

# VISUAL ANALYSIS BASED ON THE DATA OF CHINESE SURVEYING AND MAPPING JOURNALS

LI Jing<sup>a</sup>, LIU Haiyan<sup>a</sup>, GUO Wenyue<sup>ab</sup>, YU Anzhu<sup>ab</sup>

<sup>a</sup> Institute of Surveying and Mapping, Zhengzhou, China.

<sup>b</sup> State Key Laboratory of Geo-Information Engineering, Xi'an, China.

## Commission VI, WG VI/4

**KEY WORDS:** statistics, Visual analysis, research hot spots, bibliometrics, Price's Law

### ABSTRACT:

Taking four influential Chinese surveying and mapping journals as the data source, 5863 papers published during the period of 2003-2013 were obtained. Using the method of bibliometrics and visual analysis, summarizing the surveying and mapping papers in the past ten years (2003-2013), research themes, authors, and geographical distribution were analyzed. In the study, the papers of geodesy, cartography and GIS are 59.9%, more than half of all the papers. We also determine that the core author group has 131 authors, mainly of whom are from big cities. 90% of top ten cities on the number of publishing papers are capital cities or municipalities directly under the central government. In conclusion, we found that the research focus was different every year, and the research content was richness, the content of geodesy, cartography and GIS were widely researched, and the development of surveying and mapping is imbalanced in China.

## 1. INTRODUCTION

*Acta Geodaetica et Cartographica Sinica Bulletin of Survey and Mapping, Science of Surveying and Mapping, Science of Surveying and Mapping and Engineering of Surveying and Mapping* are four influential journals in the field of Surveying and mapping in China. The papers are the results and thoughts of scientific research, in some degrees, reflecting the development of Surveying and mapping, and the latest achievements in this field in China. In order to explore the development of Surveying and Mapping Science and technology in China from qualitative analysis to quantitative analysis,

In this paper, we use the methods of bibliometrics to analyze three aspects of 5863 papers published in the past eleven years (2003-2013): keywords, first author, and the first author's address. Then according to those factors to find out the changes of research hot spots, the core author group and the center research area of Chinese surveying and mapping science. At the same time, to provide information and reference for subsequent development of the field.

## 2. METHOD

### 2.1 Bibliometrics

Bibliometrics is a quantitative analysis method, with various

external feature as the research object, using mathematical and statistical methods to describe, and it is quantity analysis method using in analyzing the external features of the literature. It focuses on the analysis of the literature form features of "quantity", analyzing of the law of literature from a quantitative point, and indirect reflecting the relationship (Wang Yuefen, 2007).

### 2.2 Price's Law

Price's law is a famous literature review. It was introduced in by D. Price in small science big science in 1963. It can be formulated as follows: in the same topic, half of the papers written by a group of high productive authors, and the number of the prolific authors is approximately equal to the square root of the total number of authors (Price, 1963). This theory establishes the relationship between author and the amount of papers. It provide a review to determine the core author group in one field.

## 3. VISUAL ANALYSIS

### 3.1 The evolution of the hot research spots

Academic papers are different from general text data. They are belong to the semi-structured data, having a certain structure and

elements. Keywords in the literature, to a certain extent, reflects the research content of this article. Our study is based on the key elements of journal articles data characteristic, in years, using ROST word frequency statistics software developed of Wuhan university to statistics, in order to reflect the changes of the research content on time. And according to the paper title, abstract, key words elements such as information, we classify the papers, so as to highlight the differences in research content of surveying and mapping.

### 3.1.1 Keyword statistics

Fig.1 shows the highest frequency keywords from 2003 to 2013 without useful words. We can find that, the highest frequency keywords different from each other every year. This could reflect in the 11 years, the research hot spot is different every year. The average value of word frequency is 9.7, and the highest is 13, "comprehensive map", in 2005. As can be seen from the keywords content, from 2003 to 2007, mainly belong to the area of GIS and cartography.

