Chinese scholars in China and overseas: comparative analysis on research productivity and impact

*Wen Lou, Hui Wang and Siluo Yang**

The present article addresses the differences in research productivity and impact between Chinese scholars in China and overseas. A total of 1190 Chinese scholars in China and 1983 Chinese scholars overseas were recognized out of 6306 papers in six journals over 10 years. Research performance was evaluated by informetric indicators, including the absolute value, proportion, and average number of authors, publications, citations, and usage counts. Statistics metrics, including standard deviation and coefficient of variance, were used to test the viability of two groups. These metrics conclude that: (1) in general, Chinese scholars in China have fewer advantages than Chinese scholars overseas on all indicators; (2) Chinese scholars in China have more research potentiality than other scholars due to the faster speed of metrics' growth; and (3) Chinese scientific research in China is more developed and better at science than social science. Additionally, Chinese scholars in China have a slightly stronger impact in most research areas than Chinese scholars overseas.

Keywords: Chinese scholar, research productivity, research impact, scientific communication, usage metrics.

CHINA is facing a great tide of returnees finishing their studies or working abroad and now ready to make a living in China. The large population not only represents the large number of Chinese scholars who studied overseas but also indicates that China has increasing opportunities and beneficial policies that have been attracting students and workers abroad back to China[1]. Chinese scholars overseas have certainly participated in many important scientific projects and produced many significant outcomes[2]. The level of research in China is no longer far behind the global level and is perhaps even higher in some subjects[3]. Chinese scientific research has captured increasing attention in the global academic environment. Names of Chinese scholars have become common on papers in international journals and in awards in conferences both individually and together with scholars in other countries[4].

The two above mentioned scenarios represent the research output of Chinese scholars in China and overseas. We wonder if there is any difference in terms of research productivity and impact of those who moved abroad. In this article, we report an individual-oriented analysis on the differences between Chinese scholars in China and overseas. The differences are displayed by a performance evaluation of Chinese scholars through the following questions: (1) What is the difference between Chinese scholars working in China and overseas in terms of research productivities and impacts? (2) What are the differences between the two groups in different research areas? The potential implications are knowledge of the knowing the current situation of Chinese international communication as well as future policy on scholarly communication.

Related work

Research performance evaluation metrics

Performance evaluation involves evaluation of objects and indicators. Countries, institutions and individual scholars are all evaluation objects. Different indicators are utilized on different objects. Sometimes one indicator is sufficient for evaluation, but generally, a complex indicator system is more convincing[5].

Scholars' productivity is usually measured by the number of papers, patents, presentations and awards. Citation metrics including total citations, citations per paper, *h*-index and other deformations have been overused to measure scholars' impact[6]. Many perspectives have been discussed recently since the absolute quantity on only a few

Wen Lou is in the Faculty of Education, East China Normal University, Room 323, Building of Law and Business, East China Normal University, 500 Dongchuanrd, Minhang, Shanghai 200241, China; Hui Wang is in the Zhejiang Provincial Institute of Communications Planning, Design and Research, 89 Huancheng West Road, Hangzhou, Zhejiang 310006, China and Siluo Yang is in the School of Information Management, Wuhan University, 299 Bayi Road, Wuhan 430072, China.
*For correspondence. (e-mail: 58605025@qq.com)

measurements cannot satisfy in an era of fast-changing data[7]. For this reason, altmetrics was proposed[8]. Later Web of Science (WoS) put forward usage count[9] to adjust the utility of user browsing and saving behaviour instead of using citations only. Indeed, the relationship between the number of citations and downloads has been discussed for a long time[10], and has attracted increasing attention until altmetrics started being utilized on social media.

Performance evaluation of China

Since China entered the realm of international scholar communication, the Chinese have approached China's current developmental situation with interest, and scientists from other countries have also been highly interested in Chinese issues[11]. Previous studies have focused on two aspects. One is the approach to the performance evaluation of China itself[12]. Zhu and Willett[13] conducted a bibliometric analysis of the development of Chinese research in superconductivity. Second, a comparative analysis has been continuously discussed between China and other countries. Most studies have taken China in its entirety to uncover the difference from external objects[14,15]. Fewer studies have focused on internal comparisons, such as Chinese scholars themselves, but within and outside China. On the other hand, because the United States is a multi-national country, the analysis on scholars' performance within America is more popular[16,17].

