China Scientific and Technical Papers and Citations (CSTPC): History, Impact and Outlook

Wu Yishan, Pan Yuntao, Zhang Yuhua, Ma Zheng, Pang Jingan, Guo Hong, Xu Bo, Yang Zhiqing
Institute of Scientific and Technical Information of China (ISTIC)
15 Fuxinglu, Beijing 100038, China
wuyishan@istic.ac.cn

Abstract
This paper traces the history of China Scientific and Technical Papers and Citations database (CSTPC) since its founding in 1988. The fact that most Chinese scientists publish their research results in Chinese journals requires that China establish SCI counterparts dedicated to domestic S & T journals. The article describes the selection criteria for source journals, the approach used to adjust the structure of source journals, the criteria for selecting items to be included in the database, and the indexing method. Then it discusses the impact upon government R & D administration agencies and the science community in general by both CSTPC team and CSTPC database. Finally, the article analyzes the main factors that lead to the primary success of CSTPD. The authors encourages information workers in other non-English developing countries to build up similar databases.

Keywords: China, database, journal, impact

1. Introduction

SCI has been used as an evaluation tool in many countries and China is no exception. However, although China takes about one fifth of the world population, and China's full-time equivalent R & D researchers are as many as 925,000 man-year in 2000 [1], which ranks it as the third place in the world (with USA and Russia boasting more R & D personnel), China's share in world total of SCI papers has never passed 5%. Obviously, in order to appropriately evaluate research performance of Chinese scientists, we have to build our own SCI equivalent. Institute of Scientific and Technical Information of China (referred to as ISTIC henceforth), an institute under the Ministry of Science and Technology of China, has done exactly that. In this paper, we will reflect the history of China Scientific and Technical Papers and Citations Database (referred to as CSTPC henceforth), analyze its impact upon the publishing and evaluation practice of the Chinese scientific community, and give an outlook about its future development.

1.1 Historical Background and Description of CSTPC

As late as 1987, few Chinese knew how many papers were published by Chinese scientists in the world and no one knew how many papers were published domestically. In July of 1987, Mr. Son Jian, the then Minister of the State Science and Technology Commission, which was renamed as the Ministry of Science and Technology in later years, directed the Institute of Scientific and Technical Information of China (ISTIC) to carry out such a paper count. As a result, China Scientific and Technical Papers and Citations (CSTPC), a database dedicated to the partial

Yishan Wu, Yuntao Pan, Yuhua Zhang, Zheng Ma, Jingan Pang, Hong Guo, Bo Xu, Zhiqing Yang, *China Scientific and Technical Papers and Citations (CSTPC): history, impact and outlook*, In: Guohua Jiang, Ronald Rousseau, Yishan Wu (Eds), Proceedings of the 9th International Conference on Scientometrics and Informetrics – ISSI 2003, Dalian University of Technology, Dalian, China, 2003, 352-361.

evaluation of research performance of China's scientists and engineers, initiated by ISTIC and sponsored by the Ministry of Science and Technology (then the State Science and Technology Commission) was born. It is based on representative domestic S & T journals. In 1988, CSTPC covered 1,189 journals while in 2000, the source journals totaled 1,411.

1.2 How we selected the Source Journals?

An ISTIC research group was established in 1987. The group decided that any candidate source journal for our database has to meet four basic requirements: numerability, analyzability, availability and continuity. What is more, considering the general biased distribution of articles and citations among all the journals, as was found by Dr. Eugene Garfield, the group thought that they did not need cover all journals.

By the end of 1988, there were totally 3052 S & T journals as registered in the State General Administration for Press and Publication. In order to select the most representative ones as the source journal of CSTPC, ISTIC did the following work:

--identify all the journals already registered by internationally influenced abstract and index services, such as SCI, EI (Engineering Index), CA (Chemical Abstracts), SA (Science Abstracts, UK), РЖ (produced in the former Soviet Union), and the Current Bibliography on Science and Technology (produced in Japan);

--collect the data on core Chinese journals as analyzed by some Chinese information experts, such as 104 journal titles recognized as core journals in natural sciences by Mr. Jing Qinsu of Lanzhou University, who built up a prototype of Chinese equivalent of SCI but soon gave up [2];

--collect all the major science abstracts, such as "China Physics Abstracts", "China Mathematics Abstracts", and "China Abstracts on Mechanical Engineering, and see what journals were used by them;

--identify all the journals produced by national science societies;

--ask experts of different disciplines to identify most important journals in their specific fields.

