*Jointly published by Akadémiai Kiadó, Budapest*
*and Kluwer Academic Publishers, Dordrecht*

*Scientometrics,*
*Vol. 54, No. 1 (2002) 145–154*

**Short Communication**

# Development of the *Chinese Scientometric Indicators* (CSI)

BIHUI JIN, JIANGONG ZHANG, DINGQUAN CHEN, XIANYOU ZHU

*Documentation and Information Center of the Chinese Academy of Sciences (DICCAS)*
*Zhongguancun, Beijing (P.R. China)*

We describe the *Chinese Scientometric Indicators* (CSI), an indicator database derived from the *Chinese Science Citation Database* (CSCD). Its design is supported by the Natural Sciences Foundation of China (NSFC). In this indicator database data of a statistical nature are organized and categorized leading to ranked lists and providing bases for comparisons among Chinese institutions and regions.

## Introduction

Science and technology are the driving forces of China's rejuvenation and sustainable development (*Jiang*, 2000; *Cyranovski*, 2001). As in most countries, China's leaders, its policy makers, funding agencies and university administrative units pay close attention to research evaluation. Having a database or, if necessary, several databases reflecting all scholarly achievements is indispensable for this purpose. Moreover, a bibliographic database alone does not suffice. A scientometric database providing research-related input and output indicators is also necessary. Therefore, after setting up the *Chinese Science Citation Database* (CSCD) (*Jin* and *Wang*, 1999; 2000), the Documentation and Information Center of the Chinese Academy of Sciences (DICCAS) (*Meng* and *Wang*, 1996) designed a database of indicators meeting China's needs and conditions for research evaluation.

The Library of the Hungarian Academy of Sciences, the Centre for Science and Technology Studies (CWTS) at Leiden University (the Netherlands) and SPRU (UK), all provide examples of good practice and plenty of experiments in the field (*Braun* et al., 1985; *Moed* et al., 1985; *Skea* et al., 1992). They served as sources of inspiration in order to determine which indicators to consider. The actual indicators were designed based on an in-depth exploration of the data, record and field structure of the CSCD, the

*0138–9130/2002/US $ 15.00*
*Copyright © 2002 Akadémiai Kiadó, Budapest*
*All rights reserved*

SCI and ISI's *Journal Citation Reports* (JCR). The resulting database, the *Chinese Scientometric Indicators* (CSI), is available in print as well as on CD-ROM. It is the tool *par excellence* for probing into the genesis of scientific developments, evaluating the efficiency of scientific research and conducting comparative studies according to internationally accepted standards in China. In this indicator database data of a statistical nature are organized and categorized leading to ranked lists and providing bases for comparisons among Chinese institutions.

## China's conditions for research and research evaluation

### *Strong demands on scientific research and limited resources*

China is a vast developing country with a large population and comparatively low resources for scientific research. There were in total more than 22,000 S&T research bodies, and about 1.5 million scientists and engineers in China, in the year 1998. The R&D expenditure, however, was only 55.11 billion yuan (about 6.7 billion US dollar), accounting for 0.69% of China's GDP (*National Bureau of Statistics & Ministry of Science and Technology*, 2000, pp. 2–5, 230).

One of China's development strategies for science and technology implemented by the government is that limited resources are focused on strategic fields. Consequently, policy decisions are to be made in order to determine these strategic fields, and further, evaluation is required in order to optimally allocate (or re-allocate) provided resources.

### *Language preference*

Most of the S&T papers written by Chinese scientists are published in Chinese journals. Only a small part is published in international journals. Due to this linguistic preference China's academic exchange is only partially internationalised (*Moed*, 2000). There exist indeed two circles for academic exchange: a domestic and an international one. The Chinese language dominates in the first circle, English in the second one. Therefore, it is necessary that domestic databases, such as the CSCD, and international ones (e.g. ISI's) be combined in order to obtain adequate indicators.

