Knowledge map of Yellow River Basin’s water resources based on literature analysis

Y Y Liu¹, R G Jiang¹², J C Xie¹, S M Wang¹ and W Li¹

¹ State Key Laboratory of Eco-hydraulics in Northwest Arid Region of China, Xi’an University of Technology, Xi’an 710048, China
² E-mail: jrengui@163.com

Abstract. The Yellow River Basin (YRB) is an important economic zone in China. The issue of water ecology has attached widely attention by scholars and society. This paper systematically reviewed the status, hotspots and development tendency on water resources research in the YRB based on the China National Knowledge Infrastructure (CNKI) database from 1992 to 2019. The SATI4.0 was used to analyze the annual publication, journal sources and keywords. High-frequency keywords were extracted, and software including NetDraw, SPSS and CiteSpace were used to analyze the co-occurring keywords, clustering and research frontier. Results showed that: (1) The number of published papers had increased before 2005 and decreased after it. The quantity of paper has increased sharply since 2018, indicating that the research has a new direction. (2) The source of journals mainly concentrated on basic subjects and hydraulic engineering. (3) The cluster analysis shows that hydrologic features, water management and protection and the water rights system are hotspots. (4) The result of research frontier indicates that sensitivity, ecological protection and high-quality development of water resources have ample research perspectives. It is general trend to combine the scientific methods with traditional methods. The results are conducive to systematically understand research advances and trends of water resources under changing environment, thereby promoting the construction of ecological civilization in the YRB.

1. Introduction

With the high-speed development of global economy and the increase of population, the aquatic ecosystem is facing a huge threat. Under the global climate change and human activities, global hydrological cycle has changed which exerted great effects to multiple scale fresh water availability [1]. The Yellow River Basin (YRB) stands a significant position in socio-economic development and ecological civilization construction in China. Therefore, the water resources protection and high-quality development of the YRB have attached great concerns.

Many previous studies focused on the water resources of the YRB. For example, Peng et al. established an integrated optimization layout scheme of food production, energy development and water resources allocation based on multi-factors equilibrium intelligent algorithm in 2017. Liu and Zhang used the Penman-Monteith to analyze the sensitivity and regional differences of potential evapotranspiration to meteorological variables in the watershed in 2011. They found that the maximum temperature is the most sensitive factor to the YRB. Ren et al. found that climate change affected the hydrologic cycle in 2008. Yang et al. simulated the spatial and temporal distributions of water, and simulated seasonal changes of river discharge, evaporation and soil moisture in 2004. Besides, the...
proposal of water resources carrying capacity, water rights allocation scheme also plays an indispensable role in this area [2].

To provide a more comprehensive understanding of water resources in the YRB, this paper analyzed the research status, discipline distribution, research hotspots and developing trend of water resources in the YRB based on the core literature from China National Knowledge Infrastructure (CNKI) database using bibliometric analysis and knowledge map methods.

2. Data and methodology

2.1. Data collection

The data used for the bibliometric analysis was collected from CNKI, which has an internationally leading digital library technology. To maximize reliability and typicality, the data gathered from Scientific Citation Index (SCI), the Engineering Index (EI), core journals, Chinese Social Sciences Citation Index (CSSCI), and Chinese Science Citation Database (CSCD) in CNKI. Considering the recall ratio and precision ratio, we set the searching terms as “Yellow River Basin” AND “water resources” OR “Yellow River” AND “water resources”, and the timespan to 1992-2019. In this systematic review, all literature was in Chinese. A total of 1472 articles were retrieved. Duplicate and conference articles were not included. Considering the format problem of numbers and letters in Chinese, we made a manual cleaning by the Statistical Analysis Toolkit for Informetrics (SATI 4.0) [3]. Finally, 1349 papers were selected for bibliometric analysis.

2.2. Methods

With the development of information technology, information sources become more and more huge, and information visualization becomes very important [4]. The scientific knowledge map is an image that shows the development process and the relationship between the structure of scientific knowledge domain [5].