Year	Keyword	Word frequency	Total number	Percentage (%)
2003	DEM	11	1370	8.0%
2004	Topological relationship	11	1492	7.3%
2005	Map generalization	13	1349	9.6%
2006	Spatio-temporal data model	9	932	9.7%
2007	isohight	8	1074	7.4%
2008	Coordination transformation	10	1925	5.1%
2009	RS	10	2963	3.4%
2010	Precise point positioning	8	2294	3.5%
2011	visualization	10	2705	3.7%
2012	Coordination transformation	9	2307	3.9%
2013	Settlement monitoring	8	1192	6.7%

Fig.1 Highest frequency keywords from 2003 to 2013

Figure 2 reflects the frequency and percentage of keywords by year (from 2003-2013): broken line represents word frequency, and the bar chart indicates the percentage of this word frequency and all keywords' amount. What be seen from the chart, is that word frequency and its percentage are basically positively correlated relationship from 2003 to 2007, that is to say, the higher the keyword frequency is, the greater the percentage is.

By this way, we can illustrate the word has a certain representative word which could reflect the research hot spot. From 2008 to 2013, the total number of keywords is increasing, but the highest frequency and the highest percentage of word frequency keywords are significantly lower than before.

This shows that in recent years the research contents in comparison is more and more dispersed, and diversified.

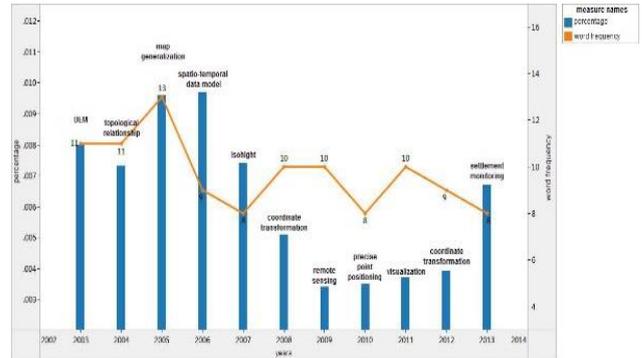


Fig.2 keyword changed by year(from 2003-2013)

### 3.1.2 The classification of the paper content

We classify the papers by title, keywords and abstract. Fig.3 display the results of classification. From Fig. 3, we can see: (1) nearly 11 years, the highest number of papers is geodesy, a total of 1407, accounting for 24% of the total number of papers, the second highest number is GIS, accounting for 19.6% of the total number of papers, and the third is cartography, accounting for 16.3% of the total number of papers. (2) The papers related to geodetic survey, GIS and cartography research sum up to 59.9%. This shows that in the field of surveying and mapping, geodesy, cartography and GIS has been widely researched and developed. (3) On the other hand, other aspects of the research also in step by step, have enriched the field of surveying and mapping.

Classes	Number	Percentage(%)
Geodesy	1407	24.0%
GIS	1150	19.6%
Cartography	953	16.3%
RS	791	13.5%
GPS	588	10.0%
Photogrammetry	603	10.3%
Instrument of surveying and mapping	102	1.7%
Ocean Surveying	65	1.1%
Engineering surveying	54	0.9%
Others	150	2.6%
Total	5863	100.0%

Fig.3 Papers' content classifications

We also reference the geographic information science and technology body of knowledge proposed UGGIS (David DiBiase, 2006) in 2006 to classify papers into eight classes: analytical methods, cartography and visualization, design aspects, data modeling, geographical calculation, geospatial data, data operation, and geographic information science and technology.

From Fig.4, we can find that surveying and mapping field is mainly based on geographical spatial data which is the basis of the surveying and mapping research. The concept of the geospatial data occupies an important position. Overall, data modeling, analysis, geographic calculation, cartography and visualization, in which areas numbers of papers accounted for the majority, are mainly related to the exploration of research methods. Secondly, the design aspects, geographic information science and technology, mainly involving the application of geographic information and services, have few publications. This shows that experts pay more attention to the exploration of law and the establishment of the model and algorithm.

