Demographic Data Visualization on Continuous Area Cartograms

Lijing Ren [a], Zhengxu Zhao [b]

School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 054300, China

[a]ren.lijing@foxmail.com, [b]zhaozx@stdu.edu.cn

Abstract. Cartogram is a map depicting attributes of spatial data by distorting two dimensional original maps while preserving their topology. In this paper, a novel cartogram thematic map is adopted to express population growth. The best fitting model is proposed to predict the population of randomly selected modern cities in a provincial area of P R China. The result is visualized with cartograms to produce useful findings for scheming a population planning. It is expected that this method will be an essence for the study of population prediction models.

Keywords: Cartogram; Population prediction; Curve fitting.

1. Introduction

The population problem is one of the major problems facing the world in the 21st century. Growing population directly affect countries' economic construction and social progress. Expanding urban population results in contradictions among population, resources and environment. The population is an important factor in economic policy and social development. Reasonable population is one of the main drivers of regional economic development. While too much or too little population may cause enormous ecological problems, which will limit or slow economic growth[1].

An effective measure to control population growth is to understand the changing rule of population and establish the reliable prediction model. Several research streams converge to provide us with a number of tools and models, such as the Malthusian population model, the Logistic growth model, the GM (1, 1) grey prediction model and the BP neural network prediction model, etc.[2]. While research on population prediction from the perspective of visualization is emerging but to date has not received wide-spread attention. Both abstract and concrete digital information can be expressed by visualization techniques. Visualization techniques not only can express data more intuitively, but also facilitate individuals from different fields transmitting and sharing data easily. As thematic maps, cartograms visualize attributes values by loading original map and coinciding each cell with corresponding attributes values. The exaggerated deformation reflects the distribution of attribute values intuitively, which helps users to understand rules of data. With the development of computer graphics and cartography, cartograms find more applications in wide-ranging fields [3].

In this paper, the least-square method is used for establishment of the population prediction model which models time as the independent variable and population as the dependent variable. A novel cartogram thematic map is adopted to express future population growth. It is an important innovation to forecast population by thematic maps, which enriches the population prediction methods.

2. Population Prediction Model

The least-square method is most commonly applied form of linear regression by minimizing sum of square deviations from each point to find the best match function. Given there are n points $(x_i, y_i)(i=1,2,...n)$ in a plane, where (x_i, y_i) is different from each other. The best fitting function $f\left(x\right)$ can be got using the undetermined coefficients method as follows [4]:

$$\mathrm{Q} = \sum_{i=1}^{n} (f(x_i) - y_i)^2 \to min \tag{1}$$

Given a curve fitting functions $f\left(x\right) = a_0 + a_1 x + a_2 x^2 + ... + a_m x^m$, the best number of fitting m minimizing the error between the fitting results and the experimental data can be selected by the least square method. The study found the low order polynomial can express the change of the population more accurately, while the differences between predicted data using the high order polynomial and actual data is large. Taking into consideration the accumulated error and the instability of high order polynomial, the quadratic polynomial is selected for population prediction. The general expression of quadratic polynomial is $f(x)=a_0+a_1x+a_2x^2$,where a_0, a_1 and a_2 are parameters to be determined.

The population of 11 cities in Hebei Province is analyzed based on the proposed model. The Hebei Province is located in the central zone of china and the north of the North China Plain. Since the birth control policy introduction, the birth rate in Hebei Province has descended from 26.73% in the early 1970s to the current 13%. At the end of 2015, the total population of Hebei Province is about 74.55 million. Low birth level creates a favorable environment for the economic development and improves population quality steadily. However, faced with a large population base in China, it is still necessary to control population growth and grasp population growth law.

The demographic data around cities of Hebei Province from 2006 to 2015 is collected from the Hebei Province Statistical Yearbook[5]. A polynomial fitting programming in Maple is developed to predict population changes. Taking Shijiazhuang City which is the capital and largest city of North China's Hebei Province for example, the final fitting quadratic polynomial function is $f(x)=913.6+14.61x-0.40575x^2$.

Table 1 Population prediction results cities in Hebei Province

City	2020 year (x=15)	2025 year (x=20)	2030 year (x=25)
Zhangjiakou	1003.75	1167.28	1319.53
Chengde	297.77	298.22	295.25
Baoding	1110.99	986.39	792.20
Cangzhou	811.66	861.79	914.26
Hengshui	449.52	620.19	700.51
Shijiazhuang	1041.45	1043.50	1024.65
Xingtai	787.87	817.81	842.16
Handan	1003.31	952.02	847.30
Tangshan	729.85	701.44	655.09
Langfang	428.33	407.98	369.56
Qinhuangdao	383.89	383.91	358.07

The population data from 2020 to 2030 of other cities in Hebei Province can also be predicted based on population prediction functions. The results are shown in Table 1.

3. Population Density Cartograms

Cartogram map was proposed by the professor Mark Newman at the Michigan University. It expresses certain attribute values through the vision effect. Each unit in the map is weighted to cause distortion of the original graphics, at the same time, keeps topological relationships of each unit. Based on certain attribute values of deformation maps, exaggerated rendering is carried out on the map, which makes users directly read specific meanings of the map expressing without reference to other information. Cartograms highlight attributes differences of maps and enhance visual recognition [6].

There are two broad categories of cartograms, area and linear. The former is also called value area map. It expresses thematic information through the ratio of attribute values and the local area to reflect the characteristics of units [7]. While the latter is mainly used for linear features by alerting the distance of maps. The area case is focused exclusively in this paper. The area cartograms have nothing to do with the unit area, but is proportional to attributes values of each unit [8].

