take the place of the classes. This research puts the density that is zero into one group, and other densities are grouped in an interval of 100 person/ha. Then we calculate landscape richness index (R), Shannon-weaver landscape diversity index (H), landscape evenness index (E) and landscape dominance index (D) (Wu, 2000).

3.2.3 Geo-statistics

On the basis of regionalized variable theory, geo-statistics researches spatial phenomena with structured and stochastic features. Geo-statistics has two main fields: 1) analysis of spatial correlation and variation of regionalized variable; 2) spatial interpolation (Wang, 1999). The semi-variance function and its model are the foundation of geo-statistics research. According to structure character of the semi-variance of population density, this paper chooses ordinary Kriging to model spatial distribution of population density and draw population density isoline.

3.3 Factors analysis method

In Shenyang built-up area, spatial heterogeneity of main physical geographical elements is small, such as climate, physiognomy and hydrology. Urban infrastructure, such as communication, water supply, heating, power supply, can satisfy living of the residents. These elements can not limit population distribution. But most public sectors' distribution is asymmetrical, such as governments, schools and hospitals, which have close relations with living of the resident. So this research chooses governments, primary schools, middle schools (including high schools), colleges, hospitals, marketplaces and park, greenbelts as the factors affecting population distribution. Data disposing includes following steps:

Firstly, to pick the elements affecting population distribution from the database, simplifing governments, primary schools, middle schools, colleges, hospitals and marketplaces as points, and incorporating parks and greenbelts into one polygon cover.

Secondly, to create multi-rings buffer data for every element with 250m-distance, and erasing them out of the research region.

Thirdly, to overlay every buffer data with the vector data of the first floor of the residential houses, then calculating area, population quantity and density for every ring in every buffer using formulae (4) and (5).

$$P_r=\sum_i A_{ir} * D_{ir} \tag{4}$$

$$D_r=P_r/A_r \tag{5}$$

where P_r represents population amount in the rth ring; A_{ir}, the area of the first floor of the residential houses that belongs to the *i*th census section and is located in the *r*th ring; D_{ir}, the population density of the first floor of the residential houses that belongs to the *i*th census section and is located in the *r*th ring; D_r, the population density of the *r*th ring; A_r , the area of the *r*th ring.

4 Results and Analysis

4.1 Choosing appropriate scale

The problem of choosing the appropriate scale is involved when using grid analysis technique to analyze character of population distribution. This research confirms the appropriate scale by landscape indexes. It can be shown in Table 1 that when *G* is 100m, *R* is 20; when *G* increases to 300m, *R* decreases to 10; with the increase of *G*, *R* decrease a littile. There is an obvious inflexion in the curve of *R* (Fig. 1). The Shannon-weaver landscape diversity index (H), which is based on the information entropy theory, decreases with the increase of the grain. The landscape evenness index (E) and dominance index (D) fluctuate with the increase of G, and the maximum E and minimum D appear respectively when the grain is 300m (Fig. 1). This suggests that there is an obvious structure feature when G is 300m for population density, and 300m is the appropriate scale to research population distribution in Shenyang.

Table 1 Landscape indexes of population density

G (m)	*R*	*H*	*E*	*D*
100	20	1.5956	0.5326	1.4001
200	14	1.5941	0.6040	1.0450
300	10	1.5182	0.6593	0.7844
400	10	1.4150	0.6145	0.8876
500	9	1.3660	0.6217	0.8312
600	9	1.2723	0.5790	0.9250
700	8	1.2080	0.5809	0.8714
800	8	1.1762	0.5656	0.9033
900	8	1.1806	0.5678	0.8988
1000	7	1.1367	0.5841	0.8092

4.2 Character of population distribution

4.2.1 Statistical character

Population density statistics when *G* is 300m are listed in Table 2. It can be seen from Table 2 that population density of Shenyang is dispersed. SD/Mean=1.58 indicates that population density distribution has variability. In general, if SD^2/Mean is more than 1, population distribution exhibits convergence. Now SD^2/Mean arrives to 266.56, which illuminates that the convergence of population is very high. For analyzing the degree of population convergence, we draw a Lorenz curve of population distribution (Fig. 2). It can be seen that the curve is segregative with the rectangle diagonal, which suggests that population distribution has character of non-equilibrium and assembling. The point (29.57%, 88.30%) is an obvious inflexion of the curve, which illuminates that about 90% of population resides in 30% area of the research region, where population is dense; and about 10% of population resides in the other 70% area of the research region, where population is sparse.

