

# *An Introduction to the Barycentre Method with an Application to China's Mean Centre of Publication*

BIHUI JIN AND RONALD ROUSSEAU

The Documentation & Information Centre of the Chinese Academy of Sciences (DICCAS),  
Beijing, P. R. China

KHBO – (Catholic School for Higher Education Bruges-Ostend) Department Industrial Sciences and  
Technology, Oostende & IBW (Informatie- en Bibliotheekwetenschap), UIA (University of Antwerp),  
Wilrijk, Belgium

---

We explain the barycentre method and show how to apply it in a practical situation. A country's mean centre of publication, or publication barycentre, is defined and its evolution is studied in the case of China's centre of publication. Data for this application are taken from the Chinese Science Citation Database (CSCD). It is shown that China's publication

barycentre slowly moves to the south and that it approaches the population barycentre. Tentative explanations for this southward movement, based on economic and educational considerations, are given. It is observed that inequality in publication between administrative units (regions) decreases slowly.

---

## *Introduction*

China's mean centre of population, or population barycentre, is defined as the imaginary point at which a flat, weightless but stiff map of China would balance if weights of identical value were placed on it so that each weight represented the location of one person (Bartlett 1985). Similarly, China's mean centre of publication (its publication barycentre) is the imaginary point at which a flat, weightless but stiff map of China would balance if weights of identical value were placed on it so that each weight represented the place of origin of one article.

In this article we explain the barycentre method and show its advantages. As an application we study a) if the publication centre coincides with the mean centre of population and b) how these two barycentres evolved over a ten-year period.

As the source for publication data we use the Chinese Science Citation Database (CSCD) (Jin & Wang 1999, Rousseau et al. 2001). China is the first non-Western country that has developed its own citation index as a complement to ISI's Science Citation Index (SCI). The People's Republic of China publishes more than 4000 journals in science and technology, but only a small number of these (about 50) is covered by ISI's Web of Science. This means that it makes little sense to use the SCI as the only source for information retrieval, or for research evaluation purposes in China. For this reason it was decided already in 1989 to develop a local, i.e. Chinese citation database at the Documentation and Information Centre of the Chinese Academy of Sciences (DICCAS, <http://www.las.ac.cn/>). Up to 1995, 315 source journals were included in the CSCD. Since 1996 this group has been expanded to yield a group of

---

Bihui Jin, The Documentation & Information Centre of the Chinese Academy of Sciences (DICCAS), 8 Kexueyuan Nanlu, Zhongguancun, Beijing 100080, P. R. China. E-mail: [csci@mail.las.ac.cn](mailto:csci@mail.las.ac.cn).  
Ronald Rousseau, KHBO, Zeedijk 101, B-8400 Oostende, Oostende & IBW, UIA, Universiteitsplein 1, B-2610 Wilrijk, Belgium.  
E-mail: [ronald.rousseau@kh.khbo.be](mailto:ronald.rousseau@kh.khbo.be) Web page: <http://users.pandora.be/ronald.rousseau>

Table 1: Most productive Chinese institutes according to the CSCD (1996)

	set of 582 journals	set of 315 journals
Qinghua University	1248	889
Beijing Medical University	916	693
Huazhong Univ. of S&T	892	639

582 core titles, accounting for approximately 14% of all science and technology journals published in China. All journals covered by SCI-expanded are among these 582. In 1996 there were 5859 Chinese institutes publishing at least one article included in the CSCD. Of these 5859 the most active institute published 1248 articles. There were also 3350 institutes with only one article covered. If we restrict ourselves to the original 315 source journals, then there are 3123 'active' institutes (in 1996); the number one in the publication list had 889 articles, and 1698 institutes are included on the basis of just one article. Table 1 gives the top three institutes for both sets and their production. These source-item relations (institutes producing articles) can roughly be described by a Lotka law with exponent 1.8 (Rousseau and Rousseau 2000).

As can be derived from its name, the CSCD only covers journals publishing articles in the fields of science and technology. Yet some multi-disciplinary journals are included such as the *Chinese Science Bulletin* and *Progress in Natural Science*. The CSCD includes journals in the following categories: mathematics, physics & mechanics, chemistry, astronomy, geosciences, biology, agriculture & forestry, medicine, engineering & technology, and environmental sciences. For the exact distribution of journals over fields we refer to (Jin & Wang 1999). The CSCD grows nowadays by more than 56,000 articles a year.

For budgetary reasons only citations to Chinese authors are included in the CSCD (this means about one-third of all citations but inclusion of foreign authors is planned for the near future). Note that this omission has no influence on the evaluation of Chinese scientists, institutes or universities. Via an authority file every university, hospital, state lab or research institute receives a unique code. Source journals of the CSCD are journals with an emphasis on basic re-

Table 2: Chinese regions and geographic co-ordinates (gc) in decimal form; E: east; N: north.