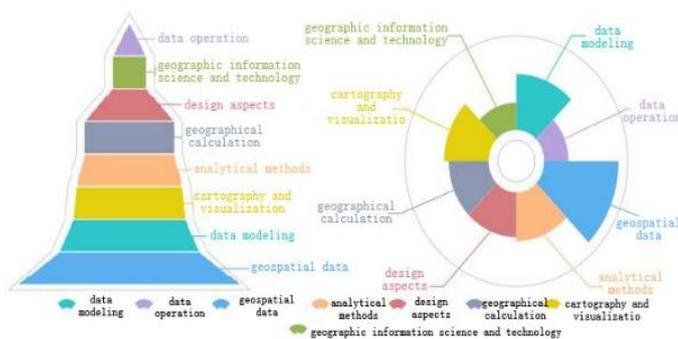


Fig. 4 Papers' content classified by UGGIS

### 3.2 Paper authors' analysis

#### 3.2.1 The number of papers published by the author

Through the statistical data, we find that between 2003 and 2013, 5822 papers are signed by authors. The highest productive author has published 45 papers, and the minimum is 1 paper, the average number is 1.59 papers.

Fig.5 reflects the connection between the number of published paper and authors. We also calculate the percentage that different numbers of published papers in all papers. All among them, the number who published only one paper is the largest, taking 45% place of all the papers. This is suggesting that the

four journals has rich resources of new authors. The paper number greater than 3 amounted to 23.9%, indicating that the four journals have a high quality author team.

The paper number	The author number	Total number	Percentage (%)
1	2620	2620	45.0%
2	579	1158	19.9%
3	218	654	11.2%
4	96	384	6.6%
5	62	310	5.3%
6	21	126	2.2%
7	6	54	0.9%
8	16	120	2.1%
9	6	99	1.7%
10	4	40	0.7%
11	6	66	1.1%
12	2	24	0.4%
13	1	13	0.2%
14	1	14	0.2%
15	1	15	0.3%
17	1	17	0.3%
18	2	36	0.6%
27	1	27	0.5%
45	1	45	0.8%

Fig.5 statistical data of author and the number of paper

#### 3.2.2 Core author group

Core author group means the collection of prolific and influential authors. Our study is in accordance with the Price law to determine the number of core author group in Chinese surveying and mapping field. Computation formula is as follows:

$$N = 0.749(Z)^{1/2} \quad (1)$$

where  $Z$  = the number of papers that the most prolific author has published

$N$  = the least amount of papers that a core author should publish

From the Fig.5, we can easily find that  $Z$  equals to 45. According to the Price law, the core authors must have published at least 5 papers. So we get 131 core authors in our data, who have published papers accounted for 17.3% of total number of papers. That shows that the authors in the field of surveying and mapping is stable and strong. Although the number of core author group is small, the abilities of them is higher.

Fig.6 summarize the author information, publishing at least 10 papers. There are 20 authors in total. From Fig.7, we can find

that the place in which published paper number is above 10, 45% in Beijing, and 15% in Wuhan.

### 3.3 Geographical distribution of core research areas

Quantity of papers, from a certain extent, reflects the scientific

research level in this area. In this paper, based on the number of paper, to determine the distribution of core research areas in China.