Based on the above work, 1189 journals were chosen as the first batch of source journals, in which the shares occupied by basic science, engineering and technology, medicine, as well as agriculture are close to the corresponding shares in the total 3052 journals. See Table 1 and Table 2.

Table 1 Distribution of Chinese S & T Journals in Terms of Subjects, 1988

Subject category	Journal titles	Share of all journals (%)
Engineering and technology	1286	42
Medicine and health	490	16
Agriculture and forestry	394	13
Total journal titles for all subjects	3052	

Source: Internal database of the State General Administration for Press and Publication, 1988

.

Table 2 Distribution of Source Journals for SCTPC in Terms of Subjects, 1988[3]

Subject category	Journal titles	Share of all journals (%)
Basic sciences	275	23.5
Engineering and technology	560	47.7
Medicine and health	193	16.2
Agriculture and forestry	129	10.8
Total journal titles for all subjects	1189	

Source: ISTIC, CSTPC

1.3 How do we adjust the structure of source journals?

Since the initial selection of source journals in 1988, they were fine-tuned each year, with new titles added and relatively poor ones phased out. China's S & T journals have witnessed great development since 1988. By the end of 2001, there had been about 4600 titles of S & T journals, among them many are about new and emerging disciplines, making the adjustment of source journals necessary. In making adjustment decisions, we ask ourselves several questions:

--Are articles in a new journal of high quality? We could rely on subject specialists to answer this question.

--Does the candidate journals follow established editing norm? For instance, is the bibliographic information complete? Do manuscripts receive peer review before they are accepted?

--Does the candidate journal have high impact? For instance, is its circulation big or small? Can it manage to attract famous scientists as its contributors?

--Does it have certain international visibility? Was it already registered by internationally influenced abstract and index services? Does the membership of its editorial board contain any internationally renowned scientists?

--Is the disciplinary and geographical representation reasonable enough? On the one hand, we give priority to journals in emerging and interdisciplinary fields. On the other hand, we heed to cover reasonable amount of journals produced in less developed provinces.

--Are relevant indicators of the current source journals satisfying enough? For each journal, we calculated its total citations, impact factor, immediacy index, proportion of papers funded by major science foundations, proportion of overseas authors, average citations of each cited paper, etc. If a journal performs very poor in terms of above indicators, we get rid of it from source journals.

In 2000, after adjustment, 1411 domestic journals (including 25 journals in English language) are selected as the source journals of CSTPC. It is safe to say that all the major journals in various disciplines have been covered.

1.4 How do we select papers to be counted?

Unlike SCI, which counts all kinds of literature in a journal as long as it is accepted as source

journal, we manually select papers to be processed into the database from all stuff in a given journal. The basic rules are as follows:

--For academic journals, all the “scientific articles” and “research letters” are selected.

--For engineering and technology journals, items to be selected are “scientific articles” and papers that explain new technology, new materials, new process and new products.

--For medical journals, items to be selected are theoretical papers, major summary report on clinical experience and substantial review articles.

No matter what kinds of journals, items such as lectures, abstracts for conference papers, speeches or pieces of practical know-how would be excluded. We hope that through such strict selection, the database would become better evaluation tools.

1.5 How do we index a paper?

Each paper is normally attributed to a single organization as indicated by the author. If more than one corporate addresses appear in a paper, the first address is used to attribute the paper. In the case of author working in hospitals affiliated to a university, the relevant paper would be attributed to both university category and hospital category, so that ranking of universities and hospitals would be possible. This paper will be counted only once despite being doubly attributed. Subject classification for each paper is done manually according to “the Discipline Classification and Code” issued by the State Quality and Technology Supervision Bureau. For every paper, our workers will ask the following questions and give corresponding codes depending on the answer: Whether is it resulted from a project funded by the National Natural Science Foundation of China or other major foundations? Whether is it contributed by overseas authors? Are co-authored papers resulted from international cooperation, inter-provincial cooperation, cooperation between researchers in different institutes in the same province, or between researchers in the same institutes? etc. In some years, we also code such information as the time lag between a manuscript was accepted and the journal was issued, the age, gender, and professional title of the first author, etc. Anyway, our indexing scope is wider than SCI.