Province	Capital	gc: E	gc:N
Anhui	Hefei	117.300	31.917
Beijing	municipality	116.433	39.917
Chongqing	municipality	103.683	30.650
Fujian	Fuzhou	119.283	26.150
Gansu	Lanzhou	103.750	36.017
Guangdong	Guangzhou	113.333	23.133
Guangxi Zhuang	Nanning	108.317	22.833
Guizhou	Guiyang	106.717	26.583
Hainan	Haikou	110.417	20.083
Hebei	Shijiazhuang	114.467	38.067
Heilongjiang	Harbin	126.683	45.750
Henan	Zhengzhou	113.633	34.750
Hubei	Wuhan	114.317	30.583
Hunan	Changsha	113.000	28.167
Inner Mongolia	Hohhot	111.617	40.817
Jiangsu	Nanjing	118.783	32.050
Jiangxi	Nanchang	115.917	28.700
Jilin	Changchun	125.333	43.833
Liaoning	Shenyang	123.433	41.833
Ningxia Hui	Yinchuan	106.317	38.500
Qinghai	Xining	101.917	36.583
Shaanxi	Xi'an	108.900	34.267
Shandong	Jinan	117.000	36.683
Shanghai	municipality	121.367	31.100
Shanxi	Taiyuan	112.500	37.833
Sichuan	Chengdu	104.100	30.617
Tianjin	municipality	117.200	39.133
Tibet = Xizang	Lhasa	91.167	29.683
Xinjiang Uygur	Urumqi	87.633	43.717
Yunnan	Kunming	102.683	25.067
Zhejiang	Hangzhou	120.117	30.300

search and theory, or research in high-tech areas (for engineering & technology journals). They should include articles describing research at the forefront of the particular field. Indicators for the evaluation of journals (when screening for possible inclusion in the CSCD) are: the journal's overall citation rate, its impact factor, its coverage by international databases, and its inclusion in the *Chinese Core Journal List* compiled by Peking University. 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. Indeed, it has been shown (Jin et al. 1997) that the 582 journals covered by the CSCD, i.e. less than 15 % of all Chinese journals, receive more than 50% of all citations. This means that the CSCD is a representative source for the analysis of Chinese science and technology.