*Introducing quantitative evaluation into research management*

About twenty years ago China did not perform any quantitative evaluation of research management. Since the reform of the S&T system, however, quantitative evaluation has been introduced into research management and decision-making related to S&T issues. In recent times the attitude with respect to quantitative methods has greatly changed. Administrators responsible for funding, journal editors, scientists, engineers and every one involved in the scientific enterprise are all coming to terms with bibliometric and scientometric methods. From a mild interest, quantitative methods have turned into an almost daily practice.

Thus it can easily be understood that the design of the *Chinese Scientometric Indicators* (CSI) was indispensable for the evaluation and the analysis of Chinese science and technology.

## Selection of statistic sources

Taking China's special situation into account, it is clear that for the quantitative analysis of the country's S&T system, neither a foreign database nor a local one can suffice. Therefore we have used information obtained from the SCI, in so far it pertains to papers written by Chinese scientists in ISI-covered journals, and all information from the CSCD, following the principle that statistical sources must be chosen in accordance with the objective of the evaluation exercise.

*Significance of the SCI as a statistic source*

Most of the several thousands of journals covered by the SCI can been considered as core international journals. The SCI covers about 15,000 publications by Chinese scientists. Of these 15,000 about 13,000 (or roughly 85%) are published in non-Chinese journals. Those publications are few compared with the total number of Chinese publications but they represent the most important research results of China. In this sense, they constitute an important window through which outside people can understand and follow the scientific and technological development of China. Together they also give rise to important parameters indicating the level of internationalisation of fields or institutions in China.

*Significance of the CSCD as a statistical source*

Despite the fact that the SCI is an important tool for science evaluation, including China's research performance (*Moed*, 2000), Chinese journals still are the main channel for the publication of research results obtained by Chinese scientists. For example, the number of Chinese S&T journals totaled 4267 titles in 1998 (*Association of Chinese Publishers*, 1999, p.74). SCI-Expended covered 31 Chinese journals, accounting for only 0.73% of the total. The number of Chinese articles covered by the SCI accounts for 4.5% of the total production in China (about 0.44 million a year). So, it is absolutely necessary to include local journals when reviewing research performance in China.

The CSCD was founded in 1989. It is the first non-English science citation database in the world (meanwhile China has a second science citation database, the CSTPC, and also social science citation databases) (*Liang*, 2001). Because of the language preference in China and the small number of Chinese journals covered by the SCI, the CSCD is more than just a local supplement to the SCI. It provides the way for a quantitative evaluation of scientific research in China, and helps designing Chinese scientometric indicators representing the real situation of scientific research in China.

When constructing the CSCD, 582 Chinese journals were selected as source journals. This selection was based on the fact that these journals were the best in their fields, emphasizing basic academic research on the frontier of science. In 1999 the set of source journals was enlarged to a total of 1064. Statistical data show that first-rate universities and research institutes publish their best results either abroad (in ISI-covered journals) or in the source journals of the CSCD. This means that the CSCD is a representative source for the analysis of Chinese science and technology, as well as for the design of the *Chinese Scientometric Indicators* (CSI).

According to our philosophy that both local and international core journals are needed for effectively assessing China's publications, we combined data from the SCI and the CSCD in the construction of the CSI.

## Profile of the CSI

The CSI is the prime set of indicators for a bibliometric and scientometric performance evaluation of China's research. The Natural Sciences Foundation of China (NSFC) supports its design. Its inception was spurred by two conditions: the availability of data (the SCI and the CSCD) and a great social (and scientific) demand.

*General description*

The CSI consists of eight subsets of indicators. These will be described shortly. The first volume contained 121 indicators. It was published, in a printed version, in 1998 and described the situation for 1997. There are 149 indicators in the 1999 volume and the 2000 volume (to be published in 2001) uses 191 indicators. Since 1999 the CSI also exists in a CD-ROM version. Meanwhile, a web-based version has been planned.

*Brief introduction to the eight subsets of indicators*

*Statistical sources.* This section contains statistics on the subject distribution of Chinese articles in SCI and CSCD journals, and on the distribution of SCI source journals by country and region. Data are given on the ranking zones of journals containing articles by Chinese scientists. This section further contains the publication distribution of articles in different subjects (for the SCI and the CSCD) and the citation distribution in the CSCD of articles in different subjects. This subset of indicators is the basis for all other ones.