However, in this study, we draw a comparison only between Chinese scholars in China and overseas, which is rare in previous studies. We believe that the first approach on the performance evaluation of Chinese scholars in China and overseas with updated metrics will elucidate insights on the current and potential distribution of the level of Chinese research. This will help derive some ideas on Chinese scientific decision policies and current education construction projects.

Methodology

Data sources and collection

Journals were selected using the following criteria: (1) choose three journals each from SCI journals and SSCI journals based on 2016 JCR; (2) subjects to which journals belong must be independent, aiming to reduce the bias of duplicate author names and multi-disciplines; (3) Select the top journal from the one and only associated subject. Journals in JCR are usually categorized into two or more subjects. If a journal is the top one in the subject but belongs to other subjects, we skip it until we find the top journal which belongs only to this subject. Therefore, the sampled journals are: *ACM Transactions on Graphics* in Computer Science Software Engineering (ACM), *American Historical Review* in History (HIS), *Biological Reviews* in Biology (BIO), *International Journal of Information Management* in Information Science and Library Science (INFOR), *Journal of Business Venturing* in Business (BIZ), *New England Journal of Medicine* in Medicine, General and Internal (MED).

Articles and article proceedings published between 2005 and 2014 were downloaded from the WoS platform. Reviews in *American Historical Review* and *Biological Reviews* were also downloaded as these two journals are more review-oriented rather than article-oriented. All full-record data were collected on 15 July 2016. We decided on the year 2014 as the end of the targeting period because papers published in 2015 and 2016 may not have been cited yet.

Identify Chinese scholars

First, this study must demonstrate who the Chinese scholars are. Given the difficulty of identifying a scholar's nationality, many studies have considered the authors' affiliations as their nationality for convenience. However, in this article, affiliations will not serve as a substitute for nationality. Hence, we define Chinese scholars in China as individuals whose original nationality is Chinese and who are now working in China. Chinese scholars overseas are individuals whose original nationality is Chinese but who are now working overseas. Here we use scholars' original nationality instead of their current nationality, because many scholars today have changed nationality. Current nationality has become more complicated than original nationality.

Identifying a scholar's nationality is difficult since few people would publicize nationality. However, we believe that educational background information could provide the most accurate information. People tend to go through the first educational stage in their mother countries. We determined the college period as the first educational stage because few people publish information about their kindergarten, primary school or high schools. Besides the college period is the most fundamental stage for a scientific career.

The identification process is as follows: (1) Manually highlight potential Chinese names. People's names apart from English or other Latin languages are very different from names in Chinese Pinyin. Distinguishing Chinese names from Western, Japanese, Korean, and most other Southeastern Asian names is simple. We highlighted names when the last names or first names appeared to be Chinese names, including deformations. For example, 'Wong' is a deformation of 'Wang'. (2) Look for college information from their homepages, CVs, LinkedIn profiles, Research Gate profiles, and other resources to determine if they were educated in China. We were surprised to find other helpful information, such as mother language and high school information, which assured us of their original nationalities. (3) Contact the remaining uncertain

Table 1. Indicators description and formulation in this paper

Indicators	Description
Total papers	The total amount of selected journals in a year or a research area
Author*	The number of Chinese authors who have publications in a year or a research area
Paper*	The number of papers only published by Chinese scholars in a year or a research area
*P**	The percentage of papers in the total papers. *P* = Paper/total papers
*R**	The ratio of papers per author. *R* = Paper/author
*TC**	The number of total citation of a paper on WoS platform
*U*1*	The count of the number of times the full text of a record has been accessed or a record has been saved in the last 180 days on WoS platform
*U*2*	The count of the number of times the full text of a record has been accessed or a record has been saved since 1 February 2013, on WoS platform
*RTC**	The total citation ratio of papers per author. *RTC* = *TC*/author
*RU*1*	The ratio of *U*1 of papers per author. *RU*1 = *U*1/author
*RU*2*	The ratio of *U*2 of papers per author. *RU*2 = *U*2/author

Indicators with a star (*) mean they contain sub-indicators, e.g. *AuthorC* refers to Chinese authors in China. *AuthorO* refers to Chinese authors overseas. *C* is short for China, and *O* is short for overseas.

authors for the information. Only 16 who could not be reached had to be abandoned in the following analysis.

