ATLANTIS PRESS
Advances in Social Science, Education and Humanities Research (ASSEHR), volume 130
2nd International Forum on Management, Education and Information Technology Application (IFMEITA 2017)

Analysis of the Main Influencing Factors for Population Density Based on the Method of Gray Correlation Analysis

Yu Xia, Cheng Wang*
School of Mathematics and Economics, Hubei University of Education,
Wuhan 430205, P. R. China
*Corresponding Author: wangc80@163.com

Keywords: Population density; Gray correlation analysis; Main influencing factors.
Abstract. Aiming at analyzing the main factors which may influence the population density, this paper establishes a gray correlation analysis method to determine the main influencing factors for population density of Wuhan City according to the data in the statistical yearbook of Wuhan City. The result shows that GDP, employment quantity and completed area of housing construction are the three main factors which affect the population density of Wuhan City.

Introduction

In 2015, with the rapid development of Optics Valley's economy, as GDP accounted for 70% of Wuhan's population, has been ranked 1 million 500 thousand at Optics Valley GDP in Wuhan accounted for half of the country, the proportion will further expand in the future, it is expected that in 2019, the "big Optics Valley" total output value will exceed 1 trillion yuan of industrial sector. Huge economic strength and sufficient talent reserves, Optics Valley has become increasingly dazzling.

In this paper, we take the region area and population density as the research object, and focus on the the factors such as Wuhan GDP, employment quantity and completed area of housing construction [1-4], then we establish a gray correlation analysis method [5-8] to determine the main influencing factors for population density of Wuhan City according to the data in the statistical yearbook of Wuhan City.

Correlation Analysis on Various Factors Affecting Population Density

Gray relational principle. Let $X=\{x_0,x_1,\mathbf{L},x_m\}$ be a gray correlation factor set, x_0 be a reference sequence, and x_i be the compare sequences, $i=\{1,2,\mathbf{L},m\}$, $x_0(k),x_i(k)$ respectively, and x_0 and x_i be the number of k points, that is

$x_0=(x_0(1),x_2(2),\mathbf{L},x_0(n))$, $x_1=(x_1(1),x_1(2),\mathbf{L},x_1(n))$, $x_2=(x_2(1),x_2(2),\mathbf{L},x_2(n))$

……

$x_m=(x_m(1),x_m(2),\mathbf{L},x_m(n))$,

if

$$r(x_0,x_i)=\sum_{k=1}^{n}w_{\ k}r(x_0(k),x_i(k)) \qquad r(x_0(k),x_i(k))=\frac{\Delta_{\min}+r\quad\Delta_{\max}}{\Delta_{0i}(k)+r\quad\Delta_{\max}}$$

where $\Delta_{0i}(k)=\left|x_0(k)-x_i(k)\right|$ is the absolute difference, $\Delta_{\min}=\min_i\min_k\Delta_{0i}(k)$ is the smallest difference between the poles, $\Delta_{\max}=\max_i\max_k\Delta_{0i}(k)$ is the maximum difference for the poles, r is the distinguish between coefficients (r in the above theorem in the actual calculation of the general is 0.5). w_k is the k-point weight, which satisfies $0\le w_k\le 1,\sum_{k=1}^{n}w_k=1$, then we call $r\left(x_0(k),x_i(k)\right)$ is the gray correlation coefficient for x_0 and x_i, and $r(x_0,x_i)$ is the gray relation degree of x_0 and x_i.

In general, when $r(x_0, x_i) > r(x_0, x_j)$, which means the gray correlation degree between x_i and x_0 is higher than that between x_j and x_0, or the influence degree of x_i on x_0 is greater than the influence degree of x_j on x_0.

Gray Relational Analysis on Factors Affecting Population Density. Let $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ represent the annual average temperature in Wuhan City, employment quantity, and completed area of housing construction, the green coverage rate of built-up areas, the mass production of public transportation vehicles, high-tech industries per 10,000 population, road area and annual GDP.

