Research progress and hotspot analysis of spatial interpolation

Li-juan Jia¹, Xin-qi Zheng¹, Jin-li Miao²

¹ China University of Geosciences (Beijing), School of Information Engineering, Beijing 100083, China
² Development Research Center of China Geological Survey, Beijing 100083, China

Abstract: In this paper, the literatures related to spatial interpolation between 1982 and 2017, which are included in the Web of Science core database, are used as data sources, and the visualization analysis is carried out according to the co-country network, co-category network, co-citation network, keywords co-occurrence network. It is found that spatial interpolation has experienced three stages: slow development, steady development and rapid development; The cross effect between 11 clustering groups, the main convergence of spatial interpolation theory research, the practical application and case study of spatial interpolation and research on the accuracy and efficiency of spatial interpolation. Finding the optimal spatial interpolation is the frontier and hot spot of the research. Spatial interpolation research has formed a theoretical basis and research system framework, interdisciplinary strong, is widely used in various fields.

1. Introduction
Spatial interpolation is a method of deriving unknown spatial data from known spatial data. Spatial interpolation based on the hypothesis of the first law of Geography: The closer the point of space is, the more likely it is to have similar features; and the farther away the point is, the less likely it is to have similar features. That is, there is a spatial correlation between geological attributes. Spatial continuous data plays an important role in many scientific researches. Scientists need accurate spatial continuous data to make reasonable speculation and interpretation of the research area. Therefore, it is necessary to know the research progress of spatial interpolation. There are many different classification methods for spatial interpolation according to different standards, Huang Xingyuan¹ et al. divides spatial interpolation into point interpolation and region interpolation based on known points and known partition data; Wu Lun² et al. divides spatial interpolation into two types: spatial interpolation and extrapolation; Based on the basic assumptions and mathematical nature of spatial interpolation, Li Xin³ et al. divides spatial interpolation into spatial statistical methods, function methods and other methods, and introduces the scope of application, algorithms and advantages and disadvantages of each method. Spatial interpolation has been applied to various fields such as Geosciences, environmental science and so on. Luo Ming⁴ et al. analyzed the progress of spatial soft data and its interpolation method; Shi Wengiao⁵ et al. made an analysis of the interpolation methods and accuracy of soil spatial attributes; He Hongyan⁶ et al. made an analysis of the research progress of spatial interpolation techniques for precipitation.

There are many achievements in the field of spatial interpolation in specific disciplines, however, there is little research on spatial interpolation itself. With the help of CiteSpace visual analysis software, this paper makes a bibliometric analysis of the subject of "spatial interpolation", which is included in the core collection database of Web of Science. The research progress of spatial
interpolation, research topics and their correlation and research trends are reviewed, and then the development process of spatial interpolation in recent years is summarized, so as to provide reference for spatial interpolation research.

2. Data and methods

The data used in bibliometric analysis comes from the core collection database of ISI Web of Science, subject headings are set as “Spatial interpolation”, the search time interval is set to 1982-2017 years, the type of document is set as “article”, in the retrieval results, we exclude the literature which has nothing to do with spatial interpolation, 1029 documents were recorded (data download deadline is October 8, 2017), the resulting data is exported in text format, and named CiteSpace recognizable file name starting with download.

This paper uses the bibliometric analysis software CiteSpace to visualize and analyze the data. CiteSpace is an information visualization software developed by Chen Chaomei, a professor at Drexel University, based on the Java platform for measuring and analyzing scientific literature data. CiteSpace software can visually display the information panorama of scientific knowledge field, identify the key literature in a scientific field, research hotspots and evolution progress. It displays the macro structure and development venation of scientific knowledge by using time-sharing dynamic visualization Atlas[7].

3. Results analysis

3.1 basic feature

The historical changes of the quantity of literature can be used to measure the progress of this research field. The annual statistical analysis of the literature data is shown in Figure 1, the data was only in October in 2017, so it was represented by dotted lines. It can be seen from the graph that the spatial interpolation related literature and cited literature are similar in the whole research interval, it has experienced three stages of slow growth steady growth and rapid growth; 1982-1990 years ago, the number of published and cited literatures was small and the growth rate was slow, which indicated that the spatial interpolation research was developing slowly for a long time; During the period of steady growth in the past 1991-2006 years, the degree and importance of spatial interpolation have been increasing; In the past 2006-2016 years, the number of documents has increased rapidly, which shows that spatial interpolation research and application become more and more important.