4.2.2 Character of spatial auto-correlation

Semi-variance function and its parameters are listed in Table 3 when *G* is 300m. It can be discovered that: 1) semi-variance function accords with spherical model; 2) there is spatial auto-correlation within a special range for population density, which decreases with the increase of

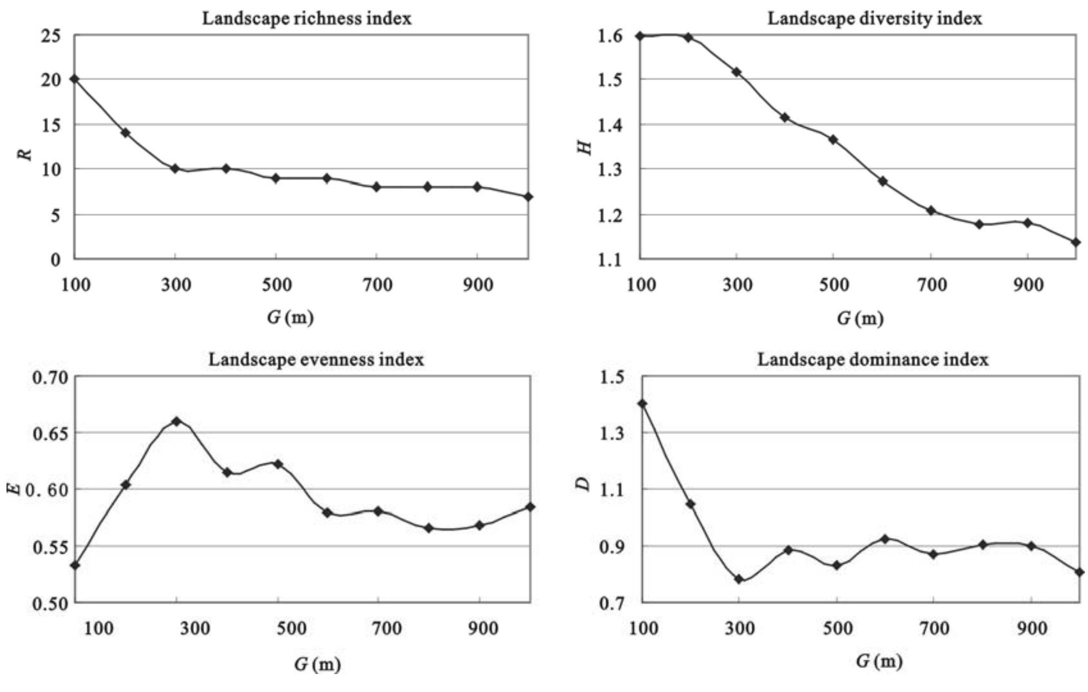

Fig. 1 Relative landscape indexes

Table 2 Population density statistics (G=300m)

Sample number	Mean (person/ha)	Max (person/ha)	Min (person/ha)	SD (person/ha)	SD/ Mean	SD^2/ Mean
3740	107.21	882.01	0	169.05	1.58	266.56

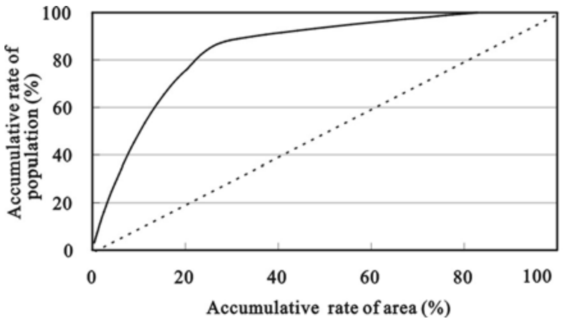

Fig. 2 Lorenz curve of population distribution in Shenyang

Table 3 Semivariogram model and parameter (G=300m)

h (m)	Model	Nugget	Partial still	Still	Nugget/ still	a (m)
900	Spherical	11649	25568	37217	0.3130	7355.7