(1) Data collection and analysis

The statistics of various factors that affect the population density can be obtained by referring to Wuhan Bureau of Statistics [3] as shown in Table 1 below.

Table 1. 2010-2015 Factors statistics for population density in Wuhan City

Employment quantity (million)	483	498	506.4	522.24	530.44	544.92
Completed area of housing construction (10,000 square meters)	5664.39	8114.17	9223.5	10103.98	10223.95	11958.87
Built-up area green coverage (%)	37.17	37.59	38.19	38.2	39.21	39.65
Every 10,000 people have public transport vehicles (standard)	15.5	15.2	14.2	13.9	16.2	13.8
The total output of high-tech industry development (million)	26380353	34489139	45560041	56044728	65996966	74992058
Road area (million square meters)	7273	7726.2	9027.2	8384	8879.63	9495
Annual output value	5565.93	6762.2	8003.82	9051.27	10069.48	10905.6

(2) Analyze the factors affecting population density by gray relational analysis method

We take the population density as a reference sequence $x_0(k), k = 1, \mathbf{L}, 6$, other factors as the comparative factor sequences $x_i(k), i = 1, \mathbf{L}, 8; k = 1, \mathbf{L}, 6$, the specific calculation process of gray correlation analysis is as follows:

Step 1: Data standardization. Initialization of the various factors, the standardization of the sequence $y_i(k), i = 0, 1, \mathbf{L}, 8; k = 1, \mathbf{L}, 6$, and dimensionless sequence, as shown in Table 2 below.

Table 2. Various factors data series

Factors	2010	2011	2012	2013	2014
Population density (person / km2)	1.0000	1.0010	1.1992	1.2104	1.2226
Average annual temperature (degrees Celsius)	1.0000	0.9819	0.9879	1.0301	1.0060
Employment quantity (million)	1.0000	1.0310	1.0484	1.0812	1.0982
Completed area of housing construction (10,000 square meters)	1.0000	1.4324	1.6283	1.7837	1.8049
Built-up area green coverage (%)	1.0000	1.0112	1.0274	1.0277	1.0548
Every 10,000 people have public transport vehicles (standard)	1.0000	0.9806	0.9161	0.8967	1.0451
The total output of high-tech industry development (million)	1.0000	1.3073	1.7270	2.1244	2.5017
Road area (million square meters)	1.0000	1.0623	1.2411	1.1527	1.2209
Annual output value	1.0000	1.2248	1.4509	1.6409	1.8255

Step 2: Calculate the absolute difference

Calculate the absolute difference based on the data in above Table 2 by the formula $\Delta_i(k) = |y_0(k) - y_i(k)|$, the result is as follows.

$\Delta_1 = (0, 0.0191, 0.2113, 0.1803, 0.2166, 0.2136)$

$\Delta_2 = (0, 0.0300, 0.1508, 0.1292, 0.1244, 0.0975)$

$\Delta_3 = (0, 0.4314, 0.4291, 0.5733, 0.5823, 0.8856)$

$\Delta_4 = (0, 0.0102, 0.1718, 0.1827, 0.1678, 0.1589)$

$\Delta_5 = (0, 0.0204, 0.2831, 0.3137, 0.1775, 0.3353)$

$\Delta_6 = (0, 0.3063, 0.5278, 0.9140, 1.2791, 1.6171)$

$\Delta_7 = (0, 0.0613, 0.0419, 0.0577, 0.0017, 0.0799)$

$\Delta_8 = (0, 0.2238, 0.2517, 0.4305, 0.6029, 0.7606)$

Obviously, we have

$\Delta_{\min} = 0$, $\Delta_{\max} = 1.6171$

Step 3: Calculate the correlation coefficient.

The correlation coefficient is calculated from the calculation result of Step2 as follows.