Table 3: Publication data according to the CSCD

Region	Publications (ranks)									
	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Anhui	658 (13)	688 (13)	659 (14)	760 (14)	792 (15)	842 (13)	881 (13)	1531 (15)	1617 (15)	1719 (14)
Beijing	6916 (1)	7024 (1)	7038 (1)	7578 (1)	7748 (1)	7926 (1)	8572 (1)	13367 (1)	14299 (1)	14517 (1)
Fujian	597 (15)	594 (15)	605 (15)	716 (15)	800 (14)	800 (14)	670 (16)	1488 (16)	1611 (16)	1548 (16)
Gansu	630 (14)	657 (14)	676 (13)	782 (13)	810 (13)	810 (13)	817 (14)	1156 (17)	1160 (17)	1086 (17)
Guangdong	1015 (9)	1035 (9)	1115 (9)	1195 (9)	1286 (9)	1286 (9)	1788 (7)	3681 (6)	4039 (5)	4238 (4)
Guangxi Zhuang	110 (25)	95 (25)	120 (25)	113 (26)	114 (26)	114 (26)	116 (25)	405 (23)	376 (25)	426 (22)
Guizhou	136 (24)	128 (24)	123 (24)	115 (24)	133 (24)	133(24)	169 (22)	328 (26)	342 (26)	338 (26)
Hainan	44 (28)	43 (28)	30 (28)	43 (28)	43 (28)	43 (28)	56 (27)	91 (28)	83 (28)	106 (28)
Hebei	287 (19)	319 (18)	366 (18)	401 (19)	394 (19)	394 (19)	414 (18)	904 (18)	950 (19)	1012 (18)
Heilongjiang	474 (17)	590 (16)	573 (16)	582 (16)	635 (16)	635 (16)	675 (15)	1719 (14)	1766 (14)	1703 (15)
Henan	287 (19)	302 (19)	317 (19)	410 (18)	432 (18)	432 (18)	407 (19)	867 (19)	971 (18)	985 (19)
Hubei	1315 (6)	1537 (5)	1540 (5)	1582 (5)	1667 (5)	1667 (5)	1994 (5)	3724 (5)	3949 (6)	3881 (6)
Hunan	494 (16)	485 (17)	550 (17)	530 (17)	600 (17)	600 (17)	622 (17)	1947 (13)	2084 (12)	2148 (12)
InnerMongolia	82 (26)	91 (26)	92 (26)	115 (24)	118 (25)	118 (25)	93 (26)	429 (22)	429 (22)	405 (23)
Jiangsu	1955 (3)	1896 (3)	2077 (3)	2294 (3)	2389 (3)	2389 (3)	2368 (3)	4967 (3)	5036 (3)	5134 (3)
Jiangxi	142 (23)	138 (23)	144 (23)	138 (23)	160 (23)	160 (23)	146 (24)	402 (24)	410 (23)	352 (24)
Jilin	1066 (8)	1321 (7)	1349 (7)	1456 (7)	1361 (8)	1361 (8)	1376 (9)	2329 (10)	2334 (10)	2213 (11)
Liaoning	1112 (7)	1261 (8)	1261 (8)	1341 (8)	1436 (7)	1436 (7)	1667 (8)	2353 (9)	2558 (9)	2623 (9)
Ningxia Hui	19 (29)	29 (29)	30 (28)	40 (29)	27 (29)	27 (29)	27 (29)	64 (29)	81 (29)	93 (29)
Qinghai	57 (27)	81 (27)	71 (27)	69 (27)	55 (27)	55 (27)	56 (27)	127 (27)	114 (27)	111 (27)
Shaanxi	1385 (5)	1379 (6)	1415 (6)	1514 (6)	1585 (6)	1585 (6)	1880 (6)	3341 (7)	3437 (7)	3584 (7)
Shandong	808 (11)	907 (11)	926 (10)	994 (10)	1120 (10)	1120 (10)	1250 (10)	2722 (8)	2815 (8)	2956 (8)
Shanghai	3303 (2)	3464 (2)	3521 (2)	3419 (2)	3489 (2)	3489 (2)	3997 (2)	6262 (2)	6400 (2)	6728 (2)
Shanxi	177 (21)	200 (21)	249 (21)	220 (21)	300 (20)	300 (20)	319 (20)	787 (20)	740 (21)	824 (21)
Sichuan	1668 (4)	1728 (4)	1745 (4)	1882 (4)	1959 (4)	1959 (4)	2003 (4)	3993 (4)	4058 (4)	4056 (5)
Tianjin	807 (12)	839 (12)	804 (12)	962 (12)	963 (12)	963 (12)	1192 (11)	2165 (11)	2043 (13)	2132 (13)
Tibet = Xizang	10 (30)	7 (30)	12 (30)	10 (30)	10 (30)	10 (30)	13 (30)	18 (30)	21 (30)	22 (30)
Xinjiang Uyгур	157 (22)	162 (22)	170 (22)	181 (22)	170 (22)	170 (22)	159 (23)	346 (25)	384 (24)	349 (25)
Yunnan	289 (18)	295 (20)	301 (20)	327 (20)	265 (21)	265 (21)	254 (21)	638 (21)	780 (20)	830 (20)
Zhejiang	821 (10)	928 (10)	923 (11)	973 (11)	985 (11)	985 (11)	1031 (12)	2063 (12)	2174 (11)	2425 (10)
Total	26821	28191	28802	30742	31846	31846	35012	64214	67061	68444

Data

As publications originating in an administrative unit (often a province) will be assigned to the capital of this province, the first set of data we need consists of the geographic co-ordinates of the capitals of China’s administrative units. Table 2 gives the administrative units (provinces, autonomous regions and municipalities) of the People’s Republic of China (for diverse reasons Hong Kong, Macao and Taiwan are not included). Chongqing became a municipality in 1997. In our data it is considered a part of Sichuan as it was before 1997. The capitals’ geographic co-ordinates (in decimal form for easy calculation) were obtained from the Getty Thesaurus of Geographic Names (Getty). Figure 1 shows a map of China and its regions.

Table 3 shows the yearly publications per region according to the CSCD. It also ranks the

Figure 1. The People’s Republic of China: administrative units (note that Macao, although an integral part of China, is not shown on the map because publication data are small, and, moreover, not used in this study).



Table 4: Centres of publication, CPub, for the period 1989 – 1998

Year	Longitude: E	Latitude: N
1989	115.9854	34.9415
1990	116.1227	35.0612
1991	116.0967	34.9960
1992	116.0317	35.0512
1993	116.0364	34.9883
1994	116.1057	34.9289
1995	116.0603	34.9233
1996	115.8168	34.5052
1997	115.7953	34.4705
1998	115.7994	34.3859

Table 5: Centres of population, CPop, for the period 1989 – 1998

Year	Longitude: E	Latitude: N
1989	113.4509	32.6575
1990	113.4549	32.6575
1991	113.4464	32.6525
1992	113.4405	32.6445
1993	113.4317	32.6369
1994	113.4220	32.6299
1995	113.4137	32.6257
1996	113.4041	32.6210
1997	113.3963	32.6159
1998	113.3855	32.6083

regions according to article production. These ranks are fairly stable over the years, with Beijing at the top, followed by Shanghai. Yet, China’s high-tech region of Guangdong ranked ninth in 1989 and went up to fourth in 1998. As the focus of this application is on publications, not on population issues, we have relegated all data concerning China’s population to appendix B.