Based on bibliometrics and scientific knowledge maps, this paper used a synthesized method that combines qualitative and quantitative analysis. We selected the amounts of articles, source journals and keywords by SATI4.0 as the research objects to analyze the characteristics from four aspects including time series distribution, subject distribution, research hotspots and research frontiers. It is helpful to provide a preliminary understanding YRB’s water resource. I. Time series analysis of literature. Sorting the annual frequency of literature according to time to analyze different research stage. II. Subject analysis. According to the source of journals to reveal its research level, subject distribution status, and its attention degree. III. Hotspot analysis. The high-frequency keywords were analyzed using g index of keywords with visualization tools (UCINET) and social network analysis (NetDraw). System clusters by SPSS were used to discuss knowledge structure and reveal the hotspots [6]. IV. Research process and frontier analysis. A keywords-burst table and a hotspot migration map generated by CiteSpace can reveal the evolution process and future development trend of the research field [7].

3. Results and discussion

3.1. Time series analysis of literature

The changes in the quantity of literature can be used to analyze the development stage and growth trend of a particular discipline [8], and reflect the research speed, scale and level. In order to understand the development status, the annual frequency of papers from 1992 to 2019 were calculated. Figure 1 shows that the subject variation since 1992. The number of papers had increased from 1995 to 2005. The upward trend was more pronounced in the latter 5 years than the first one. This is mainly because the Yellow River Commission issued a five-year plan in 2001, emphasizing reform and development of the yellow river from 2001 to 2005. Thus, this research became a focus of scholars and published papers increased. So, the frequency reached its peak in 2005. After 2005, the research in this field has moved from basic research to information research, and the pace slowed down. Then this subject was in a
fluctuating decline from 2006 to 2018. In September 2019, the ecological protection and high-quality development of the YRB has become a national strategy. The volume of papers issued in 2019 has increased dramatically.

3.2. Subject analysis

Journal can reflect subject distribution. According to the source of journals of these papers, the top 10 journals with the most papers were shown in Table 1. Among the 1349 papers, a total of 531 articles were published in the Yellow River, accounting for 39.36%. The Journal such as Yellow River, China Rural Water and Hydropower belong to water resources and hydropower engineering. Journal of Natural Resources belongs to environmental science and resource utilization journals, indicating that the water ecology and environmental issues have get considerable attention. Each journal has its own thematic category. Papers with similar category are published in same journals. It is found that more than half of the literature mainly distributed in water resources and hydropower engineering, manifesting that researchers have focused on water conservancy and hydropower projects.

| Journals                                | Quantity of Literature | Percentage (%) |
|-----------------------------------------|------------------------|----------------|
| Yellow River                            | 531                    | 39.36          |
| China Rural Water and Hydropower        | 34                     | 2.52           |
| Journal of Natural Resources            | 33                     | 2.45           |
| China Water Resources                   | 31                     | 2.30           |
| Journal of China Hydrology              | 29                     | 2.15           |
| Resources Science                       | 26                     | 1.93           |
| Advances in Water Science               | 25                     | 1.85           |
| Journal of Hydroelectric Engineering    | 24                     | 1.78           |
| Journal of Arid Land Resources and Environment | 23                 | 1.71           |
| Water Resources and Hydropower Engineering | 20                   | 1.48           |

3.3. Hotspot analysis

Title words and keywords are often selected to reveal the hotspots and discover the emerging trends in scientific research [9]. Based on the g index, this paper sorts keywords according to their frequency from largest to smallest. Keywords with a cumulative frequency greater than the square of serial number are
high-frequency keywords until the cumulative frequency is less than it. This paper extracts 44 high-frequency keywords from 2013 times totally, as shown in Table 2.

Table 2. Part of high-frequency keywords measured by g index.

| Number | Number | Keywords                        | Frequency | Cumulative |
|--------|--------|---------------------------------|-----------|------------|
| 1      | 1      | yellow river basin              | 721       | 721        |
| 2      | 4      | water resources                 | 343       | 1064       |
| 3      | 9      | climate change                  | 70        | 1134       |
| 4      | 16     | yellow river cut off            | 53        | 1187       |
| 5      | 25     | water management                | 52        | 1239       |
| 6      | 36     | water right transfer            | 51        | 1290       |
| 7      | 49     | yellow river irrigation area    | 47        | 1337       |
| 8      | 64     | south-to-North water diversion  | 45        | 1382       |
| 9      | 81     | optimal allocation of water     | 40        | 1422       |
| 10     | 100    | runoff variation                | 39        | 1461       |

3.3.1. Keyword co-occurrence analysis. In order to reveal the research direction and characteristics of the research field, an information visualization map drawn by UCINET and NetDraw. As shown in Figure 2, setting node size by “betweenness”, which can measure the connection strength between keywords through the size of them. The more times that node acts as an “betweenness”, the greater the betweenness.