Indicators description

In the following sections, several measurements are discussed (Table 1) for representing productivity and impact.

To calculate co-author papers, we used the normal count approach in which each author is counted as one publication. Author order was not considered. In addition, we determined that people with the same names and different affiliations were different people. However, if the same names were shown in one paper, we considered the same as one person working with different affiliations. In addition, if a person worked both in China and overseas, we counted the paper once both in China and overseas.

Findings

The numbers of authors in China and overseas

In total, we found 1983 Chinese authors overseas and 1190 Chinese authors in China. In different years, we calculated 2270 Chinese authors overseas and 1654 Chinese authors in China. The reason for this inequality is that some authors published more than one paper in ten years.

The first part of Figure 1 shows sequentially comparison. Generally, *AuthorC* and *AuthorO* both increase over the years. *AuthorC* is almost always less than *AuthorO*, except in 2013, because in 2013, there was a sudden increase of Chinese authors in China in computer science and medicine. Additionally, the increasing and decreasing tendencies of *AuthorC* and *AuthorO* are similar, especially from 2006 to 2011. However, in 2012 and 2013, a dramatic reversal occurs.

The second part of Figure 1 shows the comparison between different research areas. *AuthorO* is larger than *AuthorC* except in computer science. In information science, business, and medicine areas, *AuthorO* approximately doubles *AuthorC*, which implies that more Chinese overseas are good in these research areas than Chinese scholars in China. However, scholars in China are better in computer science since Chinese scholars overseas are 147 less than the number in China. We also found few Chinese scholars in history regardless of group. This finding is reasonable because the journal (*American History Review*) mainly focuses on American history, which leads to few other nationalities in this study area.

Chinese authors in China and overseas productivity performance

Six journals published 6306 papers in 10 years (shown in the first part of Figure 2). The number of publications slightly increases yearly after 2006. Chinese authors overseas published 1319 papers in 10 years. Meanwhile, Chinese authors in China published only 400 papers. P_c and P_o also show the huge gap between the two groups. P_o mainly stays at the same level (20–25%). Therefore, the trend of P_c will potentially keep growing, which means that Chinese scholars in China will publish an increasing number of papers in the future. The total number of papers in all research areas aligns with the total number of authors in Figure 1. The only exception is that P_c, is lower than P_o in computer science, which is opposite to the number of authors in computer science.

Apparently, in Figure 3 the gap between within and outside China is large as well. Even though both *PaperC* and *PaperO* increased over ten years, the absolute value of *PaperO* is at least twice *PaperC*. However, *PaperC* increased six times, which was much higher than the growth of *PaperO* (60%). The situation is similar to the comparison between P_c and P_o. Therefore, Chinese

