Visual analysis of the evolution and focus of agricultural drought research on crops

Shanli Yang[a,b,c,d], Changxiu Cheng[a,b,c,d*], Shi Shen[a,b,c,d], Jing Yang[a,b,c,d]

[a] *Key Laboratory of Environmental Change and Natural Disaster, Beijing Normal University, Beijing 100875, China*

[b] *State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China*

[c] *Faculty of Geographical Science, Beijing Normal University, Beijing100875, China*

[d] *Center for Geodata and Analysis, Beijing Normal University, Beijing 100875, China*

Abstract: Big Data analysis was applied in this study to visualize and analyze the developing trends and emerging topics in agricultural drought research. In total, 10,048 agricultural drought-related articles (From Jan 1900 to Dec 2015) were indexed from the Web of Science. First, a co-citation network of these articles was generated and a semantic cluster in Citespace was created. The cluster results showed there were two periods from the point of view of research topics and methods: (1) pre-1990, when the topic of agricultural drought research was focused on the crop morphological features such as the size of leaf area, the length of root and so on; and (2) post-1990, when the topic of agricultural drought research was focused on the genetic analysis of crops. Then the evolution of agricultural drought research was presented in details. Next, a hybrid network of countries and key noun phrases were generated to detect the major topics of crop drought research in different countries. The results showed the following region preferences for study: China, northern region; Thailand, northeast region; Canada, western region; USA, northern great plains; and France, clay soil regions. In addition, the following favored drought research topics were identified: Australia, faba beans; India, supplemental irrigation; Iran, irrigation regimes, and South Korea, gene expression.

Keywords: agricultural drought research; visualized analysis; Complex Networks; CiteSpace

Citation: Yang, S., Cheng, C., Shen, S., & Yang, J. (2018). Visual analysis of the evolution and focus of agricultural drought research on crops. International Agricultural Engineering Journal, *27*(2), 348–352.

1. Introduction

Drought is by far the most important environmental stress in agriculture and many efforts have been made to improve crop productivity under water-limiting conditions (Ashraf, 2010). Although natural selection favors mechanisms for adaptation and survival, breeding activity has directed selection towards increasing the economic yield of cultivated species. More than 80 years of breeding activities have led to yield increases in drought environments for many crop plants (Cattivelli et al., 2008). Although fundamental research has provided significant gains in the understanding of the physiological and molecular responses of plants to water deficits, there are few reviews or summaries outlining the panoramic view of this scientific field. Yan et al. (2016) used CiteSpace to analyze the history and status of crop transgenic breeding research based on the SCI database; however, their study did not systematically summarize agricultural drought development research. The present study could answer questions such as: How does agricultural drought evolve? How do the countries contribute to agricultural drought research and development? Are my ideas novel, or have they already been addressed? These questions are still posed by administrations, policy makers, and scientists.

Complex network analysis is used for mining Big Data. In particular, it has been utilized to analyze the Big Data from social media, such as Facebook, Twitter and Weibo (Topirceanu, Duma, & Udrescu, 2015; Yuan & Liu, 2015), as well as to evaluate the vulnerability and

Received date: 2018-03-23 **Accepted date:** 2018-06-15

* **Corresponding author: Cheng, C.,** State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China. Tel: +86-010-58807241, Email: chengcx@bnu.edu.cn.

stability of worldwide transportation networks (Tran & Namatame, 2015). Complex network analysis, rather than statistical analysis, has been valued and employed by scientometrics to investigate various scientific fields (Chen, Dubin, & Kim, 2014; Liu, Rui, & Yan, 2014; Liu, Sun, & Song, 2014; She et al., 2014; Shen et al., 2018). Specifically, the research field related to the Wenchuan earthquake has been analyzed using scientometrics and scientific visualization techniques (Qian, 2012). However, the analysis is only focused on literature concerning an individual earthquake. Complex network analysis has not been used to analyze the research field of agricultural drought.

In this study, three networks were constructed, i.e., a network of co-citation, a network of co-citation by cluster, and a hybrid network of countries and key noun phrases. The network of co-citation depicts the relationship of different authors and discovers the key scholars. The co-citation cluster analysis displays the macro-structural evolution of scientific knowledge. Additionally, the emerging noun phrases are detected and identified as research hotspots. The hybrid network of key noun phrase and countries were analyzed to explore research foci of different countries.

The organization of this paper is: Section 2 introduces the tools and data; Section 3 explores the data distribution, creates a co-citation network of agricultural drought, clusters the co-citation network of agricultural drought and builds a hybrid network of agricultural drought noun phrases and countries. Section 4 makes some conclusions and some insights for further research areas.