2. Impacts and Applications

2.1 What impact has CSTPC team made upon China’s science community?

CSTPD team not only builds and maintains this database, but also collects publication statistics about Chinese scientists through SCI, EI, ISTP, etc.. CSTPC team had exerted the enormous impact upon China’s scientific community through its annual press release (first of the press conference was held in 1990) of various rankings of universities, research institutes, companies and hospitals based on CSTPC statistics and the statistics produced by ISI and further processed by ISTIC, through its data and analysis service to some important organizations, such as the National Natural Science Foundation of China (NSFC), Ministry of Education and the China Association for Science and Technology (CAST), and through its close connection with editorial boards of many S & T journals.

Prompted by our open ranking of universities and research institutes, more and more organizations introduced incentives to encourage scientists to publish more papers in influential journals. China S & T Statistical Yearbook, published by the State Statistics Bureau, formally adopt the publication figures provided by CSTPC as the official figure. Some university presidents or

other university officials remarked once and again that “CSTPC team functions as a baton in a commander’s hand. We will move to any direction where your baton points”. To be specific, through CSTPC team’s constant promotion and evaluation, more and more scientists in China have been increasingly aware of the importance to publish in international journals as well as in major domestic journals so as to enhance effective scholarly exchange and establish themselves among the community. As a result, their paper productivity has been improved a great deal. In 1987, China produced 4,880 SCI papers. In 2000, 30,499 papers were recorded by SCI, forming a more than 5-fold jump over 14 years. In 1988, CSTPC recorded 85,750 papers while in 2000, the corresponding figure is 180,848, which is more than two times of the figure 12 years ago. Tables 3 through 5 and Figure 1 show the increasing trend of China’s SCI papers.

Table 3 Growth Trend of SCI Papers by Chinese Scientists

Unit: Piece

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
China	6630	6224	9617	10411	13134	14459	16883	19838	24476	30499
World total	619972	679695	752241	790638	853822	890941	916434	930479	973286	967663

Source: ISI, cleaned by CSTPC team

Table 4 Growth Rate of China’s SCI Papers against That of the World Total, 1991-2000

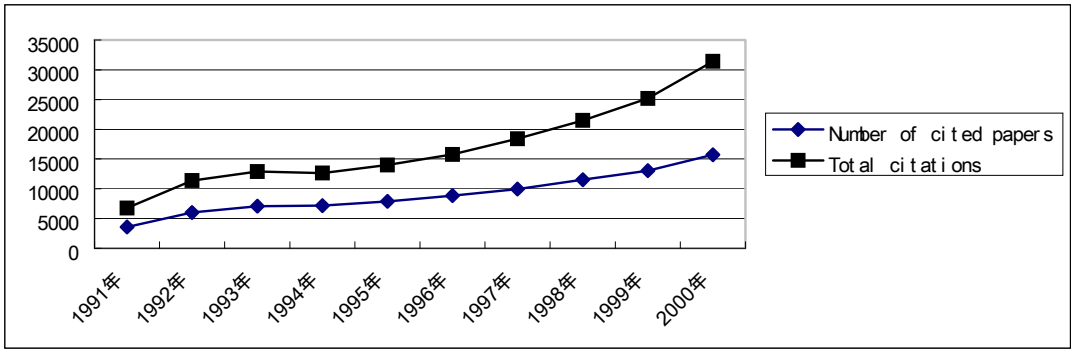
Growth rate of China over a decade（%）	360.0
Growth rate of world total over a decade（%）	56.1

Source: ISI, cleaned by CSTPC team

Table 5 Growth Trend of Citations to Papers by Chinese Scientists

year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	Growth rate over a decade
Number of cited papers	3608	5994	7060	7180	7869	8826	9952	11549	13024	15733	336%
Total citations	6771	11384	12896	12626	14000	15800	18434	21511	25173	31384	363.5%

Source: ISI, cleaned by CSTPC team



(Source: ISI, cleaned by CSTPC team)

Figure 1 Growth Trend of China's SCI papers

2.2 What impact has been produced by CSTPC database?

At first, CSTPC team only counted the number of papers through a simple database, and made relevant analysis. In 1994, the original database was transformed into a relational citation database. Since then, its application in various organizations has been increasing and its impact more visible. Table 7, Figure 2 and Figure 3 summarizes some basic statistics as derived from CSTPC.