### The barycentre method

Details about the barycentre method, and about the way calculations must be handled are given in Appendix A. Barycentres, as defined in the introductory section of this article, were calculated for every year in the period 1989–1998 (see Tables 4 and 5). The mean centre of publication is de-

noted as CPub, the population centre is denoted as CPop.

The centre of publication has moved to the south, and slightly to the west. Between 1989 and 1998 it has traversed about 64 km. In the beginning of this period it was still situated in the province of Shandong, but now it is in Henan, near the city of Shangqui. The centre of population on the other hand is situated at approximately 113.4 degrees east and 32.6 degrees north. Also this centre has moved further to the south, but much more slowly. Over the same period, it only moved about 8 km, staying within the borders of the province of Henan, in the neighbourhood of Biyang, south of the city of Zhumadian. Comparing the locations of both centres shows that the centre of publication is situated more to the north, and more to the east. In other words, the publication barycentre is closer to Beijing than the population centre (see Figure 2). For reasons of comparison we note that the centre of publication of the Scandinavian countries (based on SCI data) traversed a distance of 23 km over a 12-year period (1974–1986) (Rousseau 1989a).

We would like to point out here that the barycentre method adds a new dimension to the study of, for instance, the research output of a country, or a group of countries. Indeed, finding the barycentre yields a kind of ‘summary statistic’, similar to an average number of publications, or to an index measuring the publication inequality between regions. A simplified example will make this clear. Assume that a country has four research centres, two in the north and two in the south. In the first situation (situation A) the two centres in the north have a large research

Figure 2. China’s centre of publication and its centre of population. Both are now situated in the province of Henan.



output while the two other ones, in the south, have only a small output. Situation B is just the opposite, and in situation C two centres publish many articles, while the two other ones publish little, in each case one in the north and one in the south. The average output is the same in the three situations, and also the output inequality is the same. Yet the barycentre for situation A lies in the north, for situation B it lies in the south, and for situation C it is situated approximately in the middle. This clearly illustrates the extra dimension (yielding more information) obtained through the use of barycentres.

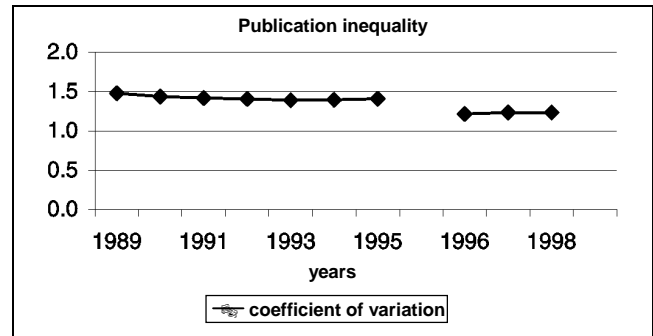
We further note that the calculations involved in obtaining the co-ordinates of a barycentre are easy. They can be done in any spreadsheet. No sophisticated mathematical or statistical package is needed. Finally, the barycentre method leads to a natural visual representation on a medium well known to everybody, namely a geographic map. In the case the units under study do not form a geographic whole, the method can be abstracted, and visualisation can take place in a regular polygon (Rousseau 1989b).

### *Tentative explanations*

We offer two tentative reasons for the southward movement of the centre of publication. The first being the fact that the southern provinces Fujian and Guandong attract a lot of high-tech industrial development (besides of course Beijing, Shanghai and the province of Shandong) and hence also more and more research is done in that area. The other one being that China's policy (in all domains) becomes less centred in Beijing.

The slow movement of the population centre has several reasons. First, we could point out that in general a big country, such as China, has a large inertia, so things, such as a centre of population do not change rapidly. Moreover, the centre of population is situated in a densely populated area. But, there are also more specific reasons: China's strategic population policy (control of growth through the one-child policy in order to improve the quality of living), China's strict rules on settlement in the past years (people can move around freely, but not settle in another city freely), and finally, explaining the small southward movement, the fact that Guandong, situated

Figure 3. Publication inequality as measured by the coefficient of variation.



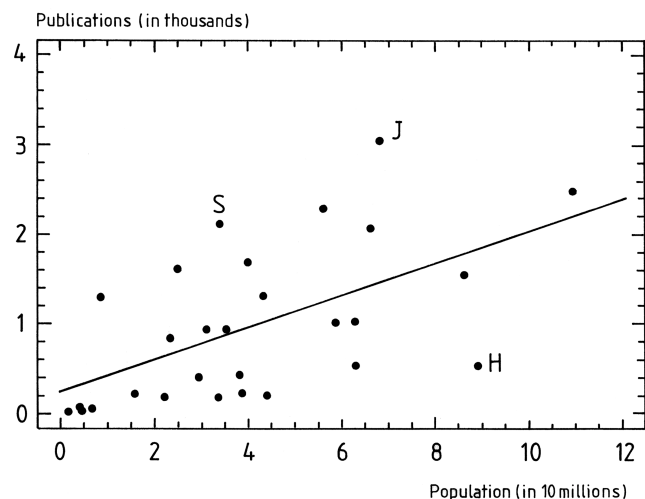
in the south, attracts many people for the development of the new economic regions.