Notes: h represents lag distance; a represents auto-correlation scale

distance. But there is no spatial auto-correlation out of the special range for population density. The special distance is auto-correlation scale (a). The auto-correlation scale is 7.36km in the research region. The nugget is 11,649, which suggests that part of spatial heterogeneity of population density is resulted from random factors. This part takes effect mainly within 900m. The still is 37,217 and the partial still is 25,568, which means that part of heterogeneity is resulted from structured characters and spatial auto-correlation. This part takes effect mainly between 0.9km and 7.36km. The score of nugget and still is 0.313, which means that spatial heterogeneity that resulted from random factors accounts for 31.3%, spatial heterogeneity that resulted from structure character and spatial auto-correlation accounts for 68.7%. The latter is major of the whole spatial heterogeneity of population density.

4.2.3 Character of spatial structure

On the basis of the character of roads distribution in Shenyang City, this research regards radiation center of city roads as urban center. Moreover, this research calculates population barycenter according to grid data whose grain is 300m, and draws population density map (Fig. 3). The population barycenter is expressed as follows:

$$X = \sum_{i=1}^{n}(C_i \times X_i) / \sum_{i=1}^{n} C_i$$
$$Y = \sum_{i=1}^{n}(C_i \times Y_i) / \sum_{i=1}^{n} C_i \qquad (6)$$

where X and Y represent the population barycenter coordinate, C_i represents the population quantity of the ith cell. X_i and Y_i represent the center coordinate of the ith cell.

For showing character of population distribution clearly, based on Fig. 3 and semi-variance analysis, we use ordinary Kriging to interpolate, and create population density simulation map (Fig. 4).

As shown in Fig. 3 and Fig. 4, population barycenter and urban center are not superposed, whose distance is 1.61km. Urban center, population barycenter and population density maximum points are separate. Population distribution shows a multi-cores structure. The close isolines are elliptic in the population density high-valued regions. There is a "U" sunken region of population density in the western of the city, and there is a low-valued region of population in the northeast of the city. According to the above-mentioned, population distribution takes on complicated multi-cores structure.