*Institutional indicators.* The number of papers and citations were obtained for universities and colleges, research institutes, and medical institutions. Following a Bradford-like approach, the SCI and the CSCD journals were subdivided respectively into four and three zones. A unified method for assessing international (SCI) and domestic (CSCD) articles has been proposed in order to determine the publication productivity of an institution (*Jin* et al., 1999).

Statistics on the core institutions with respect to publication (number of articles) and impact (number of citations) were obtained, observing (roughly) an 80–20 rule. In practice a cut, segregating core from non-core, was applied not at 80% of all ‘productions’ (articles or citations), but at 70%. Table 1 shows the basic results for 1998.

Table 1. Productivity of Chinese institutions (1998)

| |
|---|
| ***Total number of institutions producing articles: 6329*** |
| ***Total number of articles published: 80635*** |
| ***Average number of articles per institution: 12.7*** |
| |
| ***Number of core institutions: 218*** |
| ***Number of articles produced by core institutions: 57271*** |
| ***Average number of articles per core institution: 262.7*** |

The group of core institutions consists of 3.44% of all institutions, but publishes 71.02% of all articles.

*State key laboratory indicators.* There are 157 state key laboratories and more than one hundred open laboratories in China. The Chinese government gives these laboratories special support. In return these labs must accept government evaluation. Responding to this situation, the CSI provides a set of indicators for state key laboratories and open laboratories. The number of articles of these labs, number of citations, number of authors, subject distribution of articles, and the number of articles supported by the NSFC are counted and tabulated.

*Regional indicators.* The People's Republic of China is subdivided into 31 provinces and autonomous regions. There exists an unbalanced development amongst regions. This is true in the areas of economy and culture as well as S&T (*Tsui*, 1996; *Jin* and *Rousseau*, 2000). In the CSI statistics on the number of articles and subject distribution for each province and autonomous region can be found. The number of core institutions of each province and autonomous region are presented based on the data from the institution indicator subset. These data can be used to analyze the disciplinary advantages and weaknesses of each province and autonomous region.

*S&T funds indicators.* Setting up a system for the funding of science is one of the important measures taken in the research reform in China. A funding system cascading from the central government to local governments was set in place. In this system the NSFC plays a major role for the funding of basic research.

Corresponding to its role in research evaluation the CSI also procures statistics on the number of articles published with the help of these funds and gives a ranking of colleges and research institutions with respect to the number of articles supported by the NSFC.

Statistics on the average citation frequency of articles supported by the NSFC and the number of those not supported by NSFC show that the former have a larger impact than the latter. In 1999, the articles supported by NSFC received 0.25 citations on average with 5 years window, while for those not supported by NSFC the number is 0.13. This is illustrated in Table 2.

Table 2. Average citation frequency of NSFC-funded articles versus non-NSFC-funded ones

| | A. Number of articles | | B. Citations | | C. Impact = B/A | |
|---|---|---|---|---|---|---|
| Year | 93–97 | 94–98 | 93–97 | 94–98 | 1998 | 1999 |
| NSFC-funded | 50,225 | 59,169 | 9,056 | 14,792 | 0.18 | 0.25 |
| Not NSFC-funded | 152,645 | 171,094 | 15,532 | 22,390 | 0.10 | 0.13 |

*Indicators of research collaboration.* Collaborative research is a general trend in modern science. The number of articles written in collaboration with other nations, those resulting from collaboration among provinces and autonomous regions and, finally, articles written in a collaborative effort among universities, institutes and enterprises have all been counted and tabulated.

Statistics on the co-authorship degree for articles of different disciplines show that this index in 1999 is the highest in medical science, namely 4.15, and the lowest in mathematics (1.86). At same period our observation also show in general that the articles of Chinese authors covered by SCI have a larger co-authorship degree than the articles covered by CSCD.

International collaboration is of the utmost importance for the development of Chinese S&T. Our observations show that most articles written in international collaboration are published in journals covered by the SCI. In 1999 there were 4101 such articles in the SCI and 2362 in the CSCD.