2. Tools and data

2.1. Research tools

In this study, CiteSpace was used as the research tool to assist in the analysis. CiteSpace is a Java application developed by Dr. Chaomei Chen for analyzing and visualizing emerging trends and changes in scientific literature (Zhao & Wang, 2011). It can produce networks consisting of nodes and links. The resultant networks visualize patterns via different indicators. The color of the links shows the first year of co-occurrence of the two nodes. The size of the circle is proportional to the frequency of the appearance in the literature. The frequency represents the degree of recognition in academic circles, and it reflects the academic contribution of corresponding literature. Purple rims of nodes indicate pivotal points with high betweenness centrality. The betweenness centrality stats to the node act as a bridge in the development of a scientific field linking research in different time periods. The red circle represents the node having a phenomenon of a citation burst.

2.2. Data collection

The input data for CiteSpace were retrieved from the SCI database via the Web of Science, and were based on papers with “drought” as the topic published during the period from 1900 to 2015. The search was limited to “articles”; the field was limited to “agriculture”; and the language was limited to English. The resultant dataset contained a total of 10,048 records. A bibliographic record in SCI contains fields such as author, title, abstract, keywords and references.

3. Results

Figure 1 represents the total number of articles retrieved from the Web of Science from 1900 to 2015; which in this domain appeared primarily during the early 1990s and gradually increased to reach a total of 950 articles published in 2015. There are two important years of particular interest: 1990: the number of articles after 1990 increased and more than doubled that in 1991; and 2002: the number of articles also increased substantially after 2002.

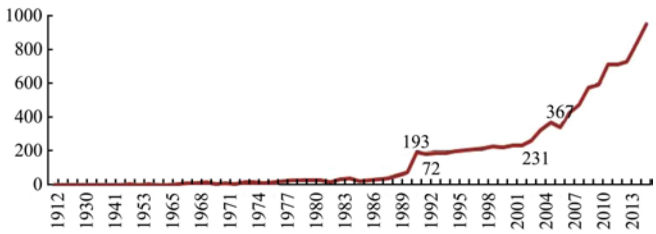

Fig. 1. Yearly published articles.

3.1. Creating a co-citation network of agricultural drought

The co-citation network (Popp, et al., 2016) was employed to study the evolution of the subject. In CiteSpace, a time span of 1990 to 2015 and a slice length of 1 year were set by the authors. The node type selected was “cited reference”. The “top N per slice” was set as “50”, which means that 50 documents with the highest cited frequency were selected for each “time slice”. The nodes and lines in the network were then generated automatically. The overview of the co-citation network in the field of agricultural drought from 1990 to 2015 is presented in Fig. 2. The color of a link in Fig. 2. varies from blue to yellow and finally to red, which indicates the year of the corresponding literature that is co-cited. The earliest publication of literature in the field corresponds with the blue lines and the latest publication with the red lines in the network. The size of the circle is proportional to the cited frequency of the literature and purple rims of nodes indicate pivotal points. Fig. 2. shows that the study of agricultural drought began in the blue area, then the green area, and finally the red area.

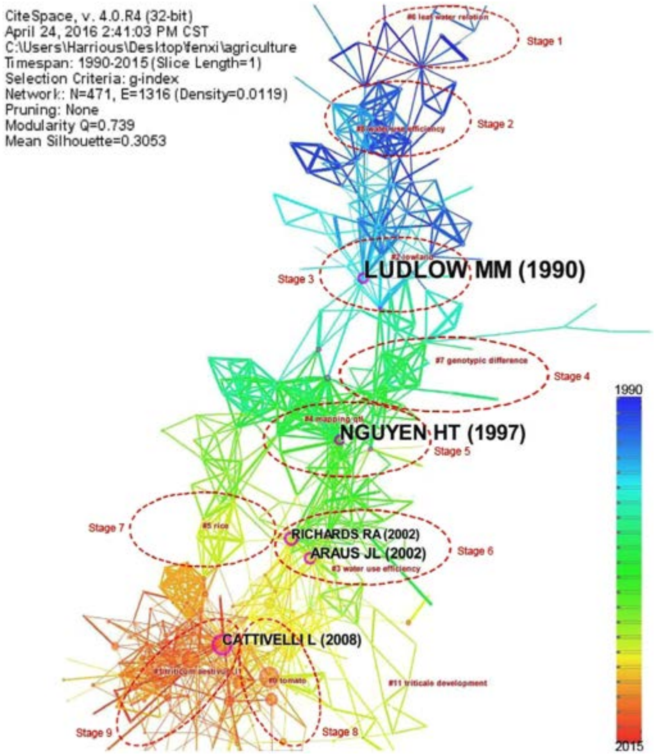

Fig. 2. Cluster view of the co-citation network.