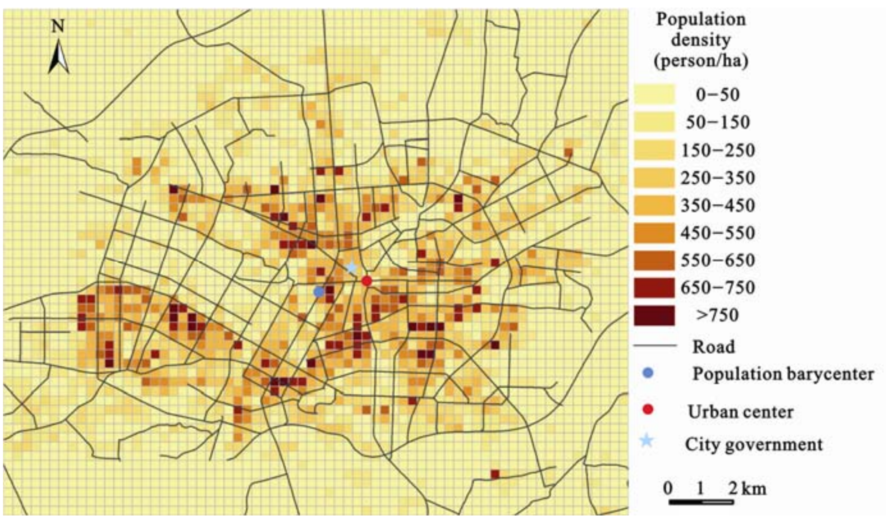

Fig. 3 Population density map of Shenyang

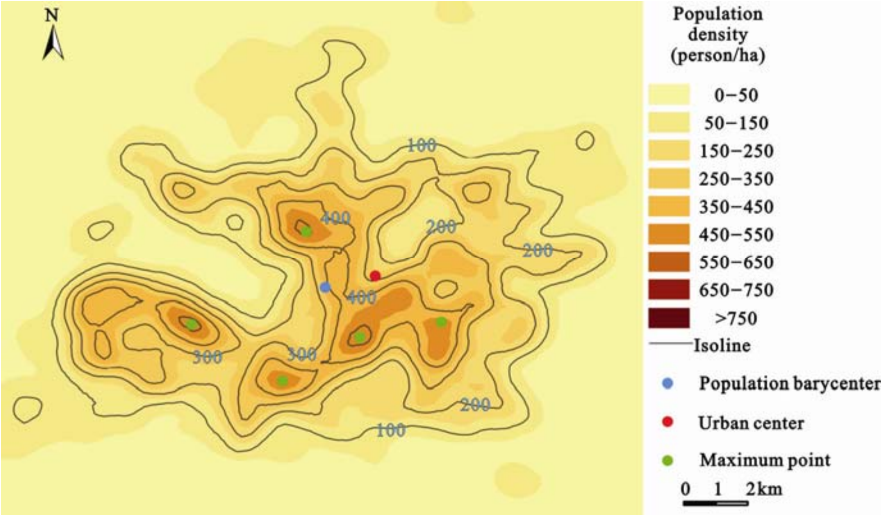

Fig. 4 Population density simulation map of Shenyang

4.3 Factors affecting population distribution

Based on the results calculated by sector 3.3, the curves of relationship between distance to the elements and population density are drawn (Fig. 5). As shown in Fig. 5, population density fluctuates with the increase of distance to parks and greenbelt, but the period is not certain. For the other elements, population density decreases with the increase of distance to these elements, and there are obvious inflexions around 1km. For describing these elements' impacts on population distribution accurately, in the light of Table 3, this research choose linearity, logarithm, exponent and power models to simulate population density and distance for every element, and choose the best model whose R^2 is the greatest (Table 4). It can be shown in Table 4 that all the best models are logarithm model and all R^2 are more than 0.86, which illuminates that the effects of models simulating are fine. These models prove that the distribution of governments, primary schools, middle schools, colleges, hospitals and marketplaces affects population distribution directly. With the increase of distance to these elements, population density decrease sharply. When the distance arrives at a certain extent, the rate of population density decreasing becomes slower, and the population density trends to zero. Moreover, when overlaying population density map with distribution map of these elements, it can be discovered that these factors assemble in the regions with high population density.

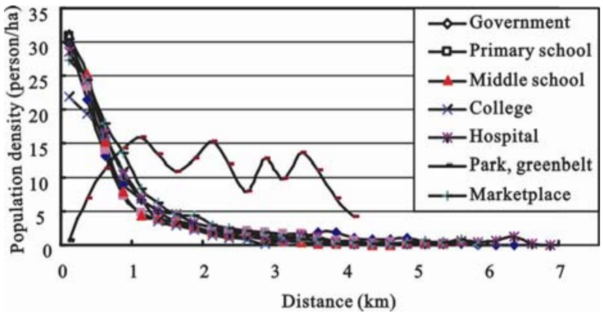

Fig. 5 Relation curves of distance to every element and population density

Table 4 Conditions of every factor and population density-distance model

Factor	Number	Ring number	Model	R^2
Government	357	26	$y = -69.989\text{Ln}(x) + 103.2$	0.8845
Primary school	160	22	$y = -78.466\text{Ln}(x) + 103.77$	0.8813
Middle school	164	23	$y = -80.347\text{Ln}(x) + 106.19$	0.8672
College	196	15	$y = -76.047\text{Ln}(x) + 86.002$	0.9223
Hospital	149	28	$y = -71.115\text{Ln}(x) + 109.6$	0.8786
Marketplace	131	24	$y = -78.16\text{Ln}(x) + 113.89$	0.9225