International collaboration between China and other countries focuses on a small number of preferred nations: USA, Japan, Germany, UK, France, Canada, Australia, Italy, South Korea, and the Netherlands. Collaboration with these ten countries led to 3456 articles (in 1999), accounting for 84.27% of all collaborative articles (covered by the SCI)**.** Data show a trend of even more concentration. Table 3 provides more detail about the number of collaborations with these countries. Within Asia China collaborates most with Japan, South Korea, Singapore, and to a lesser extent India (*Arunachalam* and *Doss*, 2000).

Table 3. Distribution of collaborative papers

| | A. China as first author | | B. China as co-author | | C. All Chinese collaborations (A+B) | |
|---|---|---|---|---|---|---|
| Country | 1998 | 1999 | 1998 | 1999 | 1998 | 1999 |
| USA | 505 | 521 | 858 | 752 | 1363 | 1273 |
| Japan | 321 | 413 | 308 | 371 | 629 | 784 |
| Germany | 211 | 259 | 283 | 239 | 494 | 498 |
| UK | 150 | 171 | 219 | 183 | 369 | 354 |
| France | 82 | 100 | 171 | 148 | 253 | 248 |
| Canada | 85 | 77 | 122 | 98 | 207 | 175 |
| Australia | 91 | 84 | 102 | 106 | 193 | 190 |
| Italy | 42 | 41 | 146 | 145 | 188 | 186 |
| Korea | 51 | 58 | 103 | 90 | 154 | 148 |
| The Netherlands | 45 | 52 | 67 | 89 | 112 | 141 |
| Total | 1583 | 1776 | 2379 | 2221 | 3962 | 3997 |

*Author indicators.* The distribution of gender, age, academic degree (MSc. or Ph.D.) for Chinese authors based on data obtained from the original publications (in about 50%

of the cases this information is provided in Chinese journals) are collected. These data are useful for analyzing and evaluating the sociological structure of Chinese S&T personnel.

The statistics with respect to the age of authors for the years 1998 and 1999 show that the mode of the age curve of authors is situated in the 31–35 range. Altogether 61039 articles fall in this range, accounting for 27.82% of the total. In this same period, authors between 20 and 40 published 68.58% of all articles. These data show that China's active scientists are young (see also Figure 1 and Table 4). The dip for the ages between 46 and 50 is probably due to the effects of the Cultural Revolution.

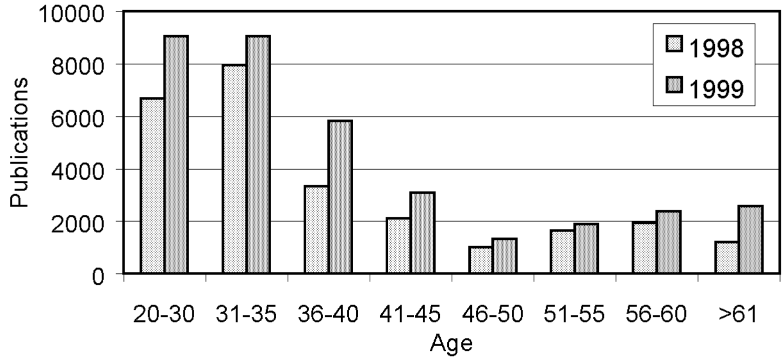


Figure 1. Age distribution of Chinese scientists publishing in a part of journals covered by the CSCD

Table 4. Age distribution of Chinese authors

| | 20–25 | 26–30 | 31–35 | 36–40 | 41–45 | 46–50 | 51–55 | 56–60 | >60 |
|---|---|---|---|---|---|---|---|---|---|
| 1998 | 6681 | 7941 | 3325 | 2105 | 1007 | 1642 | 1951 | 1197 | 811 |
| 1999 | 9051 | 9043 | 5821 | 3090 | 1341 | 1905 | 2369 | 2563 | 1021 |
| 1998 (%) | 25.85 | 30.72 | 12.86 | 8.14 | 3.90 | 6.35 | 7.55 | 4.63 | 4.16 |
| 1999 (%) | 25.73 | 25.70 | 16.54 | 8.78 | 3.81 | 5.41 | 6.73 | 7.28 | 4.04 |

*Literature indicators.* Citation analysis is an important tool in journal evaluation. For a worldwide assessment of journals the JCR contain the best (or most often used) statistics. Similar statistics on the cited frequency of Chinese journals and their impact factors have been obtained from the CSCD.