Assuming administrative areas around the cities in Hebei Province within the prediction years not changing, the population destiny of different years around the cities can be calculated according to the population model. The population destiny from 2020 to 2030, respectively is expressed by the cartogram thematic map plugin in the ArcGIS system. The population density is imported from the Shapefile files and resulting thematic maps are exported as SVG files [9].

A comparison of cartogram base map and the Hebei Province map showing the population destiny is provided in Fig. 1-Fig. 4. The administrative divisions of Hebei Province are shown on Fig. 1 which is treated as the base map. The cartograms of population density around cities in Hebei Province respectively for 2020, 2025, 2030 are shown on Fig. 2- Fig. 4. It is intuitive to observe the change of population around the cities in different years.

As shown in cartograms, the bigger administrative division have bigger local population density. The administrative division areas of Langfang City and Qinhuangdao City are small, but become large after deformation. The change keeps relative stable in years. While Zhangjiakou City and Chengde City have greater administrative area, the surface area decreases after deformation. In terms of growth rate, the area of Zhangjiakou City continues to increase while Tangshan City continues to decrease. Other cities maintain a relatively stable growth rate.

Fig. 1 Administrative divisions of Hebei

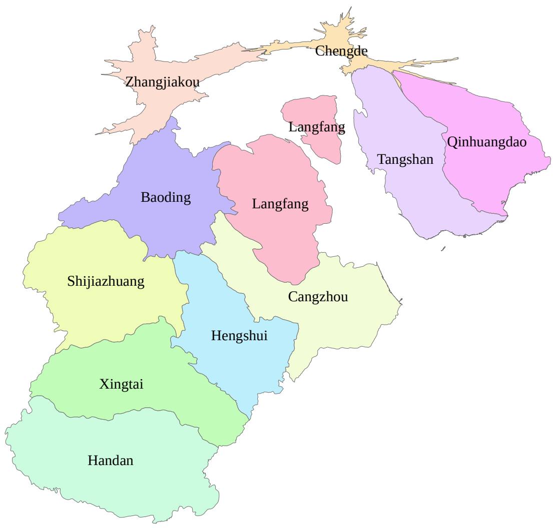

Fig. 2 Population of Hebei in 2020

Fig. 3 Population of Hebei in 2025

Fig. 4 Population of Hebei in 2030

From the predicted population results, it can be seen that population of Hebei Province will maintain relatively stable in the future. One of reasons is that the future fertility groups mostly are the only child. the high social consumption increases the cost of raising children [10]. While geographical advantages make the population of cities such as Langfang City which is located approximately midway between Beijing and Tianjin and Qinhuangdao City which is a port city on the coast of China continue to grow.

4. Conclusion

The human overpopulation is one of the major problems facing the china currently. Establish a reliable prediction model is an effective measure to make a population planning. In this paper, a population growth model based on the curing fitting method is proposed. The result is visualized with cartograms in which the sizes of cites have been rescaled according to their population. Different from conventional maps which are limited to the topographic acreage and cannot reflect the feature differences between geographical elements, the novel thematic maps provide visual clues to the identity of the distorted objects. The technique expressing demographic data by cartogram thematic maps enrich the study of population prediction methods. This research has only focused on just demographic data of a provincial area. In the future, statistical studies of higher-level administrative area throughout China, Asia and even the world will be carried out.

References

[1]. B. Renoust, G. Melançon, and T. Munzner, Detangler: Visual Analytics for Multiplex Networks, Computer Graphics Forum, Vol. 34 (2015), p. 321–330.

[2]. Z. Zhao and J. Fan, an Investigation into Small World Phenomenon in Engineering Informatics, the Third International Multi-Conference on Computing in the Global Information Technology, 2008. ICCGI ’08, (2008), p. 125–130.

[3]. J. H. Kämper, S. G. Kobourov, and M. Nöllenburg, Circular-arc cartograms, 2013 IEEE Pacific Visualization Symposium (PacificVis), (2013), p. 1–8.

[4]. A. J. Elliot and M. A. Maier, Color Psychology: Effects of Perceiving Color on Psychological Functioning in Humans, Annual Review of Psychology, Vol. 65 (2014), p. 95–120.

[5]. The people’s government of Hebei province. Hebei Economy Book 2005-2014 year [M]. Beijing: China Statistics Press., China Statistics Press, (2014).

[6]. A. L. Hugine, S. A. Guerlain, and F. E. Turrentine, Visualizing surgical quality data with treemaps, Journal of Surgical Research, Vol. 191 (2014), p. 74–83.

[7]. W. Cui, S. Liu, L. Tan, C. Shi, Y. Song, Z. Gao, H. Qu, et al., TextFlow: Towards Better Understanding of Evolving Topics in Text, IEEE Transactions on Visualization and Computer Graphics, Vol. 17 (2011), p. 2412–2421.

[8]. T. Schreck and D. Keim, Visual Analysis of Social Media Data, Computer, Vol. 46 (2013), p. 68–75.

[9]. D. Thom, H. Bosch, S. Koch, M. Wörner, and T. Ertl, Spatiotemporal anomaly detection through visual analysis of geolocated Twitter messages, Visualization Symposium (PacificVis), 2012 IEEE Pacific, (2012), p. 41–48.

[10]. S. Barlowe, Y. Liu, J. Yang, D. R. Livesay, D. J. Jacobs, J. Mottonen, and D. Verma, WaveMap: Interactively Discovering Features From Protein Flexibility Matrices Using Wavelet-based Visual Analytics, Computer Graphics Forum, Vol. 30 (2011), p. 1001–1010.