3.2. Clustering the co-citation network of agricultural drought

Table 1 presents clustering results based on co-citation references. From Table 1, the evolution in agricultural drought research can be followed and can be divided into two levels: the macro level and the micro level. Prior to 1990, authors focused on the crop surface and the adaptation of the plant to the environment (Ludlow & Muchow, 1990), and after 1990, the scholars moved their interest into molecular studies (Champoux et al., 1995; Fukai & Cooper, 1995).

Table 1 Reference clustering result.

Cluster ID	Size	Silhouette	Mean (Year)	Label (LLR)
1	55	0.857	2008	*Triticum aestivum* l; common bean; wheat population
0	57	0.683	2006	Tomato; glycine betaine; salicylic acid
11	8	0.981	2006	Cereal; early stage; triticale development
5	27	0.902	2004	Rice; genetic analysis; heat tolerance
9	20	0.934	2004	Soil respiration; precipitation pulse; birch effect
3	49	0.798	2001	Water use efficiency; grapevine; chile
4	43	0.792	1997	Rainfed lowland rice; upland rice; mapping qtl
7	24	0.951	1995	Durum wheat yield; Mediterranean basin; genotypic difference
10	9	1	1991	Tall fescue; endophyte effect; persistence
2	51	0.866	1990	Lowland; selection; cycle
8	23	0.936	1988	Water-use efficiency; alfalfa germplasm; variation
6	24	0.987	1985	Leaf-water relation; cell-membrane stability; water-stress

The focus of the evolution in agricultural drought can be divided into nine stages. Stage 1 occurred in the year 1985, when scholars focused on leaf-water relationships. Stage 2 occurred in 1988 with a focus on water use efficiency. Stage 3 occurred in 1990 with a focus on the lowland environment and the survival environment of crops (Bolanos & Edmeades, 1993). Stage 4 was in 1995 with a focus on genotypic differences, which represented a big shift. Stage 5 occurred in 1997 when the focus was on mapping qtl (Nguyen, Babu, & Blum, 1997). Stage 6 occurred in 2001 with a focus on water use efficiency again (Richards et al., 2002; Araus & Royo, 2002). Stage 7 occurred in 2004 with a focus on rice and soil respiration. Stage 8 occurred in 2006 with a focus on tomatoes (Tian et al., 2006) and cereal. Finally, stage 9 occurred in 2008 with a focus on Triticum aestivum l (a type of wheat). From Table 1, every cluster in the silhouette had a value greater than 0.6, indicating that every cluster was efficient and valid.

3.3. Creating a hybrid network of agricultural drought noun phrases and countries

An analysis of terms such as noun phrases can help us identify hot topics of crop drought research in different countries. Fig. 3. shows a hybrid network of noun phrases and countries. Noun phrases are shown as rectangles, whereas countries are shown as circles. The size of the circles in the map is proportional to the total articles from a country from 1900 to 2015. The circles showing purple indicate that the country is a key node in this hybrid network. In Fig. 3., a number of conclusions can be drawn. First, the USA plays a leading role in this field of study, as it has the highest number of articles (China is second followed by India and Iran). Second, different countries have specific places to study drought, for example, for China it is the northern region, for Thailand the northeast region, for Canada the western region, and for America the northern great plain. Third, every country has a different perspective on drought, for example, France studies summer dormancy, Australia studies faba bean, India studies supplemental irrigation, Iran studies water deficit stress and irrigation regimes, Germany studies leaf area index, South Korea studies gene expression, Brazil studies irrigation systems, Zimbabwe studies soil fertility, and Switzerland studies kernel weight, and Mexico studies moisture regimes.

