

Bibliometric Study of Literature on Bibliometrics

Swapan Kumar Patra, Partha Bhattacharya, & Neera Verma

Abstract

This paper analyses growth pattern, core journals and authors' distribution in the field of bibliometrics using data from *Library And Information Science Abstracts (LISA)*. Growth of literature does not show any definite pattern. Bradford's law of scattering is used to identify core journals and determines 'Scientometrics' as the core journals in this field. Lotka's law was used to identify authors' productivity patterns. It is observed that authors' distributions do not follow original Lotka's law. Study also identified 12 most productive authors with more than 20 publications in this field.

1. INTRODUCTION

The terms bibliometrics and scientometrics have been introduced simultaneously by Pritchard, Nalimov and Mulchenko in 1969. Pritchard defined the term 'Bibliometrics' as 'the application of mathematical and statistical methods to books and other communication medium'. Nalimov and Mulchenko defined 'Scientometrics' as 'the application of those quantitative methods which are dealing with the analysis of science viewed as an information process'. So, scientometrics is the measurement of science communication, and bibliometrics deals with more general information processes. Although, famous Bradford's law (1934) of scattering, Lotka's law (1926) of scientific productivity are regarded as milestones in bibliometrics, but bibliometrics/scientometrics research actually started in late sixties. Later in the seventies and eighties, bibliometrics research took a distinct shape and emerged as a prominent discipline. Major boost to the scientometrics research was with the publication of the journal 'Scientometrics' in late seventies particularly devoted to bibliometrics/scientometrics. With the advent of information and communication technology (ICT), web technology and availability of different

databases online, the field of bibliometrics gain a momentum. Increasing CPU speed and online availability of various databases makes bibliometrics research much easier and no longer a manual task. This study aims to find out the growth pattern, core journals, authorship pattern and productive authors in this field.

2. METHODOLOGY

In this work, the data for analysis was downloaded from *LISA (Library And Information Science Abstracts)* produced by Cambridge Scientific Abstracts (CSA). *LISA* is an international abstracting and indexing service specially designed for library and information professionals. It covers about 440 periodicals from more than 68 countries and in 20 different languages. It started indexing since 1969 and updated every two weeks. Its subject coverage includes all aspect of Library and Information Science (LIS) including artificial intelligence, book reviews, CD-ROMs, computer science applications, information centres, information management, information science, information storage and retrieval, information technology, internet technology, knowledge management, librarianship, libraries and archives, library

management, library technology, library use and users study, medical informatics, online information retrieval, publishing and book selling, records management, telecommunications, technical services, world wide web. The search terms used were "Bibliometrics" OR "Scientometrics" OR "Webometrics" for retrieval of records. Up to September 2005, 3781 records were retrieved for the bibliometric analysis.

3. DATA ANALYSIS

The retrieved data contains mostly journal articles. Only 53 conference proceedings were found. These records are further analysed using Microsoft Excel and Access for getting further bibliometrics indicators. The results shows that there is no definite pattern of literature growth in the field of bibliometrics. *Scientometrics* is the core journal, which covers mostly (41%) of the total literature coverage. Author with single publication is more predominant, about 77 % authors have only one publication and English is the predominant language of publication.

3.1 Growth of Literature

The *LISA* started indexing Bibliometrics literature in 1968. After that there was a growth of literature (Fig. 1). The highest literature growth occurs in 1999 where 208 records were indexed. In 2005 till September 106 records has been indexed. For the last

few years' trend it can be concluded that literature on bibliometrics has no definite growth pattern. As usual in other subject areas, for the first few years the growth of literature was exponential (1972-78, 1996-98) after that growth is declining. The exponential growth of literature in seventies may be due to the emergence of bibliometrics as a new discipline. Usually a new subject field produce a huge amount of literature at the initial stage of its growth causes exponential literature growth. In nineties the exponential literature growth was due to the advent of computer and information technology and particularly the web technology. With the increasing CPU speed and availability of databases in various formats foster this growth of literature. However, by observing the last few years' trend it can be concluded that about 175 reference per year produced in the area of bibliometrics/scientometrics.