### *Inequality of publication versus inequality of population*

It is clear that there is a large inequality in publication between China's administrative units. Inequality can be measured in different ways, the simplest ones being the coefficient of variation and the Gini index. The coefficient of variation is nothing but the standard deviation divided by the mean. We will not go into details about a general theory of inequality (concentration, evenness) but refer the reader to Egghe & Rousseau (2001) and Rousseau (2000). Let us just say that concentration can be described as the relative apportionment of items among the sources present. In this example sources are China's different regions, and items are the publications originating from these regions. Figure 3 shows the change in publication inequality over the period 1989–1998. Generally, the inequality between the publication outputs of China's different regions decreases slowly. The update of CSCD's database increased this process considerably.

The population inequality in this period as measured by the coefficient of variation is about 0.68. This number is considerably smaller than that of the publication inequality. Such a result, however, is not surprising. It has been observed on several occasions (Allison 1980, Rousseau 1992) that in general the inequality in 'use' is larger than the inequality in 'availability'. Examples of availability-use pairs are publications and their citations; available CDs in a public library and the number of loans; and in our case, population and publications. It was suggested (Rousseau

Figure 4. Linear regression line for the variables 'population' and 'publication'.



1992) that the underlying mechanism for this phenomenon is a kind of positive reinforcement. Activities that are rewarded continue to be performed, while activities that are not rewarded tend to be stopped.

*Is there a linear relation between population and publication?*

In this section we investigate whether there is a linear relation between population and publication. Population is the simplest explicative variable for the number of publications (more refined variables, to be studied in subsequent work, are the S&T budget of the region, the regional GDP, the number of skilled workers, the numbers of university graduates, etc.). As a first step we just consider here the number of habitants of the region. Applying a linear regression between the variables 'population', measured through the average of the ten-year period, and the variable 'publication', also measured through its average value, yields the following linear relation:

$$\text{publication} = 251.7 + 0.178 * \text{population} \quad (2)$$

'Publication' has been considered as the variable depending on the independent variable 'population'. In equation (2) population is expressed in 10.000s (see Table 7). Figure 4 shows the data and the best fitting regression line. The municipalities of Beijing and Shanghai are left out of this equation, because they would have distorted the result (they could not even be placed in Figure 4).

Pearson's product moment correlation coefficient for this linear regression is 0.57. A t-test, with null hypothesis 'no correlation': yields a t-value of 3.536. As the acceptance region for the null-hypothesis at the 95% level is [-2.06 ; + 2.06], we reject the null-hypothesis (no correlation) and conclude that the variables population and publication are correlated. For details on this test we refer the reader to Egghe & Rousseau (1990, I: 3.8.5.)

Observing which provinces deviate most from this general relationship, and trying to explain this deviation is then the next (and most interesting) step. It seems that Jiangsu and Shaanxi have a higher number of publications than expected, while Henan has a significantly lower number of publications (these provinces are indicated with the letters J, S and H in Figure 4). Jiangsu is a province with a lot of industry and is very active in high-tech related areas; it has moreover its fair share of universities and research institutes. The province of Shaanxi on the other hand, has a lot of universities and institutes for higher education. Finally, Henan is a densely populated agricultural province. These factors could, at least partially, explain the observed differences.

*A partial corroboration of our findings*

The production of scientific knowledge in a region depends on the general economic 'climate' and specific local factors. Such factors related to S&T resources are talent and potential of the inhabitants (scientists and engineers), local S&T expenditures and the presence of universities, scientific engineers and science and industrial parks. The province of Guandong has many highly educated inhabitants and receives a lot of S&T money both from national as well as local authorities. This explains its quick rise. The province of Shaanxi on the other hand has many universities and research institutes between its borders. Finally, Jiangsu's prominence is based on all of these factors.

In a recent article You Guangrong (2000) studied the competitiveness of China's regions. Besides a ranking, he also provided a broad categorisation based on weighting different aspects related to S&T (including publications). He distinguished five categories, indicated as: I: very strong, II: strong, III: moderate, IV: weak, V: very

weak. The complete ranking is given in Table 6. It is clear that the regions we detected as strong in publication output or stronger than expected in publication output coincide with highly competitive regions.

Note that the original list consists of 31 regions. This is because it includes Chongqing (at rank 22).

The Spearman rank correlation (Egghe & Rousseau 1990, formula I: 3.42) of the 1998 publication output according to the CSCD and this regional S&T competitive ranking is 0.92, which is, of course, highly significant. This shows that, although S&T competitiveness and scientific publishing are not the same, they are highly correlated. One could say that the results of scientific research fuel the country's technological progress.