Note: *y* represents population density (person/ha); *x* represents the nearest distance to the relevant factor in any point of the research region (km)

5 Conclusions

Using the fifth national census data as basic data and using areal interpolation method, this paper analyzes character of urban population distribution of Shenyang City, Northeast China. Furthermore, this paper analyzes the factors affecting the population distribution of the city. The main conclusions of this paper include:

Firstly, there is an obvious structure character of population distribution in the grid with a grain of 300m which is appropriate scale when researching population distribution in Shenyang City.

Secondly, urban population distribution has character of assembling while population density distribution takes on variability in Shenyang City. And 90% of population resides in about 30% of area of the research area.

Thirdly, spatial auto-correlation scale of population density is 7.36km. Spatial heterogeneity of population density results from spatial auto-correlation mainly.

Fourthly, urban center, population distribution barycenter and population density maximum point separate each other, and population density takes on multi-cores character.

Fifthly, the distribution of primary schools, middle schools, colleges, hospitals and marketplaces affects the population distribution directly. With the increase of distance to these elements, population density decreases as logarithm. These elements assemble in the regions with high population density.

References

Batty M, Longley P A, 1994. *Fractal Cities: A Geometry of Form and Function*. London: Academic Press, Harcourt Brace & Company.

Berry B J, Kim H M, 1993. Challenges to the monocentric model. *Geographical Analysis*, 25: 1–4.

Chen Yanguang, Xu Qiuhong, 1999. Studies on the fractal geometric model of allometric growth relationships between area and population of urban systems. *Journal of Xinyang Teachers College* (*Natural Science Edition*), 12(2): 198–203. (in Chinese)

Chen Yanguang, 2000. Derivation and generalization of Clark's model on urban population density using entropy—Maximizing methods and fractal ideas. *Journal of Central China Normal University* (*Nat. Sci.*), 34(4): 489–492. (in Chinese)

Clark C, 1951. Urban population densities. *Journal of Royal Statistical Society*, l14: 490–496.

Feng Jian, 2002. Modeling the spatial distribution of urban population density and its evolution in Hangzhou. *Geographical Research*, 21(5): 635–646. (in Chinese)

Feng Jian, 2004. *Restructuring of Urban Internal Space in China in the Transition Period*. Beijing: Science Press. (in Chinese)

Naroll R S, Bertalanffy L Von, 1956. The principle of allometry in biology and the social sciences. *General Systems Yearbook*, 1: 76–89.

Newling B E, 1969. The spatial variation of urban population densities. *Geographical Review*, 59: 242–252.

Pan Zhiqiang, Liu Gaohuan, 2002. The research progress of areal interpolation. *Progress in Geography*, 21(2): 146–152. (in Chinese)

Sherrantt G G, 1960. A model for general urban growth. In: Churchman C W et al. (eds.). *Management Sciences, Model and Techniques: Proceedings of the Sixth International Meeting of Institute of Management Sciences* (Vol. 2). Elmsford, N Y: Pergamon Press, 147–159.

Small K A, Song S, 1994. Population and employment densities: Structure and change. *Journal of Urban Economics*, 36: 292–313.

Smeed R J, 1961. The traffic problem in towns. In: *Manchester Statistical Society Papers*. Manchester: Norbury Lockwood.

Stewart J Q, Warntz W, 1958. Physics of population distribution. *Journal of Regional Scienc*, 1: 99–123.

Tanner J C, 1961. Factors affecting the amount travel. In: *Road Research Technical Paper No.51*. London: Department of Scientific and Industrial Research.

Wang Zhengquan, 1999. *Geostatistics and It's Application for Ecology*. Beijing: Science Press. (in Chinese)

Wu Jianguo, 2000. *Landscape Ecology-pattern, Process, Scale and Hierarchy*. Beijing: Higher Education Press. (in Chinese)