Table 6 Growth Trend of CSTPC papers Unit: piece

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Piece of papers	94434	98966	101983	107492	107991	116239	120851	133341	162779	180848

Source: ISTIC, CSTPC

Table 7 Growth Trend of Citations Received by CSTPC Papers

	1994	1995	1996	1997	1998	1999	2000
citations	100748	104758	172385	280476	335314	466611	554324

Source: ISTIC, CSTPC

In 2001,all the papers in CSTPC cited totally 535,291 references, including journal articles, monographs, reports, conference proceedings, theses and dissertations, patents, industrial standards, etc. The distribution of these literature types are shown in Figure 2.[4]

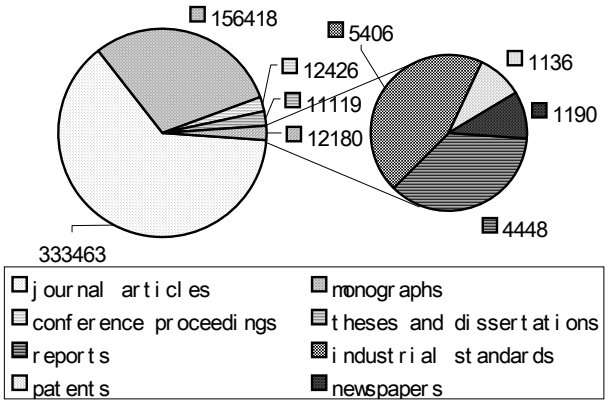
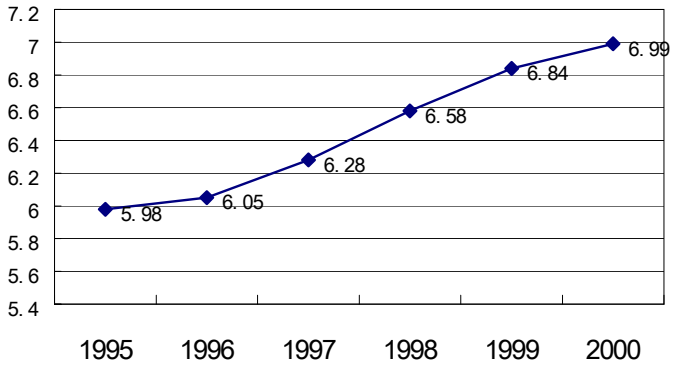


Figure 2 Distribution of Literature Types of References as Recorded in CSTPC

Source: ISI, cleaned by CSTPC team

In the past, many Chinese journals did not give enough page space for references in order to cut cost. Some editorial boards even recommend to contributors or demanded that the number of references should be less than a given number, and some went so far as to delete the reference list given by the authors and indicated that "References here have been deleted by the editor". Through the wider and wider influence of CSTPC, being selected as the source journal of CSTPC has become a great honor. Those journal producers who belittled references in the past gradually realize that if they do not publish complete references as it should, they will have no hope of being selected as the source journal, or excluded from the new source journal list. Therefore, the average number of references per paper in Chinese journals have been increasing in recent years, as shown in Figure 3.



Source: ISTIC, CSTPC

Figure 3 Increasing Trend of the Number of References per Paper, 1995-2000

As China's status of science and technology in the world and the average quality of Chinese journals get improved, more and more overseas scientists start collaborating with Chinese counterparts and they publish their solely-authored co-authored papers in Chinese journals. Table 8 indicates the distribution of countries or regions whose scientists coauthored with Chinese scientists in 2000.

Table 8 Distribution of Countries or Regions Whose Scientists Coauthored with Chinese Scientists as First Authors, 2000

Country or region	Number of joint papers with Chinese scientists as first author	Country or region	Number of joint papers with Chinese scientists as first author
USA	587	Sweden	27
Japan	566	Belgium	19
Hong Kong (SAT)	339	Denmark	19
Germany	183	Italy	19
UK	133	Switzerland	17
Australia	94	Finland	13
France	79	New Zealand	11
Canada	74	Poland	11
Singapore	48	Spain	9
Korea	44	Austria	7
Russia	33	Israel	7
Taiwan, China	32	Norway	6
The Netherlands	30	India	5

Source: ISTIC, CSTPC

2.3 What applications have been made of CSTPC?

-(1) Editing of Chinese S & T Journal Citation Report

Based on the database, CSTPC group published its first Chinese S & T Journal Citation Report in 1996. It is warmly welcomed by journal producers as well as journal administration agencies. Table 9 and Table 10 are some statistics derived from our Chinese S & T Journal Citation Report series.