According to the co-occurring map, the research status of water resources in the YRB are as follows:
I. Apart from the YRB and Water Resources, Climate Change has the highest frequency with 70 times. It is obvious that basic sciences such as climate, runoff, and precipitation are gradually lucubrated. II. The problems of Yellow River cut off, water pollution, and sediment have been widely concerned, and the coping solutions were proposed. III. Water resources management, water rights systems have also attached wide attention, and government constantly proposed new requirements. IV. The frequency of ecological environment, sustainable development, and renewable capacity of water resources are relatively high, which has great practical significance for the ecological protection and high-quality development of the YRB. V. The combination of traditional approach and modern information technology is an effective tool to study the current hotspots and difficult issues.
3.3.2. **Keyword clustering analysis.** The clustering analysis is used to classify multiple research variables (keywords) [10], which reflects the intellectual structure. Figure 3 shows that keywords are divided into 6 categories by the squared Euclidean distance between them.

The first category mainly focuses on the management and protection of water and soil, and its application and practice. Category II devotes the management and rational use of water. Category III reflects the problem of uneven spatial and temporal distribution of water resources in China. These problems can affect the economy, and the demand for optimal allocation of water resources is more urgent. Category IV mainly focus on the impact of the changing environment and human activities on the water resources from the view of the basic science. Category V concentrates on the interruption of YRB and related issues, the causes and solutions are also discussed. Category VI studies the Yellow River Delta and wetlands and the sustainable development.

![Figure 3. The keywords clustering map.](image)

3.4. **Research process and frontier analysis**

Research frontiers are the research issues and fields with the feature of the newest, advanced and potential development [11]. Importing data into CiteSpace and setting keyword as node to extract the top 30 high-frequency keywords of each two years. Simplifying the network with pathfinder algorithm to highlight the main structural features. The pathfinder algorithm utilizes the triangle inequality formula to calculate the weight of pace between two keywords. If the weight of pace of the two keywords is greater than their distance, then delete this path [12]. As shown in Figure 4, the circle represents the keywords. It can infer that the research content of different period according to temporal progression in terms of keywords to get the research progress. Keywords in the black rectangle box are the current research hotspot and also are the future research direction of this field.

Frequency Burst in CiteSpace can detect mutation frontier terms. If the frequency of a term has significantly increased in a short period of time, we can infer that this term represents a developing trend [13]. Table 3 shows the top 1 or 2 keywords each year in strength of frequency burst. The progress bar turns into red when a new keyword occurs. The strength reflects the importance of the keywords.
Based on the theme path evolution map and the keywords burst, the phase of the research is as follows:

(1) Early Enlightenment Stage (1992-1995). The research at this stage mainly focused on the planning, development of water resources in the YRB. Apart from that, the impacts on economy also play an important role. It is demonstrated that scholars pay more attention to the water resources planning and the economic benefits to society. It is found that a few researchers started to study the optimal allocation of water resources since 1993, but plenty of this study occurred in 2010.

(2) Practical Development Stage (1996-2005). This period mainly continued the direction of previous research. On this stage, lots of hydraulic infrastructure were built to improve the water security system. The research on water dispatch and hydraulic engineering has become mainstream. In 2001, the proposal of National Tenth Five-Year Plan made ecological construction and sustainable development necessary to pay more attention. Scholars also came up with the Digital Yellow River Project in 2001, which lead to an informatization research stage.

(3) Test and Introspection Phases (2006-2018). At this stage, the impact of human activities and the water resources carrying capacity was taken seriously. The optimal allocation of water and sediment is still a problem. Additionally, water right system was established. It is difficult to solve the shortage of water resources fundamentally. Scholars explored it from the perspective of macro-control and policy, combined with modern information technologies such as remote sensing.