### Conclusion

We have explained the barycentre method and shown, based on data from the Chinese Science Citation Database that China's mean centre of publication moves southward, and is now situated in the province of Henan. It slowly approaches China's mean centre of population (also situated in Henan). We have also shown that the inequality in scientific output between China's administrative units slowly decreases over time. Further, as a general trend, provinces with a larger population produce more scientific articles. This, of course, is not surprising. Interesting cases are those provinces that deviate from this general trend, such as Jiangsu, Shaanxi and Henan. We have also noted the fast rise of Guangdong. This article shows that it is possible to do publication studies for China based on a local source. Of course, it would be better still to combine the facts included in Chinese and Western databases. This is precisely what we hope to accomplish in the near future.

As to the method used, we would like to state that it could be applied in many other contexts. Other examples in the library context are the number of loans in all libraries from a geographic area (country, province, county), the number of library visits, the number of library clients (independent of the actual number of loans, or visits), and so on. In this case libraries are the sources, playing the role of China's provincial capitals.

Table 6: Regional S&T competitive ranking in 1998

Regions	1998	1998
Anhui	18	IV
Beijing	1	I
Fujian	14	III
Gansu	24	IV
Guangdong	2	I
Guangxi Zhuang	20	IV
Guizhou	27	V
Hainan	28	V
Hebei	13	III
Heilongjiang	11	III
Henan	12	III
Hubei	7	II
Hunan	16	III
Inner Mongolia	26	IV
Jiangsu	4	I
Jiangxi	25	IV
Jilin	17	III
Liaoning	6	II
Ningxia Hui	29	V
Qinghai	30	V
Shaanxi	15	III
Shandong	5	II
Shanghai	3	I
Shanxi	23	IV
Sichuan	9	II
Tianjin	10	III
Tibet = Xizang	31	V
Xinjiang Uygur	21	IV
Yunnan	19	IV
Zhejiang	8	II

The first factor one considers in explaining the results obtained through the barycentre method, is the number of inhabitants of the library's service area. Indeed, the more inhabitants, the more loans, visitors, etc. If the factor 'population' would explain all differences between libraries, the centre of population would coincide with the barycentre of loans, library visits, persons with a library card, and so on. If not, the difference between these two centres gives us information about the geographic inequality of the subject under study. The next step in a research project would then be to find an explanation for this observed difference.

Finally, calculating barycentres at different times introduces a dynamic aspect to this kind of study. Does the barycentre move? In one particular direction, or rather erratically? Does the movement of the barycentre follow that of the population centre or not? Again, explanations are

in order. All this shows that the barycentre method deserves more attention from the LIS community than it has received till now.

### Acknowledgement

The authors thank the Natural Science Foundation of China (NSFC) for supporting their research. They also thank Professor Liang Liming (Henan Normal University) for stimulating discussions concerning the coefficient of variation as a measure of concentration. Finally, they also thank the referees and the editor for constructive suggestions to improve the presentation of this article.

### References

Allison, P.D. 1980. Inequality and scientific productivity. *Social Studies of Science* 10: 163–79.  
 Bartlett, A.A. 1985. U.S. population dynamics. *American Journal of Physics* 53: 242–48.  
 Egghe, L. & R. Rousseau. 1990. Introduction to Informetrics. Quantitative methods in library, documentation and information science. Elsevier: Amsterdam.  
 Egghe, L. & R. Rousseau. 2001. Symmetric and asymmetric theory of relative concentration and applications. *Scientometrics* 52: 261–90.  
 Getty Thesaurus of Geographic Names. N.D. URL: <http://www.getty.edu/research/tools/vocabulary/tgn/index.html> [Viewed November 10, 2001].  
 Jin, B., Cheng J., Jiang X. & Jiang Y. 1997. Selection and evaluation of the source journals of the Chinese

Science Citation Database (in Chinese). *Chinese Journal of Scientific and Technical Periodicals* 8(4): 17–21.  
 Jin, B. & Wang, B. 1999. Chinese Science Citation Database: its construction and application. *Scientometrics* 45: 325–32.  
 Rousseau, B. & R. Rousseau. 2000. LOTKA: a program to fit a power law distribution to observed frequency data. *Cybermetrics* 4: paper 4. URL: <http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html> [Viewed November 10, 2001].  
 Rousseau, R. 1989a. Kinematical statistics of scientific output. Part I: geographical approach. *Revue française de bibliométrie* 4: 50–64.  
 Rousseau, R. 1989b. Kinematical statistics of scientific output. Part II: standardized polygonal approach. *Revue française de bibliométrie* 4: 65–77.  
 Rousseau, R. 1992. Concentration and diversity of availability and use in information systems: a positive reinforcement model. *Journal of the American Society for Information Science* 43: 391–95.  
 Rousseau, R. 2000. Concentration and evenness measures as macro-level scientometric indicators. Paper presented at the Second International Symposium on Quantitative Evaluation of Research Performances and Sixth National Conference on Scientometrics and Informetrics (October 23–25, Shanghai, China).  
 Rousseau, R., Jin, B., Yang, N. & Liu, X. 2001. Observations concerning the two- and three-year synchronous impact factor, based on the Chinese Science Citation Database. *Journal of Documentation* 57: 349–57.  
 You, G. 2000. The evaluation of regional competitiveness in science and technology (in Chinese). Paper presented at the first all-China Symposium on the Evaluation of S&T, Beijing.