![Figure 1. Time distribution map of published and cited documents.](image)

In order to reveal the development of spatial interpolation research more clearly, CiteSpace software was used to analyze its national characteristics and subject category characteristics. Centrality refers to the correlation between subjects, the higher the centrality, the stronger the correlation between disciplines, the higher the activity, the lower the contrary. In the national characteristics, from the frequency point of view, the most published countries are the United States,
but also the earliest research countries, followed by China, Canada, Germany, Spain, the United Kingdom and Turkey. According to the centrality, Germany (0.25), Holland (0.19), France (0.18), China (0.14) and the United States (0.11) were ranked; Generally speaking, the more publications are published, the more active the country is, and it plays a combined role in international cooperation. From the subject category, the top five are environmental science and ecology, engineering, environmental science, meteorology and atmospheric science and geology. It can be seen from the table 1 that spatial interpolation is the research topic of multidisciplinary concern, and is widely used in various university disciplines and research fields, There are different degrees of correlation among disciplines, and interdisciplinary is strong.

| order | Literature quantity | Centrality | Year | Subject category                          |
|-------|---------------------|------------|------|------------------------------------------|
| 1     | 216                 | 0.25       | 1991 | ENVIRONMENTAL SCIENCES & ECOLOGY         |
| 2     | 204                 | 0.29       | 1993 | ENGINEERING                              |
| 3     | 182                 | 0.24       | 1991 | ENVIRONMENTAL SCIENCES                   |
| 4     | 161                 | 0.07       | 1994 | METEOROLOGY & ATMOSPHERIC SCIENCES       |
| 5     | 145                 | 0.07       | 1993 | GEOLOGY                                  |
| 6     | 142                 | 0.04       | 1991 | WATER RESOURCES                          |
| 7     | 142                 | 0.07       | 1993 | GEOSCIENCES, MULTIDISCIPLINARY           |
| 8     | 107                 | 0.18       | 1995 | COMPUTER SCIENCE                         |
| 9     | 97                  | 0.13       | 1986 | AGRICULTURE                              |
| 10    | 78                  | 0.17       | 1992 | MATHEMATICS                              |

3.2 Literature co-citation analysis
Literature co-citation analysis can help people to analyze the evolution of a theme through key nodes, clustering and color. The literature was co cited with 1029 cited citations and 24452 indirect cited literatures using CiteSpace. The co-citation map is generated as shown in Figure 2. From the graph, it can be seen that the spatial interpolation network has strong research focus, and there are many branches, showing clearly the correlation between clustering groups and clustering groups. There are eleven clusters in the atlas, respectively #0 artificial neural network, #1 grid size, #2 mountainous region, #3 different interpolation technique, #4 daily precipitation, #5 deterministic problem, #6 exploring method, #7 comparing geospatial technique, #8 middle yarra river catchment, #9 missing precipitation record, #10 kriging interpolation precision.

#0, #3, #5, #6 and #7 represent the theory of spatial interpolation in clustering groups, and the #0 clustering group has earlier reference years than other groups, and began to appear in 1990, it shows that the neural network has been applied to spatial interpolation in the early stage. #3 is a clustering group with different spatial interpolation, and spatial interpolation is widely used in different fields, different data types, such as sampling density, sample space distribution and other factors will lead to different spatial interpolation effects. #5 is a deterministic interpolation method, which mainly uses mathematical functions such as inverse distance weighting method and radial basis function to predict spatial interpolation. #6 is a method of exploration, and the spatial phenomena in the real world are very complex, which makes it difficult to meet the needs of ordinary spatial interpolation methods, computer technology is used to study various complex spatial modeling problems, which provides an important perspective and tool for the study of spatial interpolation. Wong\cite{8} et al. Combine neural network, self-organizing method and fuzzy mathematics to propose a new method for interpolating water in space, tabios\cite{9} et al. estimate annual precipitation in different places by comparing different interpolation techniques, Daly\cite{10} and Lin\cite{11} provide the basis for the development of interpolation technology. Group #2, #4, #8 and #9 focus on the practical application and case study of spatial interpolation, #2 was concentrated in 1985-1995 years, and it appeared earlier, which indicated that spatial interpolation began to be applied in the field of geology in the early stage; The spatial interpolation target is the estimation of the insufficient or missing data, and the interpolation is used to
predict the complete spatial distribution of the observed variables in the region\cite{12}, different spatial interpolation methods are used for data types and variables in different disciplines. Li\cite{13} analyze the characteristics of commonly used methods, and they are widely used in various disciplines, such as meteorology and water resources, environmental science and geology and other fields; In addition, the spatial interpolation method has also been applied in the fields of social economy and medicine. \#1 and \#10 represent the accuracy and efficiency of interpolation, with the wide application of spatial interpolation method, the accuracy and accuracy of spatial interpolation are becoming more and more important; The higher the interpolation accuracy is, the closer the spatial prediction is to the real situation. Wang\cite{14} and Ramesh\cite{15} respectively use computational grids to create continuous surface by spatial interpolation, which improves the computational efficiency and accuracy.

![Figure 2. Spatial interpolation co-citation network map.](image)

### 3.3 Keyword analysis

Keywords are highly abstract topics, covering the core point of a paper, so keyword analysis is conducive to mining the research hot spots and the evolution of the hot spots. CiteSpace software is used to analyze the keywords of the research object, and keywords co-occurrence network time-zone evolution is shown in figure 3. The words of the high number display for larger nodes.
Figure 3. Spatial interpolation keywords time-zone map.