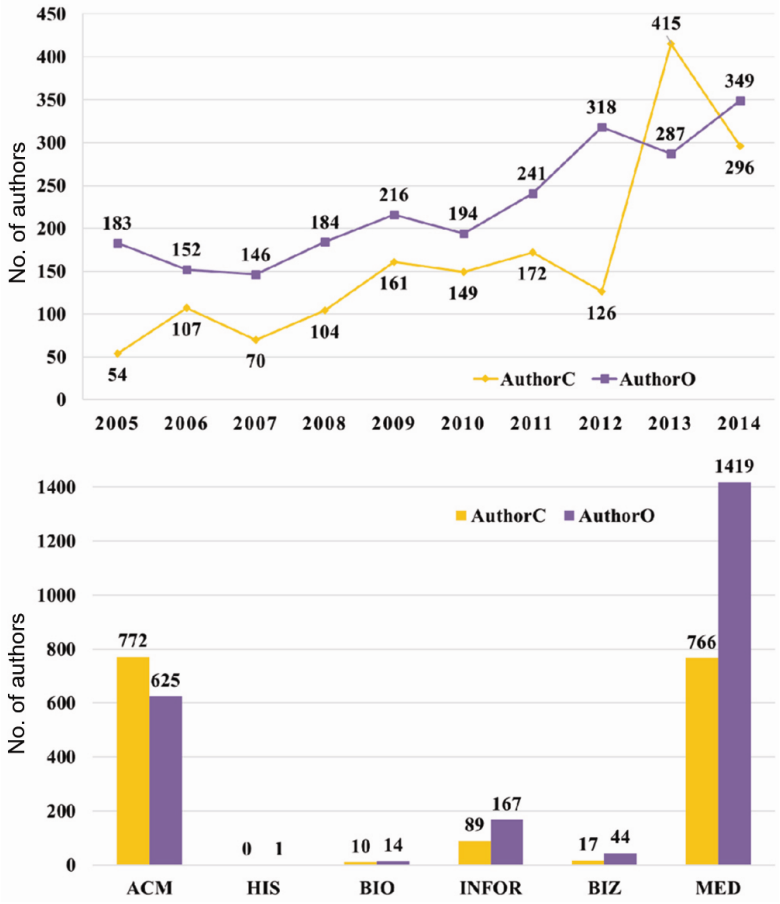

Figure 1. Comparison of the numbers of Chinese authors.

authors in China may publish more papers in the future. R_o has dropped slightly in ten years (ranges from 0.7 to 0.5), which means that two scholars co-author a paper every year. However, R_c has slightly increased in ten years but ranges from 0.2 to 0.3, which means that three or four scholars co-author approximately one paper every year. The comparison shows that Chinese authors in China publish fewer papers per person, while the degree of co-authoring is higher for Chinese authors in China than for Chinese authors overseas.

In different research areas (second part of Figure 3), Chinese scholars in China still published papers fewer than Chinese scholars overseas. For instance, even though there were more Chinese authors in China in computer science (Figure 1), their productivity in ACM was less than that of the group overseas. In addition, in medicine, double the number of Chinese authors overseas published seven times the number of papers as Chinese authors in China.

Chinese authors in China and overseas impact analysis

Total citations (*TC*) tend to go down over ten years both within and outside China, while *U1* and *U2* tend to rise. This finding is because citations have a time lag, and the newer the paper, the fewer citations it may have. However, usage count (especially *U2*) is a cumulative amount, which always increases.

The difference between the absolute values of indicators within and outside China is a ten-digit, regardless of the total numbers or the means. Coefficient of variation (CV) is useful in this case. CVs from within China are twice those from outside China. Therefore, the impact of papers published by Chinese authors in China is more unpredictable, and the impact of papers published outside China is stable. Taking *TC* as an example, TC_o has a four- or five-year cycle, while TC_c continues to rise at the beginning, but recently has a two-year cycle. We explain that papers published by Chinese authors in China are active.

TC has gone through a dramatic route over ten years. Both RTC_c and RTC_o have tended to decrease yearly with only slight internal growth. RTC_o decreases greatly from more than 200 to less than 50. RTC_c basically remains under 50. Therefore, there is a large overall gap in Chinese author impact between China and overseas. However, the gap degree varies in different research areas. The largest gap of *RTC* between the two groups is in medicine followed by biology. However, in business,