### Appendix A: Calculation of barycentres

The centre of publication, or publication barycentre, defined in the introductory section and denoted as C<sub>Pub</sub>, has co-ordinates (c<sub>x</sub>, c<sub>y</sub>) calculated according to the following formula (Rousseau 1989a,b):

$$c_x = \frac{\sum_{i=1}^n m_i L_{i,x}}{M}, \quad c_y = \frac{\sum_{i=1}^n m_i L_{i,y}}{M} \quad (1)$$

with  $M = \sum_{i=1}^n m_i$

Here  $n$  denotes the number of administrative units used in this study (i.e. 30);  $L = (L_{i,x}, L_{i,y})$  denotes the geographic location of the  $i$ -th administrative unit (usually the co-ordinates of the capital);  $m_i$  denotes the ‘mass’ of the  $i$ -th administrative unit and  $M$  is the total ‘mass’ of the system under study. For C<sub>Pub</sub> the  $m_i$ ’s are numbers of publications, for the population centre of China, C<sub>Pop</sub>, the  $m_i$ ’s will be population. Clearly,  $c_x$  and  $c_y$  are weighted averages of the location co-ordinates,  $L_{i,x}$  and  $L_{i,y}$ . Weighting is done according to the number of publications (or the population). By its construction, the barycentre always belongs to the convex hull of the  $L$ -vectors.

We recall, as mentioned in the text, that this calculation can be performed for other source-item relationships too.

## Appendix B: Population (in ten thousands)

Source: CHINA POPULATION STATISTICS YEARBOOK 1994–1999, published by China Statistics Press: Beijing (1994–1999)