We set $r = 0.5$, then we have

$$x_{i(k)} = \frac{0 + 0.5 \times 1.6171}{\Delta_i + 0.5 \times 1.6171}.$$

Substituting the data in Table 2 into this formula, we can directly calculate the correlation coefficient, the result is as follows.

$x_{1(k)} = (1, 0.976903, 0.782821, 0.817694, 0.788743, 0.791020)$

$x_{2(k)} = (1, 0.964242, 0.842821, 0.862254, 0.866704, 0.892380)$

$x_{3(k)} = (1, 0.652090, 0.653286, 0.585105, 0.581317, 0.477361)$

$x_{4(k)} = (1, 0.987562, 0.824767, 0.815714, 0.828168, 0.835745)$

$x_{5(k)} = (1, 0.975371, 0.740675, 0.720293, 0.820020, 0.706860)$

$x_{6(k)} = (1, 0.725263, 0.605036, 0.469380, 0.387292, 0.333333)$

$x_{7(k)} = (1, 0.929545, 0.950717, 0.933427, 0.997949, 0.910077)$

$x_{8(k)} = (1, 0.783224, 0.762592, 0.652537, 0.572833, 0.515280)$

Step 4: Calculate the gray correlation degree and analyze the dominant factors

Take $w_k = \frac{1}{6}$, then compare factor x_i and reference factor x_0, the gray correlation degree r_i is

$$r_i = \frac{1}{6}\sum_{k=1}^{6} x_i(k).$$

Substituting the data in Table 2 into this formula, we can directly calculate the result as follows.

r_1=0.8612, r_2=0.9047, r_3=0.8820, r_4=0.6582, r_5=0.8272, r_6=0.5867, r_7=0.9536, r_8=0.7144.

By comparison, we have

$$r_7 > r_2 > r_3 > r_1 > r_5 > r_8 > r_4 > r_6$$

Obviously, Wuhan's annual GDP has the greatest impact on population density, followed by the employment quantity and the completed area of housing construction. Therefore, these three factors are the main factors affecting population density in Wuhan City.

Conclusions

In this paper, we use quantitative method to present a number of factors that affect the population density, and then use gray relational analysis to quantitatively analyze and extract three main factors. This combination of qualitative and quantitative ideas has some new ideas, it is worth promoting. Optics Valley region has developed rapidly in recent years, so relative to the data of previous years, the data in recent years can better reflect the future development of the region. In order to make the model more reasonable, we can change the weight of the calendar year from when calculating the correlation coefficient, and then calculate the correlation coefficient to give the most relevant factors.

Acknowledgments

This work is supported by the 2017 Excellent Youth Project of Hubei Provincial Department of Education (No. Q20173007).

References

[1] Z.Q. Luo: Derivation and Application of Arbitrary Polygon Area Formula (University of Electronic Science and Technology of China, Chengdu 2005).

[2] Y. Gao: Simulation of Downscaling of County Population Density Based on GIS Data (China's Population, Resources and Environment 2017).

[3] Wuhan Statistical Yearbook on http://www.whtj.gov.cn/

[4] M. Grove: Population density, mobility, and cultural transmission, Journal of Archaeological Science 74 (2016) 75-84.

[5] Y.H. Liu, C.C. Gao and Y.Y. Lu: The impact of urbanization on GHG emissions in China: The role of population density, Journal of Cleaner Production 157 (2017) 299-309.

[6] C.J. Rao, X.P. Xiao, M. Goh, J.J. Zheng and J.H. Wen: Compound mechanism design of supplier selection based on multi-attribute auction and risk management of supply chain, Computers & Industrial Engineering 105 (2017) 63-75.

[7] M.M. Rahman: Do population density, economic growth, energy use and exports adversely affect environmental quality in Asian populous countries? Renewable and Sustainable Energy Reviews 77 (2017) 506-514.

[8] C.J. Rao, M. Goh, Y. Zhao and J.J. Zheng: Location selection of sustainability city logistics centers, Transportation Research Part D: Transport and Environment 36 (2015) 29-44.