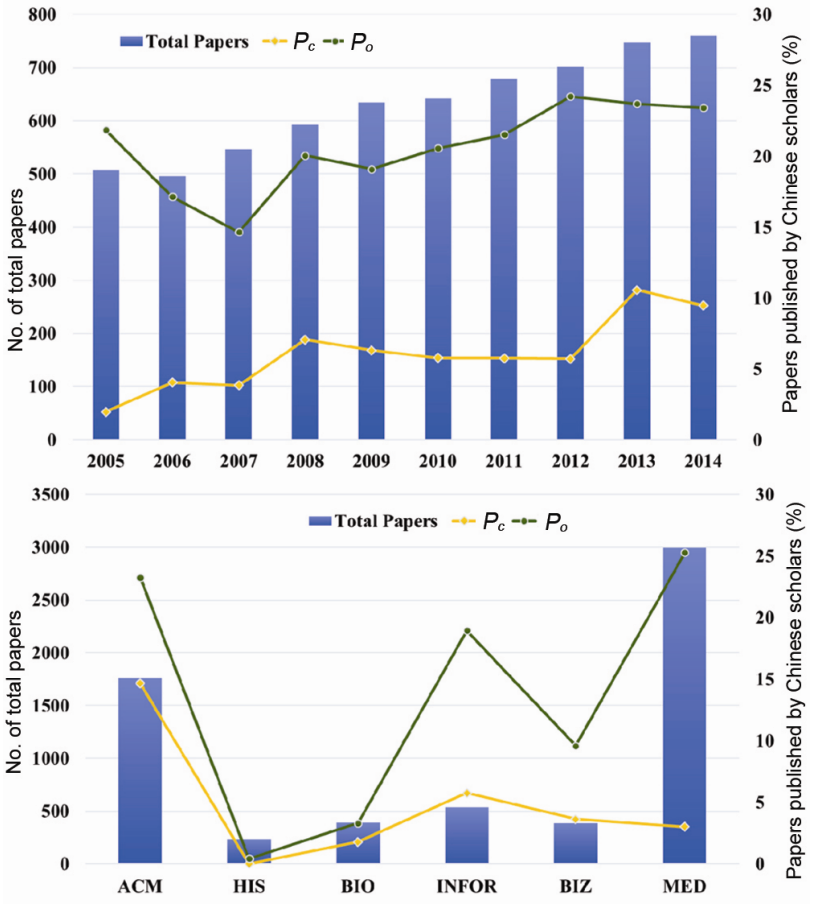

Figure 2. Comparison of total papers and *P*.

RTC_c is larger than RTC_o. We could say that Chinese authors overseas have more impact than Chinese authors in China in biology and medicine compared to other areas.

The usage count numbers in China are not comparable to those overseas. Both $RU1_c$ and $RU2_c$ are always below $RU1_o$ and $RU2_o$, and the latter are twice as many as the former almost every year. Similar to the other *R**, *RU1* and *RU2* in the two groups also changed inversely over the years. Therefore, scholar exchange not only affects publishing performance but also impacts results.

Conclusion and discussion

Our results show the big difference between Chinese scholars in China and overseas on both research productivity and impact. (1) One of the differences between the two groups is based on static data. Chinese scholars in China have fewer advantages than Chinese scholars overseas on all indicators, namely, the absolute and proportional values of authors, publications, citations and usage counts. In fact, these four aspects of indicators represent research power, output capability[18], formal impact, and informal impact individually. Therefore, Chinese scholars in China lack these abilities comparatively. Chinese papers in total share a small proportion of the world's paper in this study. Only when Chinese scholars can publish papers in top journals, can they reach the top level of research. (2) The second difference between the two groups is the difference in tendency. Chinese scholars in China have more potential development than overseas. Regardless of which research area or which indicator is used, the metrics of Chinese scholars in China have a faster speed of development.

As for the second research question, the differences vary between the two groups in different research areas as follows. (1) Regardless of the shortage of Chinese scholars in China generally, they are good at some indicators in some research areas. For example, in computer science, the number of Chinese scholars in China is larger than Chinese scholars overseas. In other words, the research power of Chinese scholars in China in computer science is stronger than the other group. (2) For the output capability represented by the number of publications, the gaps between two groups vary for different research areas. Larger gaps exist in computer science and medical science. (3) From the perspective of impact, on an average, papers written by Chinese scholars in China