## Conclusion

We described the CSCD and its derived product: the *Chinese Scientometric Indicators.* The CSCD and CSI were constructed and designed in order to provide data and share resources with researchers in the fields of scientometrics and bibliometrics, and to provide a useful tool and a set of reference data for research managers and decision-makers. The CSCD and CSI, however, can only be used for the evaluation and for scientometric studies of local, i.e., Chinese research. We hope to collaborate with scientists who are interested in international comparisons among countries to review the international position of Chinese science and technology. Though both products play already an important reference role in the effective evaluation of domestic S&T activities, we are convinced that their role, now and in the future, will still increase.

*

We thank Ronald *Rousseau* (KHBO, Belgium) for editorial help during the preparation of this article. This work was supported by NSF China, Grant 79870014.

## References

ARUNACHALAM, S., DOSS, M. J. (2000), Mapping international collaboration in science in Asia through coauthorship analysis, *Current Science*, 79 (5) : 621–628.

Association of Chinese Publishers (Ed.) (1999), *China Publication Yearbook 1999*. Beijing: Chinese Publication Yearbook Press.

BRAUN, T., GLÄNZEL, W., SCHUBERT, A. (1985), *Scientometric Indicators: A 32-Country Comparative Evaluation of Publishing Performance and Citation Impact*. Singapore: World Scientific.

CYRANOVSKY, D. (2001), A great leap forward, *Nature*, 410 : 10–12.

JIANG, Z. (2000), Science in China, *Science*, 288 : 2317.

JIN, B., ROUSSEAU, R. (2000), China's mean center of publication. Manuscript submitted for publication.

JIN, B., WANG, B. (1999), Chinese Science Citation Database: its construction and application, *Scientometrics*, 45 : 325–332.

JIN, B., WANG, B. (2000), The development and application of the CSCD, *Chinese Journal of Scientific and Technological Periodicals*, 11 (1) : 14–16.

JIN, B., WANG, S., WANG, B., ROUSSEAU, R., WU, Z., LIU, X., ZHU, X. (1999), A unified method of counting international and domestic articles, *Journal of Management Sciences in China*, 2 (3) : 59–65.

LIANG, L., WU, Y., LI, J. (2001), Selection of databases, indicators and models for evaluating academic performance of Chinese universities, *Research Evaluation,* to be published.

MENG, G., WANG, B. (1996). The library and information system of the Chinese Academy of Sciences, *Libri*, 46 : 52–58.

MOED, H. F. (2000), Bibliometric assessment of research performance based on the Science Citation Index: opportunities for China. To be published in: *Proceedings of the Second International Symposium on Quantitative Evaluation of Research Performance*, Shanghai, China, 23–25 October 2000.

MOED, H. F., BURGER, J. M., FRANKFORT, J. G., VAN RAAN, A. F. J. (1985), The use of bibliometric data for the measurement of university research performance, *Research Policy*, 14 : 131–149.

National Bureau of Statistics and Ministry of Science and Technology (Ed.) (2000), *China Statistical Yearbook on Science and Technology 2000*, Beijing: China Statistics Press.

SKEA, J. E. F., MARTIN, B. R., LING, N. (1992), Assessing university departments: some problems and partial solutions. In: *Science and Technology in a Policy Context* (VAN RAAN et al., Eds). Leiden: DSWO Press, pp. 106–120.

TSUI, K.-Y. (1996), Economic reform and interprovincial inequalities in China. *Journal of Development Economics*, 50 : 353–368.

Received October 12, 2001.

*Address for correspondence:*
BIHUI JIN
Documentation and Information Center
of the Chinese Academy of Sciences (DICCAS)
8 Kexueyuan Nanlu
Zhongguancun, Beijing, 100080, P.R.China
E-mail: csci@mail.las.ac.cn