Author's name	Author's work unit	Author's location	Number of papers
Chen Junyong	StateBureau of Surveying and Mapping	Beijing	45
Li Deren	Wuhan University	Wuhan	27
Yang Yuanxi	Xi'an Institute of Surveying and mapping	Xi'an	18
Huang Wenqian	PLA Dalian Naval Academy	Dalian	18
Dang Yamin	Chinese Academy of Surveying and mapping	Beijing	17
Hu Mingcheng	Chinese Academy of Surveying and mapping	Beijing	15
Zhang Chuanyin	Chinese Academy of Surveying and mapping	Beijing	14
Ning Jinsheng	School of Geodesy and Geomatics,Wuhan University	Wuhan	13
Zhong Yexun	Guangxi Teachers Education University	Nanning	12
Wang Renxiang	Xi'an Institute of Surveying and mapping	Xi'an	12
Zhang Hanwei	National Astronomical Observatory of Chinese Academy of Sciences	Kunming	11
Hu Shengwu	Henan Polytechnic University	Jiaozuo	11
Gui Qingming	Information Engineering University	Zhengzhou	11
Zhang Qingpu	Chinese Academy of Surveying and mapping	Beijing	11
Yang Minghu	Chinese Academy of Surveying and mapping	Beijing	11
Chen Jun	National Basic Geographic Information Center	Beijing	11
Hua Xianghong	School of Geodesy and Geomatics,Wuhan University	Wuhan	10
Liu Jiping	Chinese Academy of Surveying and mapping	Beijing	10
Li Jun	Geographic Information Engineering Center of Southwest Jiao Tong University	Chengdu	10
Lin Zongjian	Chinese Academy of Surveying and mapping	Beijing	10

Fig.6 Author information

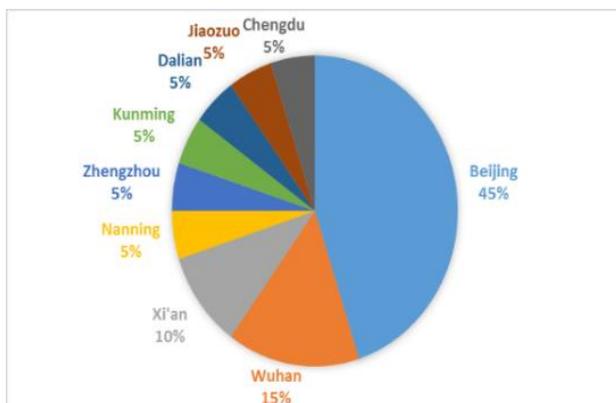


Fig.7 Geographical distribution of authors (number of published paper above 10)

In this paper, we count the first author's place. In Fig.8, we list the top ten cities. From the statics results ,it is easy to be found:(1) Beijing ranks the top, in which number of papers is 18.7% in all papers. The second is Wuhan, and the third is Zhengzhou. (2)The number in top ten cities accounts for 70% of the total number. (3)The study found that among the top 10 cities, Beijing, Shanghai are municipalities directly under the central government. In addition to Xuzhou, the rest are all capital cities. This shows that in the developed economy and large population places, there are abundant education resources and academic communication is convenient. It is the core research area of the

surveying and mapping field.

Ranking	City	Number of papers	Percentage %
1	Beijing	1094	18.70%
2	Wuhan	910	15.50%
3	Zhengzhou	633	10.80%
4	Xi'an	322	5.40%
5	Nanjing	294	5.00%
6	Changsha	238	4.10%
7	Chengdu	190	3.20%
8	Shanghai	185	3.20%
9	Xuzhou	126	2.10%
10	Harbin	119	2.00%
Total	-----	4111	70.00%

Fig.8 Top 10 cities in the number of paper publishing

Figures 9 and 10 are drawn the geographical distribution by using Tableau visualization software. Based on 5 article number for the node, we consider the place paper number above 5 as the high post area, less than 5 and greater than zero as the low post area.

From Fig.9, the high post area mainly distribute in central and eastern regions, and only Urumchi is in the west of China. High post areas are mainly the provincial capital, municipality directly under the central government and other large cities.

From Fig.10, the low post areas are also in central and eastern regions, and southeastern coastal cities. The scale is relatively smaller. Overall, regardless of the high post or the low post areas, they mainly distribute in central and eastern regions. The scale of the city and the economic development directly affect the level of the scientific research.