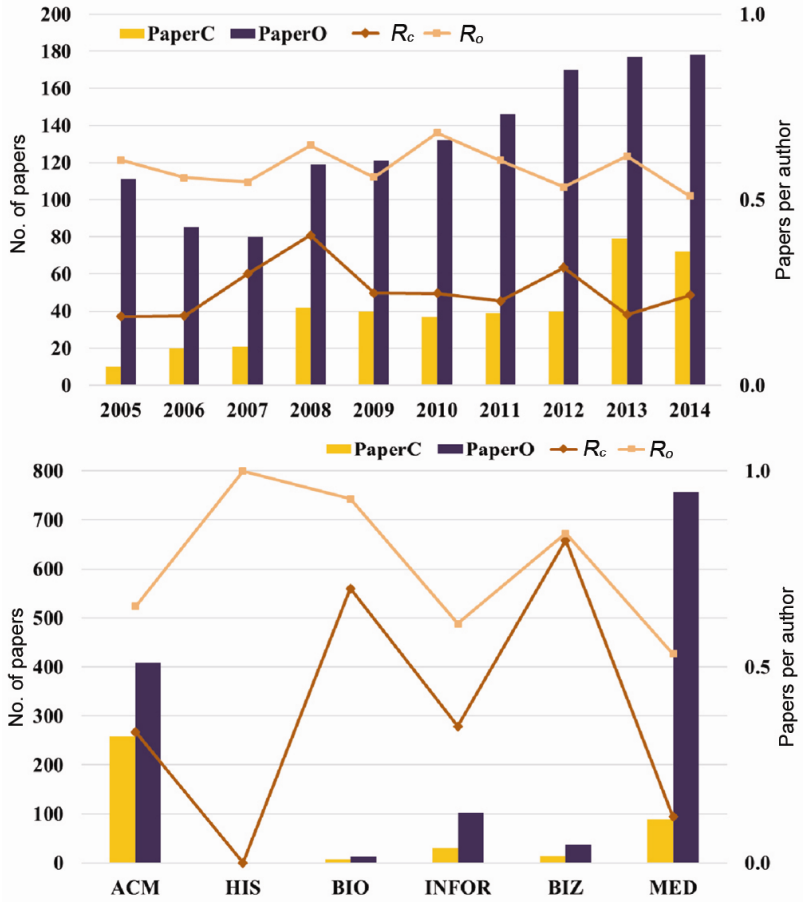

Figure 3. Comparison of paper and *R*.

Table 2. Statistics of *TC*/*U*1/*U*2 over time

PY	TC_o	TC_c	$U1_o$	$U1_c$	$U2_o$	$U2_c$
2005	38,978	2589	297	27	2082	200
2006	33,798	2410	258	24	1832	218
2007	24,812	3563	194	59	1311	326
2008	26,259	7745	278	76	1692	722
2009	30,551	9070	539	108	2698	901
2010	30,225	3254	384	114	2874	769
2011	29,060	4279	514	107	3713	920
2012	26,156	1725	563	82	4572	775
2013	17,573	4695	614	320	5399	2469
2014	12,073	1767	563	211	3240	1142
Total	269,485	41097	4204	1128	29413	8442
SD	7707.464	2485.488	154.718	90.110	1318.059	653.083
Mean	26,948.500	4109.700	420.400	112.800	2941.300	844.200
CV	0.286	0.605	0.368	0.799	0.448	0.774

have a slightly stronger impact in computer science, information science, library science, business, and medicine than papers written by Chinese scholars overseas. Meanwhile, in biology, Chinese scholars overseas have more impact both generally and on an average.

Using only one journal as a data source representing one research area is controversial. However, despite the small scale of publication, one representative for a research area is a common way to conduct informetrics analysis. We used a primary approach to determine the differences between Chinese scholars in China and overseas. We should work on more journals and other metrics in the future to improve this study. Two further considerations are demonstrated below.

Chinese scientific research is certainly progressing, but there is much room for improvement in the future. Therefore, studying or working abroad for Chinese scholars is necessary. In fact, our data on Chinese scholars in China included scholars in Hong Kong and Taiwan, which enlarged the data on Chinese scholars in China. If the data only included papers from mainland China, the gap would be even bigger, as we could expect.