Region	Population (rank)									
	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Anhui	5469.2 (8)	5660.7 (8)	5744.0 (8)	5817.5 (8)	5870.0 (8)	5937.9 (8)	5999.6 (8)	6054.0 (8)	6109.2 (8)	6152.2 (8)
Beijing	1024.7 (25)	1035.7 (25)	1042.9 (25)	1048.7 (25)	1056.9 (25)	1068.2 (25)	1077.0 (25)	1083.2 (25)	1092.3 (25)	1097.8 (25)
Fujian	2889.0 (18)	2999.8 (18)	3039.0 (18)	3066.9 (18)	3099.2 (18)	3126.9 (18)	3164.6 (18)	3210.6 (18)	3237.1 (18)	3260.8 (18)
Gansu	2170.8 (21)	2229.9 (21)	2258.0 (21)	2288.1 (21)	2318.6 (21)	2352.4 (21)	2388.4 (21)	2427.8 (21)	2456.6 (21)	2483.6 (21)
Guangdong	6025.0 (5)	6246.3 (5)	6349.0 (5)	6463.2 (5)	6581.6 (5)	6691.5 (5)	6788.7 (5)	6898.8 (4)	7013.7 (4)	7115.6 (4)
Guangxi Zhuang	4150.0 (11)	4241.6 (10)	4294.5 (10)	4359.4 (10)	4408.8 (10)	4455.1 (10)	4502.1 (10)	4545.5 (10)	4588.5 (10)	4622.2 (10)
Guizhou	3184.0 (17)	3237.0 (17)	3271.4 (17)	3301.0 (17)	3332.3 (17)	3380.6 (17)	3419.5 (17)	3459.5 (16)	3495.5 (16)	3536.7 (16)
Hainan	638.8 (27)	651.2 (27)	661.5 (27)	671.3 (27)	681.8 (27)	691.4 (27)	702.4 (27)	714.1 (27)	724.5 (27)	733.3 (27)
Hebei	5888.4 (7)	6116.8 (6)	6183.2 (6)	6249.3 (6)	6309.6 (6)	6366.0 (6)	6420.5 (6)	6461.0 (6)	6508.1 (6)	6555.3 (6)
Heilongjiang	3442.4 (15)	3488.9 (15)	3510.7 (15)	3526.2 (15)	3538.9 (15)	3557.6 (15)	3576.8 (15)	3605.1 (15)	3628.5 (15)	3642.0 (15)
Henan	8231.1 (2)	8564.4 (2)	8687.0 (2)	8811.5 (2)	8914.6 (2)	9005.0 (2)	9108.8 (2)	9203.1 (2)	9292.7 (2)	9372.7 (2)
Hubei	5223.9 (9)	5373.5 (9)	5446.8 (9)	5513.6 (9)	5590.5 (9)	5656.8 (9)	5727.1 (9)	5776.4 (9)	5838.8 (9)	5890.6 (9)
Hunan	6013.6 (6)	6110.6 (7)	6167.0 (7)	6209.0 (7)	6248.0 (7)	6305.9 (7)	6356.7 (7)	6403.9 (7)	6444.1 (7)	6482.2 (7)
Inner Mongolia	2112.2 (22)	2149.4 (22)	2164.8 (22)	2178.5 (22)	2198.0 (22)	2217.4 (22)	2237.2 (22)	2263.0 (22)	2288.5 (22)	2310.2 (22)
Jiangsu	6535.9 (4)	6671.7 (4)	6733.9 (4)	6767.5 (4)	6800.7 (4)	6831.3 (4)	6868.4 (4)	6908.1 (5)	6948.4 (5)	6983.1 (5)
Jiangxi	3695.2 (13)	3761.4 (13)	3801.9 (13)	3827.0 (13)	3857.2 (13)	3893.7 (13)	3938.6 (13)	3981.0 (13)	4026.0 (13)	4070.6 (13)
Jilin	2395.4 (20)	2440.2 (20)	2459.7 (20)	2474.0 (20)	2496.1 (20)	2515.6 (20)	2550.9 (20)	2579.1 (20)	2600.1 (20)	2603.2 (20)
Liaoning	3876.0 (12)	3917.4 (12)	3938.5 (12)	3957.9 (12)	3982.9 (12)	4007.2 (12)	4034.0 (12)	4056.8 (12)	4077.1 (12)	4090.4 (12)
Ningxia Hui	454.8 (28)	465.7 (28)	473.8 (28)	482.3 (28)	490.9 (28)	503.9 (28)	512.4 (28)	521.2 (28)	528.9 (28)	536.6 (28)
Qinghai	427.3 (29)	434.8 (29)	439.4 (29)	443.1 (29)	446.3 (29)	451.0 (29)	456.2 (29)	462.6 (29)	466.5 (29)	470.3 (29)
Shaanxi	3198.0 (16)	3275.0 (16)	3309.9 (16)	3340.3 (16)	3369.7 (16)	3401.6 (16)	3431.9 (16)	3457.7 (17)	3482.7 (17)	3501.1 (17)
Shandong	8180.9 (3)	8423.6 (3)	8534.0 (3)	8579.8 (3)	8620.4 (3)	8652.6 (3)	8701.2 (3)	8747.0 (3)	8809.7 (3)	8871.5 (3)
Shanghai	1276.4 (24)	1283.4 (24)	1287.2 (24)	1289.4 (24)	1294.7 (24)	1298.8 (24)	1301.1 (24)	1304.4 (24)	1305.4 (24)	1306.6 (24)
Shanxi	2774.4 (19)	2845.2 (19)	2883.4 (19)	2919.1 (19)	2955.5 (19)	2990.9 (19)	3025.7 (19)	3059.2 (19)	3091.3 (19)	3113.3 (19)
Sichuan	10700.3 (1)	10813.4 (1)	10886.8 (1)	10942.9 (1)	11022.4 (1)	11084.3 (1)	11162.9 (1)	11238.2 (1)	11307.6 (1)	11375.4 (1)
Tianjin	856.9 (26)	870.5 (26)	876.6 (26)	882.7 (26)	889.6 (26)	894.5 (26)	898.6 (26)	902.4 (26)	905.1 (26)	910.7 (26)
Tibet = Xizang	215.9 (30)	218.1 (30)	221.8 (30)	225.3 (30)	228.9 (30)	232.0 (30)	235.6 (30)	239.3 (30)	242.7 (30)	245.4 (30)
Xinjiang Uygur	1454.2 (23)	1498.7 (23)	1528.0 (23)	1554.1 (23)	1577.1 (23)	1605.3 (23)	1637.3 (23)	1675.6 (23)	1705.8 (23)	1733.6 (23)
Yunnan	3642.2 (14)	3694.5 (14)	3734.7 (14)	3767.1 (14)	3802.1 (14)	3837.1 (14)	3873.5 (14)	3909.4 (14)	3944.6 (14)	3983.3 (14)
Zhejiang	4208.9 (10)	4234.9 (11)	4261.4 (11)	4285.9 (11)	4313.3 (11)	4341.2 (11)	4369.6 (11)	4400.1 (11)	4422.3 (11)	4446.9 (11)
Total	110356.1	112954.3	114190.8	114342.6	116296.6	117353.7	118467.6	119548.1	120582.3	121497.2