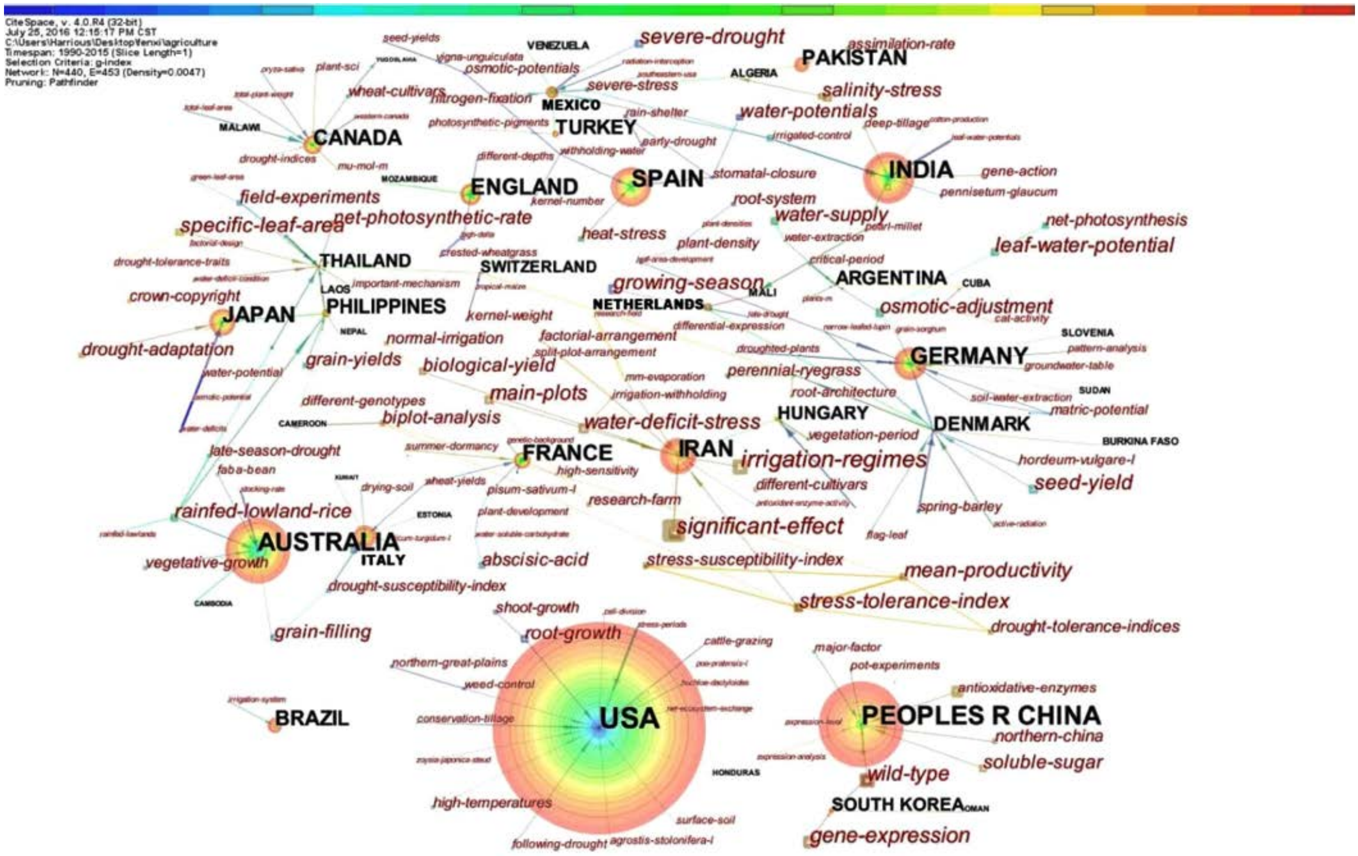

Fig. 3. A hybrid network of countries (regions) and emerging topics of 440 nodes and 453 links.

4. Discussion

This paper takes advantage of CiteSpace to conduct a comprehensive analysis of agricultural drought research. CiteSpace is a powerful information visualization tool to help uncover the hotspot changes and space-time evolution of various co-citation networks. For agricultural drought, yearly publications sustained a notable growth as a whole from 1990 to 2015. The evolution of the research can be divided into nine stages: the focus of leaf-water relation in 1985, the focus of water use efficiency in 1988, the focus of lowland in 1990, the focus of genotypic differences in 1995, the focus of mapping qtl in 1997, the focus of water use efficiency in 2001, the focus on rice in 2004, the focus on soil respiration in 2004, the focus on tomatoes in 2006, the focus on cereal in 2006, and the focus on Triticum aestivum L in 2008. In fact, the study of agricultural drought can be divided into three main stages. The first stage focused on the surface analysis of the crops such as the size of leaf area, the length of root and other morphological characteristics. The second stage focused on the adaptation of the plant to the environment. The third stage focused on the genetic analysis of crops. These three stages also can be divided into two levels: the macro level and the micro level.

From Figure 2, this study can identify two important articles that summarize the most advanced methods to help crops to resist drought during that time. The first study was by Ludlow in 1990, who discussed the importance of improving crop yields in water-limited environments and contrasted the methods used to improve crop yields. The other study was conducted by Cattivelli in 2008, who analyzed three main approaches for improving the drought tolerance of crop plants.