3.2 Language of Publications

It is also important to note the language of publication. As usual English is the predominant language of publications (Fig. 2). Out of the 3781 records retrieved up to September 2005, English occupies the first position with 3110 records (82.25%). Russian language is occupying second position with 172 records (about 4.54 %). Also 85 records found in other languages. English is the dominant language in Bibliometrics because

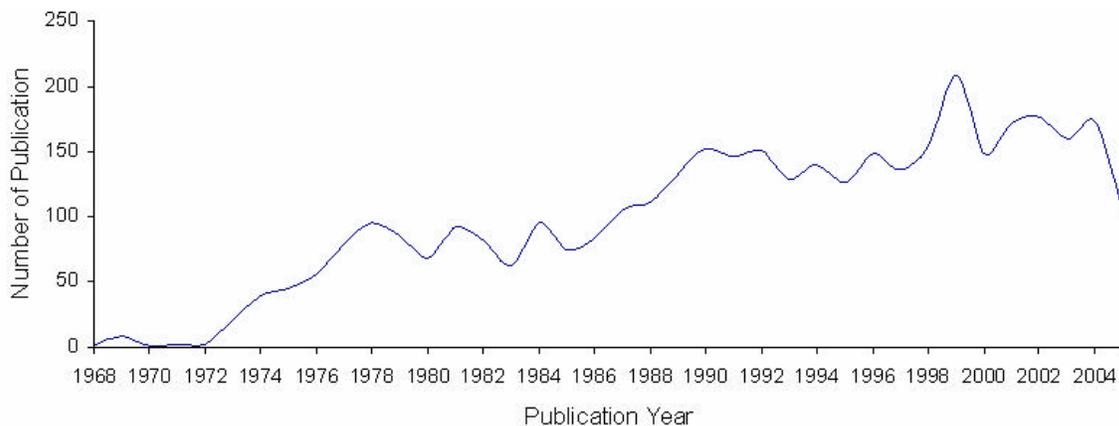


Fig 1. Growth of literature

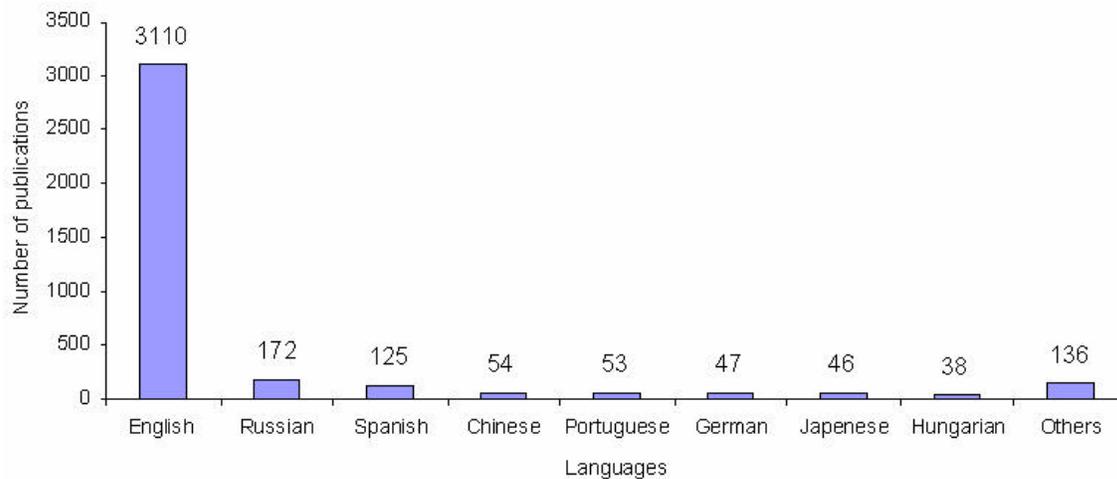


Fig 2. Language of publications

English is official language in many countries and many conference proceedings are published in English language only.

3.3 Core Journals

From the data analysis, it has been found that about 280 journals publish 3781 articles. *Scientometrics* is in the top with 1571 articles, which is about 41.54% of total publications, followed by *Journal of American Society of Information Science and Technology*, which published 197 articles (5.21%) as shown in table 1.

Bradford's law³ has been employed to identify the core journals in a given field. Bradford's law reveals a pattern of how literature in a subject is distributed in journals. According to Bradford's law of scattering, 'If

scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several other groups of zones containing the same number of articles as the nucleus'. Bradford's law is useful for the librarians because it helps in identifying the core sets of journals, which publish the most contents of a given field. So, in modern day financial crunches, a typical Bradford analysis suggests which journals will be included in a library collection⁴. Fig. 3 shows the typical Bradford curve where journals are plotted against their productivity. *Scientometrics* is identified as the core journal, which publishes most of the content because it is fully devoted to bibliometrics/scientometrics. The top four journals contain 50% of the research output.