Table 9 Total Citations Received by CSTPC Source Journals in 2000

Total citations	2000	
T	Journal titles	Share of total(%)
T>1000	25	1.77
500<T≦1000	90	6.38
200<T≦500	297	21.05
100<T≦200	345	24.45
50<T≦100	299	21.19
T≦50	355	25.16

Source: ISTIC, CSTPC

(2) Use of CSPTC data in various evaluations

Due to the data reliability and full coverage of source journals, CSTPC data have been used in a lot of decisions made by government agencies or non-government associations. The following applications are just a few examples:

- Evaluation Framework for the first batch of projects supported by the National Development Plan for Major Basic Research Fields, Ministry of Science and Technology;

- International position of China in various disciplines, a project conducted by CSTPC team and commissioned by China National Natural Science Foundation of China (NSFC);
- Decision on which journals should be granted with the special funds established by China Association of Science and Technology (CAST);
- Decision on which journals should be granted with the “Excellent S & T Journals Funds”, which is maintained by China National Natural Science Foundation of China (NSFC);
- The Second National Journal Award, which is jointly organized by the State General Administration for Press and Publication and Ministry of Science and Technology.
-

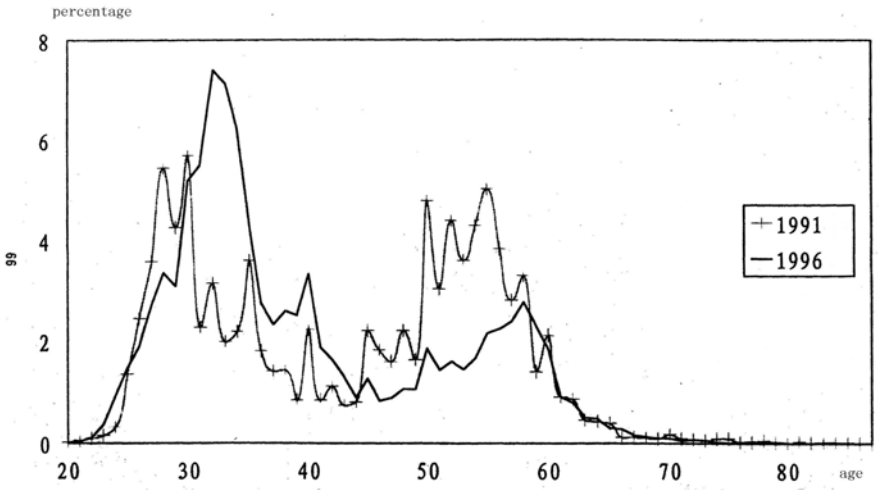
Table 10 Impact Factors (IMP) of CSTPC Source Journals

	1999 年		2000 年	
IMP	Journal titles	Share of total (%)	Journal titles	Share of total (%)
IMP>1	7	0.51	20	1.42
0.5<IMP≦1	97	7.07	108	7.65
0.1<IMP≦0.5	838	61.08	887	62.86
0.05<IMP≦0.1	274	19.97	238	16.87
0.01<IMP≦0.05	137	9.99	123	8.72
IMP≦0.05	19	1.38	13	0.92

Source: ISTIC, CSTPC

(3) Use of CSTPC data in policy recommendations

During the so-called Cultural Revolution from 1966 to 1976, all colleges and universities were cancelled. As a result, there existed a “low valley” in terms of publication amount of scientists aged 40-50 in the early 1990s, as can be seen in Figure 4. CSTPC team published a report in 1994 to alert the government about the urgency to accelerate the training of young scientists. [5]The report was well received by the science community. After about 5 years of efforts by both government and the scientific community, the “age gap crisis” was ameliorated to some extent, as our statistics shows. Scientists in their early thirties become the most productive group. We note that is phenomenon has also been discussed by Jin et al. [6].