The analysis of different countries' research focus shows that research of agricultural drought is conducted worldwide. Every country has a special perspective in agricultural drought. In conclusion, the results of this study can help us understand the development of agricultural drought research using Bibliometric and Scientometric methods. This can help researchers review the history and find a new breakthrough point in the study of agricultural drought.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [Grant numbers: 41771537].

[References]

Araus, J. L., & Royo. C. (2002). Plant breeding and drought in c3 cereals: what should we breed for. *Annals of Botany,* 89, 925–40.

Ashraf, M. (2010) Inducing drought tolerance in plants: recent advances. *Biotechnology Advances*, *28*(1), 169–183.

Bolanos, J., & Edmeades, G. O. (1993). Eight cycles of selection for drought tolerance in lowland tropical maize. 2. responses in reproductive behavior. *Field Crops Research*, *31*(3-4), 233–252.

Cattivelli, L., Rizza, F., Badeck, F. W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., Marè, C., Tondelli, A., & Stanca, A. M. (2008). Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field Crops Research*, *105*(2008), 1–14.

Champoux, M. C., Wang, G., Sarkarung, S., Mackill, D. J., O'Toole, J. C., & Huang, N. (1995). Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. *Theoretical and Applied Genetics,* 90: 969–981.

Chen, C., Dubin, R., & Kim, M. C. (2014). Emerging trends and new developments in regenerative medicine: a scientometric update (2000-2014). *Expert Opinion on Biological Therapy,* 14, 1295–317.

Fukai, S., & Cooper, M. (1995). Development of drought-resistant cultivars using physiomorphological traits in rice. *Field Crops Research, 40*(1995), 67–86.

Liu, G., Rui, J., & Yan, J. Sciatic nerve injury repair:a visualized analysis of research fronts and development trends. Neural Regeneration Research, 2014, 18: 1716–1722.

Liu, G., Sun, H., & Song, X. (2014) Visualizing and mapping the research on patents in information science and management science. *Malaysian Journal of Library and Information Science*, 19, 87–103.

Ludlow, M. M., & Muchow, R. C. (1990). A critical evaluation of traits for improving crop yields in water-limited environments 1. *Advances in Agronomy*, 43, 107–153.

Nguyen, H. T., Babu, R. C., & Blum, A. 1997, Breeding for drought resistance in rice: physiology and molecular genetics considerations. *Crop Science*, *37*(5), 1426–1434.

Popp, J., Kovács, S., Balogh, P., & Jámbor, A. (2016). Co-authorship and co-citation networks in the agricultural economics literature: the case of central and eastern Europe. *Eastern European Economics*, *54*(2), 153–170.

Qian, G. (2012). Scientometrics analysis on the research field of wenchuan earthquake. *Disaster Advances*, *5*(4), 704–707.

Richards, R. A., Rebetzke, G. J., Condon, A. G., & Van Herwaarden, A. F. (2002). Breeding opportunities for increasing the efficiency of water use and crop yield in temperate cereals. *Crop Science*, 42, 111–121.

She, J., Zhang, X., Wang, W., & Pablos, P. O. D. (2014). Mapping the impact of social media and mobile internet on Chinese academia's performance: a case on telemedicine research 2005-2013. *Journal of Universal Computerence,* 20, 2005–2015.

Shen S, Cheng C, Yang J, et al. Visualized analysis of developing trends and hot topics in natural disaster research. Plos One, 2018, *13*(1): e0191250.

Topirceanu, A., Duma, A., & Udrescu, M. (2015). Uncovering the fingerprint of online social networks using a network motif based approach. *Computer Communications*, *73*(2015), 167–175.

Tian, F., Li, D., Zhu, Z., Fu, Y., Wang, X., & Sun, C. (2006). Construction of introgression lines carrying wild rice (oryzarufipogongriff.) segments in cultivated rice (oryza sativa l.) background and characterization of introgressed segments associated with yield-related traits. Theoretical and Applied Genetics, (2006)112, 570–80.

Tran, Q. H. A., & Namatame, A. (2015). Worldwide aviation network vulnerability analysis: a complex network approach. *Evolutionary and Institutional Economics Review*, 12, 349–373.

Yan, Z., Zheng, H., Zhao, J., Chuan, L., Zhang, X., Tan, C., & Sun, S. (2016). Analysis of development situation of crop transgenic breeding research based on SCI database. *Journal of Agricultural Science and Technology*, *2016*(02), 208–215.

Yuan, W., & Liu, Y. (2015). A mixing evolution model for bidirectional microblog user networks. *Physic A Statistical Mechanics and Its Applications*, *432*(2015), 167–179.

Zhao, R., & Wang, J. (2011). Visualizing the research on pervasive and ubiquitous computing. *Scientometrics*, *86*(3), 593–612.