(4) Research Breakthrough Phases (2018 till now). Scholars mainly investigate the spatial distribution and comprehensive evaluate the future evolution of water resources, especially emphasize

---

**Figure 4.** The hot migration of keywords.

**Table 3.** Part of keywords with the strongest citation bursts.

| Keywords                              | Strength | Begin | End   | 1992 - 2019       |
|---------------------------------------|----------|-------|-------|--------------------|
| economy                               | 3.6066   | 1992  | 1999  |                    |
| yellow river cut off                  | 13.3007  | 1996  | 2000  |                    |
| countermeasure                        | 10.8659  | 1996  | 2000  |                    |
| sustainable development               | 3.8612   | 2000  | 2008  |                    |
| the source region of the yellow       | 4.3973   | 2006  | 2019  |                    |
| climate change                        | 8.9679   | 2007  | 2019  |                    |
| water right transfer                  | 6.3902   | 2009  | 2013  |                    |
| optimal allocation of water           | 4.2741   | 2010  | 2015  |                    |
| water management                      | 4.4045   | 2011  | 2016  |                    |
| human activity                        | 3.9749   | 2012  | 2019  |                    |
the sensitivity of water resources. A national strategy for ecological protection and high-quality development in the YRB was proposed in 2019.

4. Conclusions

- The year of 2005 is a landmark year for water resources research in the YRB. Prior to 2005, the hotspots mainly focus on water utilization, water saving and water resources protection. After 2005, something new occurs, such as integrated management and optimal allocation of water resources.
- Recent research emphasized on the water ecology and hydropower engineering. At present, a large number of scholars focused on basic science and water conservancy projects. However, few studies focus on the water resources management and planning in the YRB, which should pay more attention in the future research.
- Hydrologic features, water resources management, water resources integrated planning, and water rights system are hotspots. Many problems in the YRB promoted water resources carrying capacity, ecological protection and high-quality development to become important trends for future research. The rapid development and wide application of modern information technology provide strong technical support for the future research.
- In recent years, the data exchange under the background of big data is intricate, and the development of digital scientific research needs to emphasize and standardize scientific data management. Vigorously promoting the development and utilization of scientific data resources and open sharing, which will effectively promote the integration and development of scientific literature metrology.

Acknowledgments

This paper was partly funded by the National Key Research and Development Program of China (Grant No. 2016YFC0401409), National Natural Science Foundation of China (Grant No. 51679188, 51509201), Young Technology Star in Shaanxi Province of China (Grant No. 2020KJXX-092). The authors thank the editor and anonymous reviewers for their valuable comments and suggestions.

References

[1] Xu H M, Liu L L, Wang Y, Wang S, Hao Y and Ma J J 2019 Hydrol. Earth Syst. Sci. 23 4219-31
[2] Zhou P and Leydesdorff L 2006 Res. Policy 35 83-104
[3] Liu Q Y and Ye Y 2012 J. Inform. Resour. Manag. 2 50-8
[4] Li W, Jiang R, Zhao Y, Xie J, Zhu J and Cao R 2019 J. Coast. Res. 93 9-15
[5] Chen Y, Chen C M, Liu Z Y, Hu Z G and Wang X W 2015 Studies Sci. Sci. 33 242-53
[6] Egghe L 2006 Scientometrics 69 131-52
[7] Chen C M 2010 J. Am. Soc. Inf. Sci. Technol. 61 1386-409
[8] Wang Y, Liu W, Li G, Yan W and Gao G 2019 Water 11 20
[9] Zhang Y, Huang K, Yu Y and Yang B 2017 J. Clean. Prod. 149 70-9
[10] Zhao L M and Zhang H 2019 Inform. Sci. 37 97-104
[11] Chen Q Q, Zhang J B and Huo Y 2016 Agric. Econ. 62 429-45
[12] Chen C M, Chen Y, Hou J H and Liang Y X 2009 J. China Soc. Sci. Tech. Infor. 28 401-21
[13] Song J, Zhang H and Dong W 2016 Scientometrics 107 1111-47