Studying or working outside China requires different policies in different circumstances. We have seen fewer works by scholars in social science than in science. Under the double world-class project, universities with a strong background in science could improve science even more if they sent out scholars to further their world-class disciplines. Universities with a comprehensive discipline background could improve social science if they expanded social science studies outside China, so that they could build world-class universities. For instance, American history has been scarcely discussed by Chinese scholars. We did not remove it from our data because it may remind Chinese scholars of a research blank.

1. Hao, J., Wen, W. and Welch, A., When sojourners return: Employment opportunities and challenges facing high-skilled Chinese returnees. *Asian Pac. Migr. J.*, 2016, **25**, 22–40.
2. Zou, Y. and Laubichler, M. D., Measuring the contributions of Chinese scholars to the research field of systems biology from 2005 to 2013. *Scientometrics*, 2017, **110**, 1615–1631.
3. Zong, F. and Wang, L., Evaluation of university scientific research ability based on the output of sci-tech papers: a D-AHP approach. *PLoS ONE*, 2017, **12**, e0171437.
4. Yang, S. and Wang, F., Visualizing information science: author direct citation analysis in China and around the world. *J. Informetr.*, 2015, **9**, 208–225.
5. Radicchi, F. and Castellano, C., Analysis of bibliometric indicators for individual scholars in a large data set. *Scientometrics*, 2013, **97**, 627–637.
6. Radicchi, F. and Castellano, C., A reverse engineering approach to the suppression of citation biases reveals universal properties of citation distributions. *PLoS ONE*, 2012, **7**, e33833.
7. Bornmann, L., Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *J. Informetr.*, 2014, **8**, 895–903.
8. Priem, J., Taraborelli, D., Groth, P. and Neylon, C., Alt-metrics: a manifesto. 2010, pp. 1–5; doi:10.1371/journal.pmed.0030291.
9. Wang, X., Fang, Z. and Sun, X., Usage patterns of scholarly articles on Web of Science: a study on Web of Science usage count. *Scientometrics*, 2016, **109**, 917–926.
10. Moed, H. F., Statistical relationships between downloads and citations at the level of individual documents within a single journal. *J. Am. Soc. Inf. Sci. Technol.*, 2005, **56**, 1088–1097.
11. Zhou, P. and Leydesdorff, L., The emergence of China as a leading nation in science. *Res. Policy*, 2006, **35**, 83–104.
12. Liu, W., Hu, G., Tang, L. and Wang, Y., China's global growth in social science research: uncovering evidence from bibliometric analyses of SSCI publications (1978–2013). *J. Informetr.*, 2015, **9**, 555–569.
13. Zhu, Q. L. and Willett, P., Bibliometric analysis of Chinese superconductivity research, 1986–2007. *Aslib Proc.*, 2011, **63**, 101–119.
14. Wong, C. Y. and Goh, K. L., The sustainability of functionality development of science and technology: Papers and patents of emerging economies. *J. Informetr.*, 2012, **6**, 55–65.
15. Larsen, P. O., Maye, I. and Ins, M. von, Scientific iutput and impact: relative positions of China, Europe, India, Japan and the USA. *Collnet J. Scientometer. Inf. Manage.*, 2008, **2**, 1–10.
16. Walsh, J. P., The impact of foreign-born scientists and engineers on American nanoscience research. *Sci. Public Policy*, 2015, **42**, 107–120.
17. Lee, S., Foreign-born scientists in the United States-do they perform differently than native-born scientists? Ph D thesis, Georgia Institute of Technology, 2004, vol. 36, pp. 27–30.
18. Veugelers, R., Towards a multipolar science world: trends and impact. *Scientometrics*, 2010, **82**, 439–456.

ACKNOWLEDGEMENTS. We thank the support from Professor Star Xing Zhao, the youth project of National Social Science Foundation of China (grant number 17CTQ025), and the National Social Science Foundation of China (grant number 16BTQ055).

Received 28 November 2017; revised accepted 9 February 2018

doi: 10.18520/cs/v115/i1/49-55