Source: ISTIC, CSTPC

Figure 4 Age Distribution versus Productivity of Chinese Scientists, 1991 and 1996

X: Age of authors

Y: Share of papers produced by scientists at certain age in total, in percentage

3 Prospects

Since China's accession to WTO, the Chinese government will gradually loosen its control in many areas and only focus on those areas where government should bear main responsibility. Therefore, the Ministry of Science and Technology is now not willing to associate itself with the ranking of institutes according to paper productivity and citations. ISTIC will continue to release such ranking data, however, since they are welcomed by the public. In addition, CSTPC team will build new database consisting of English S & T journals published by Chinese organizations. It is our belief that since English has been de facto international language in science community, English journals will play increasingly significant role in facilitating scholarly exchange between Chinese and foreign scientists. A dedicated database will help accumulate useful data and guide relevant journals to improve their quality. When we collect sufficient data on the quality of domestic English S & T journals, we will try to recommend some of the best English journals to major international abstract and index services. In the mean time, Wanfang Data Company, a specialized database company with its majority of shares held by ISTIC, is promoting the sale of CSTPC through various approaches. The income will help improve the financial status of CSTPC team. We are very optimistic about the future of CSTPC.

4 Discussion

When ISTIC researchers began counting China's S & T papers, no one expected that CSTPC team and CSTPC database would become so popular. At present, the members of CSTPC team are often invited to give lectures in different occasions. The CSTPC database are consulted so frequently that sometimes the daily work of the group is disrupted. Why is it so?

First, SCI of ISI set an example before us so CSTPC team always know that their hard work would get rewarded.

Second, CSTPC team constantly tries new things based on their understanding of user needs. For instance, most of domestic English journals are not good enough to be accepted by SCI, but they will not be highly cited by Chinese scientists, which drags those journal producers into a quagmire. If we design a dedicated database for these English journals and give corresponding indicators, the journal producers will know what their problems are and how they could improve. Based on such considerations, CSTPC team launched such a new database in 2002.

Third, linguistic diversity is very important. Although English is the working language of international scientific community, it will not be a wise choice for scientists in non-English countries to publish only in English journals. Linguistic diversity is the prerequisite of cultural diversity, while cultural diversity leads to alternative problem solving approaches, which would benefit the whole humankind. Therefore, non-English databases like CSTPC will be useful in other countries, in particular developing countries. Spain has a database for domestic Spanish journals, called ICYT, but they have not developed an equivalent of JCR. [7] We hope to see more of such database in the world.

Fourth, the existence of other similar databases constitutes a competitive pressure, which forces CSTPC team to improve its products and services constantly. Chinese Science Citation Database, a database developed by the Documentation and Information Center, Chinese Academy of Sciences, is also very influential among the Chinese science community[8], while Chinese Social Science Citation Index, co-developed by Nanjing University and Hong Kong University of Science and technology, is more popular among Chinese scholars of social science and humanities[9]. It should be noted, though, that increasingly widespread application of such databases further creates a stronger demand for them. Therefore, the existence of other databases also helps CSTPC to some extent, and vice versa.

ACKNOWLEDGEMENT This study was supported by a grant (No. 70073007) from the National Natural Science Foundation of China.

References

[1] Ministry of Science and Technology, Survey Data on China's R & D Resources, May 2002 (in Chinese)

[2] Jing Qin-shu, China's Core Journals in Natural Sciences—Application of China Natural Science Citation Index, World Books, 1988, No.1 (in Chinese)

[3] ISTIC, Statistics and Analysis of China's S & T Papers, Science, Technology and Development, 1990, 5 (2): 1-9 (in Chinese)

[4] ISTIC, Chinese S & T Papers Statistics and Analysis 2000, an annual report released by ISTIC,

December 2001 (in Chinese)
[5] ISTIC, Chinese S & T Papers Statistics and Analysis 1993, an annual report released by ISTIC, December 1994 (in Chinese)
[6] Jin, BH, Li, L. & Rousseau, R. Production and productivity of Chinese scientists as a function of their age: the period 1995-1999, These Proceedings, 2003.
[7] Bordons, M. , Fernandez, M.T. & Gómez, I. . Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country, Scientometrics, 2002, 53 (2): 195-206
[8] Jin, BH & Wang, B. Chinese Science Citation Database: its construction and application, Scientometrics, 1999, 45: 325-332
[9] Su, XN, Han, XM. & Han, XN., Developing the Chinese Social Science Citation Index, Online Information Review, 2001, 25 (6): 365-369

A revised version of this contribution has been published as:

Yishan Wu, Yuntao Pan, Yuhua Zhang, Zheng Ma, Jingan Pang, Hong Guo, Bo Xu, Zhiqing Yang (2004). China Scientific and Technical papers and Citations (CSTPC): History, impact and outlook. Scientometrics, 60(3), 385-397.