Table 1. Core journals with their number, percentage and publishers

Name of the journal	No. of publications	% of publications	Publisher
Scientometrics	1571	41.54	SpringerLink
Journal of the American Society for Information Science & Technology	197	5.21	John Wiley & Sons
Nauchno Tekhnicheskaya Informatsiya	107	2.82	Moskova
Revista Espanola de Documentacion Cientifica	95	2.51	Madrid, Centro Nacional de Informaci3n y Documentaci3n Cientifica

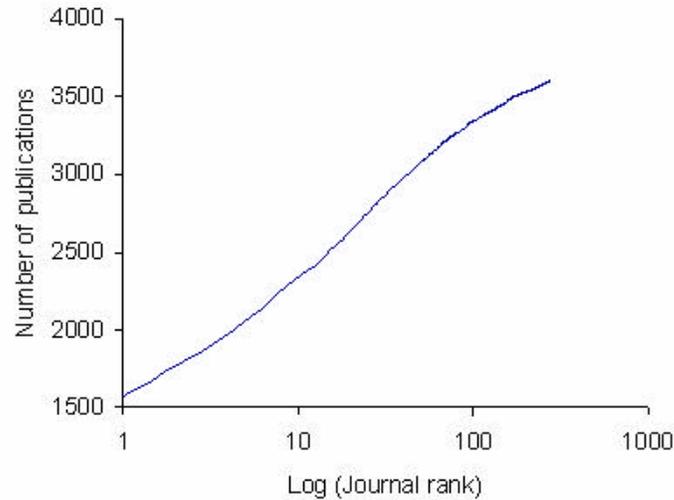


Fig 3. Bradford curve showing journal rank and number of publications

These journals are *Scientometrics*, *Journal of the American Society for Information Science & Technology*, *Nauchno Tekhnicheskaya Informatsiya*, *Revista Espanola de Documentacion Cientifica*. These journals may be regarded as the important sets of journals in the field.

3.4 Authorship Pattern

From the data, it has been found that about 4,000 authors publish 3,781 articles, which is 0.94 articles per author. It means single authorship is very common in this field. About 3,106 (77.65%) authors have only one publication and 470 (11.75%) authors have two publications. Table 2 shows the number of authors and their corresponding publications. Lotka's Law⁵, an inverse, square law, is used to find authors productivity patterns. It states that for every 100 authors contributing one article, 25 will contribute 2, 11 will contribute 3, and 6 will contribute 4 each. There is general decrease in performance among a body of authors following $1:n^2$. This ratio shows that some produce much more than the average. According to Lotka's law of scientific productivity, only six percent of the authors in a field will produce more than 10 articles. The general form of Lotka's law⁵ can be expressed as:

$$y=c/x^n$$

Using this formula and the modifications given by Pao, Fang^{6,7,8} the value of C for bibliometrics literature is determined to be 0.64 and $n=2.09$. Kolmogorov-Smirnov Goodness-of-Fit Test shows that bibliometrics literature does not follow Lotka's law. Figure 4 shows the Log-Log plot of authors' number and their corresponding publications.

Table 2. Authorship pattern

No. of publications	No. of authors	% of authors
1	3106	77.65
2	470	11.75
3	172	4.30
4	69	1.72
5	48	1.20
6	32	0.80
7	26	0.65
8	15	0.37
9	12	0.30
10	9	0.22
11	5	0.12
12	5	0.12
13	6	0.15
14	3	0.07
15	2	0.05
16	3	0.07
17	4	0.10
19	1	0.02