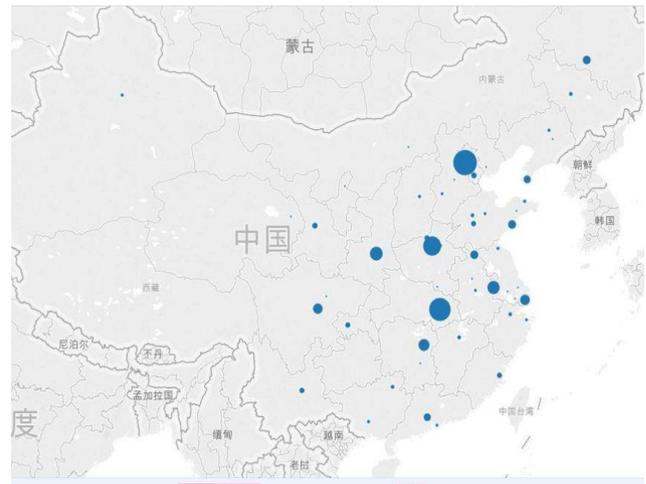


Fig.9 Geographical distribution (the number of paper publishing >5)



Fig.10 Geographical distribution (the number of paper publishing ≤5)

#### 4. CONCLUSION

The above analysis show that journal paper data has very high value in quantitatively describing the research situation and development of a discipline. Through the way of count and visualize the paper data of surveying and mapping, we can draw the following conclusion:

- (1) Hot spots: the research focus is different every year, and the research content is richness, diversity since 2008. From 2003 to 2007, surveying and mapping research focus areas are mainly concentrated in GIS and mapping.
- (2) Research content: geodesy, cartography and GIS have been widely researched; The geographical spatial data is the basis of the research of the surveying and mapping discipline; For researchers, they are more committed to seek for the scientific method and law.

(3) Authors: the discipline has an experienced, practical, high theoretical author team. Although their number is small, they hold up the development of surveying and mapping.

(4) The discipline of surveying and mapping is mainly distributed in the central and eastern regions. It can be seen is that the development of surveying and mapping is imbalanced in China.

#### ACKNOWLEDGEMENTS (OPTIONAL)

This work is funded by State Key Laboratory of Geo-information Engineering (NO. SKLGIE2015-M-3-1) and the National Natural Science Foundation of China (Nos. 41471387,41501446).

#### REFERENCES

Tang Guyu, Liu Zhiyuan, 2003a, Text visualization research review. *Journal of computer-aided design and graphics*, 25 (3), pp. 273-285.

David DiBiase, Michael DeMers, Ann Johnson, Karen Kemp, Ann Taylor Luck, etc. 2006. *Geographic information science and technology body of knowledge*, Association of American Geographers

Liu Haiyan, Sun Qun, Xiao Qiang, etc, 2006a. In the digital mapping of multi-source data (information) comprehensive application. *Journal of surveying and mapping science and*

*technology*, 23(3), pp. 161-164.

Zhou Zhao, Liu Haiyan, Li Shaomei, etc.2008a, Modern map language. *Journal of surveying and mapping engineering*, 17 (2), pp. 6-8.

Chen Sheng, Zeng Xingji, Liang Jun, 2008a. Based on statistical data of GIS visualization study. *Computer engineering and design*, 29 (14), pp. 3757-3759.

Wang Yuefen, 2007. The comprehensive research of literature metrology and content analysis [D]. Nanjing university of science and technology.

Wang Kai, 2013. News text set visualization model research [D]. China university of geosciences (Beijing).

Lin Hongfei, Gao Tian, 2000a The visual representation of Chinese text. *Journal of northeastern university: natural science edition*, 21 (5), pp. 501-504.

Zhong Wenjuan, 2012a. Based on the price law and the core of the comprehensive index method the author evaluation - in the library construction, for example. *Science and technology management research*, 32 (2), pp. 57 to 60.

Wang Yuefen, 2007. "A synthetic research of bibliometric method and content analysis method," Ph.D. Dissertation of Nanjing University of Science and Technology

APPENDIX (OPTIONAL)

Larger visions :