Time-zone evolution diagram can directly reflect the research preface and its derivative relationship in different fields at different time periods, and then identify the future development direction. A lot of spatial interpolation research focus is in the period of 1992-2000, such as Kriging, interpolation, model, geostatistics, spatial interpolation and so on; The center of these keywords is high, and the relationship between keywords is relatively close; In this stage, the research and discussion of spatial interpolation theory, method and algorithm are focused. Spatial interpolation, model, Kriging and other keywords show higher centrality, which is the core node of the whole research field; geostatistics and other keywords provide the theoretical basis for spatial interpolation. 2001-2010 years is the rapid development stage of spatial interpolation research, resulting in a variety of research topics; The period that the number of key words more, higher centrality as variability, temperature, climate, rainfall, prediction and regression analysis and uncertainty, interpolation method, variable, elevation and so on. Through the analysis of key words, we found that the research of this period mainly focused on the optimization of methods and case study. Study on spatial interpolation method for the breakthrough method itself has been basically at a standstill, the new method is very difficult to produce, has been studied from the research of the method itself transferred to the improvement of the method, emphasizing practicality. Spatial interpolation is an indispensable tool in the data processing of natural science and Social Sciences, especially in hydrology, climate, meteorology, environment and geology. 2011-2017 years keywords show small centrality and more keywords, which shows that the spatial interpolation research subject enters the branch expansion state. Spatial interpolation goes towards diversified development; While spatial interpolation is applied to various disciplines, spatial interpolation is more widely applied according to the intersection and connection between disciplines.

4. Conclusion
In this paper, the use of information visualization software CiteSpace to spatial interpolation of relevant literature information analysis, the spatial interpolation with the time development matures, identify 11 clustering knowledge group, and analyzes the development and evolution of the relationship between groups; and the evolution of theme of each period and research hotspots. In general, the spatial interpolation research started earlier and the discipline intersection is strong; It shows that the spatial interpolation method has broad prospects for application in the study; There is no optimal method for spatial interpolation itself, only the most appropriate method for the actual situation of each discipline research area. Sample data need to be fully analyzed and experimental comparisons are made to select the best method. It is important to improve the method based on the
interpolation method and the characteristics of the data itself and the subject to obtain better spatial interpolation method.

Reference

[1] Huang Xingyuan and MA Jinsong 2008 *Introduction to Geographic Information System. Third Edition* (Higher Education Press) p93-103.

[2] Wu Lun 2001 *Geographic Information System: principles, methods and Applications* (Science Press) p178-192.

[3] Li Xin, CHENG Guodong and LU Ling 2000 Comparision of spatial interpolation methos *J. Progress in Earth Sciences*. 15(3) 260-5.

[4] Luo Ming and Pei Tao 2009 Renew on Soft Spatial Data and its Spatial Interpolation Methods *J. Progress of Geographical Science*. 28(5) 663-72.

[5] He Wenjiao, Yue Tianxiang, Shi Xiaoli et al. 2012 Research progress of spatial interpolation method and its accuracy for soil continuous attributes *J. Journal of natural resources*. 2012(1) 163-75.

[6] He Hongyan, Guo Zhihua and Xiao Wenfa 2005 Review on spatial interpolation techniques of rainfall *J. Journal of ecology*. 24(10) 1187-91.

[7] Chen Yue, CHEN Meichao, LIU Beiyuan, et al. 2015 The methodology function of CiteSpace mapping knowledge domains *J. Science studies*. 33(2) 242-53.

[8] Wong K W, Wong P M, Gedeon T D, et al. 2003 Rainfall prediction model using soft computing technique *J. Soft Computing*. 7(6) 434-8.

[9] Iii, Guillermo Q. Tabios, and J. D. Salas 1985 A comparative analysis of techniques for spatial interpolation of precipitation *J. Jawra Journal of the American Water Resources Association*. 21(3) 365-80.

[10] Daly C, Helmer E H and Quinones M 2003 Mapping the climate of Puerto Rico, Vieques and Culebra *J. International Journal of Climatology*. 23(11) 1359-81.

[11] Zhu Qian, Zhang Wanchang and Yu Junhui 2004 The spatial interpolations in GIS *J. Journal of Jiangxi Normal University (NATURAL edition)* 28(2) 183-8.

[12] Li J 2008 A Review of Spatial Interpolation Methods for Environmental Scientists *J. Record - Geoscience Australia*. 137.

[13] Wang S and Armstrong M P 2003 A quadtree approach to domain decomposition for spatial interpolation in Grid computing environments *J. Parallel Computing*. 29(10) 1481-1504.

[14] Teegavarapu, Ramesh S. V, T. Meskele, and C. S. Pathak 2012 Geo-spatial grid-based transformations of precipitation estimates using spatial interpolation methods *J. Computers & Geosciences*. 40(3) 28-39.