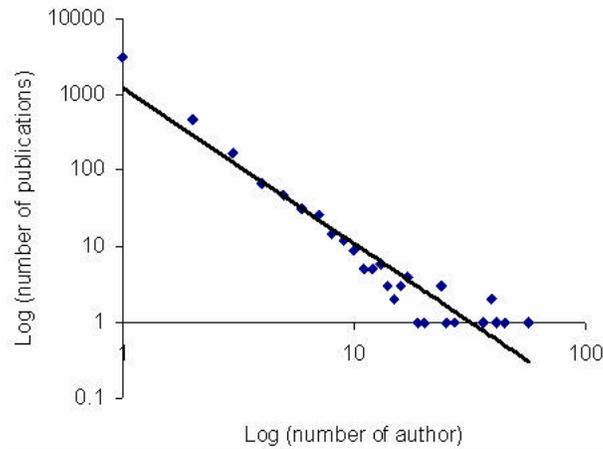


Fig 4. Log-Log plots of number of author and the number of publications

20	1	0.02	Braun, T.	43	1.13	1982-2005
24	3	0.07	Gupta, B. M.	41	1.08	1977-2005
25	1	0.02	Moed, H. F.	34	0.89	1986-2005
27	1	0.02	Leydesdorff, L.	30	0.89	1986-2004
36	1	0.02	Vinkler, P.	27	0.71	1986-2004
39	2	0.05	Van Raan, A.F. J.	26	0.68	1986-2005
41	1	0.02	Garg, K. C.	25	0.66	1985-2003
45	1	0.02	Sengupta, I.N.	20	0.52	1973-1992
57	1	0.02	Total	466	12.32	
Total	4000	100				

Table 3 shows the 12 most productive authors identified with more than 10 publications. L. Egghe is the most productive author in this field with 62 publications, which accounts about 1.63% of the total publications. Although India's contribution in the field of bibliometrics is 288 (7.61%), three Indian authors B.M. Gupta, K.C. Garg and I.N. Sengupta are ranked in the 3rd, 10th and 12th positions with 41, 24, 20 publications. Top 12 authors account about 10% of total publications.

Table 3. Most productive authors

Author	No. of public-ations	% of public-ations	Productive yrs
Egghe, L.	62	1.63	1985-2005
Glanzel, W.	60	1.58	1983-2005
Rousseau, R.	53	1.40	1986-2005
Schubert, A.	45	1.19	1982-2005

4. CONCLUSION

Bibliometrics is an important field of information science because it represents a unique set of techniques for the monitoring and analysis of information resources and for the management of knowledge in social and organisational contexts. Bibliometric methods are used in studies of properties and behaviour of recorded knowledge, for analysis of the structures of scientific and research areas, and for evaluation of research activity and administration of scientific information. Various statistical methods are applied to study to measure, authorship, citation and publication pattern, and the relationship within scientific domains and research communities and to structure of specific fields. In this sense, bibliometrics is also relevant for researchers, policy and decision makers and also researchers outside the library and information science (LIS) field

to track the trend in the specific field in their research work. Moreover, bibliometrics studies should be encouraged to evaluate research performance of a particular field of research in a country. Even national science policy can be decided on the basis of bibliometrics/scientometrics study. It is expected that more and more subject experts would take keen interest in this area of study.

REFERENCES

1. Pritchard, A. Statistical bibliography or bibliometrics? *Journal of Documentation*, 1969, **24**, 348-49.
2. Nalimov, V.V. & Mulchenko, B.M. *Scientometrics*, Nauka, Moscow, 1969 (in Russian).
3. Bradford, S.C. Sources of information on specific subjects. *Engineering: An Illustrated Weekly*, 1934, **3550**, 85-86.
4. Ungern-Sternberg, S.V. Bradford's law in context of information provision. *Scientometrics*, 2000, **49**(1), 161-86.
5. Lotka, A.J. The frequency distribution of scientific productivity, *Journal of the Washington Academy of Science*, 1926, **16**, 317-23.
6. Nichols, P.T. Empirical validation of Lotka's law. *Information Processing and Management*, 1986, **22**(5), 417-19.
7. Pao, M.L. Lotka's law: A testing procedure. *Information Processing and Management*, 1985, **21**(4), 305-20.
8. Fang, P.H. & Fang, J.M. A modification of Lotka's function for scientific productivity *Information Processing and Management*, 1995, **31**(1), 133-37.

Contributors: **Sh. Swapan Kumar Patra**, Department of Biochemical Engineering & Biotechnology, Indian Institute of Technology Delhi, New Delhi-110016. e-mail: skpatra@gmail.com

Sh. Partha Bhattacharya, FITT, Indian Institute of Technology Delhi, New Delhi-110016. e-mail: parthab@fitt.iitd.ac.in

Ms. Neera Verma, Department of Biochemical Engineering & Biotechnology, Indian Institute of Technology Delhi, New Delhi-110016. e-mail: neeraverma@hotmail.com