INDIAN JOURNAL OF BIOTECHNOLOGY (IJBT): A SCIENTOMETRIC ANALYSIS

R.Lakshmi Sankari

Research Scholar and Assistant Librarian,
Vinayaka Mission's Engineering College, Salem
E-mail: r.lakshmisankari@gmail.com

Dr.K.Chinnasamy

Associate Professor,
DLIS, Madurai Kamaraj University, Madurai
E-mail: drkchinna@yahoo.co.in

ABSTRACT

Scientometric is a type of research method used in library and information science. "The study of the use of documents and patterns of publication in which mathematical and statistical methods have been applied". It utilizes quantitative analysis and statistics to describe patterns of publication within a given field. The study presents the result of detailed analysis from the data collected by applying the Bibliometric techniques. The Scientometric analysis of 436 articles / papers published in the Indian Journal of Biotechnology during the years 2004-2008 are taken up and analysed based on various parameters such as year-wise distribution of papers, authorship productivity, geographical distributions of contribution etc. Applicability of Lotka's law is validated from the values n = 3.15. The numbers of contributions form Maharashtra at the National level and from India at the International level is significant. The study reveals that the average number of contribution per volume has come around more than 80.

Keywords: Scientometric, Growth, Collaboration Measures, Author Productivity, Lotka's Law, Indian Journal of Biotechnology (IJBT).

INTRODUCTION

Scientometrics is a branch of science that describes the output traits in terms of organizational research structure, resource inputs and outputs, develops benchmarks to evaluate the quality of information output. Scientometric studies characterize the disciplines using the growth pattern and other attributes. These studies have potential particularly in assessing the emerging disciplines. Scientometrics is one of the most important measures for the assessment of scientific productions. Macias-Chapula argues that "scientometric indicators have become essential to the scientific community to estimate the state-of-the-art of a given topic" (quoted In Lolis et. al. 2009). Scientometrics is related to and has overlapping interests with Bibliometrics and Informetrics. Scientometrics involves the application of quantitative methods to the history of science. It is the science of measuring science; the measurement involves counting artifacts to the production and use of information, and arriving conclusions from the counts. Sangam et al

(2010) say that the changes in the size of literature over specific period may be termed as growth of literature. In the information era science is becoming increasingly interdisciplinary and problem oriented, often requiring the coming together of a group of people who complement each other in terms of function and expertise which termed as collaboration (Subramanyam, 1983). The conclusions are drawn on the basis of the regularities that are revealed in the data.

DEFINITIONS

The term"Scientometrics" in the literature; Scientometrics is the quantitative study of the disciplines of science based on published literature and communication. This could include identifying emerging areas of scientific research, examining the development of research over time, or geographic and organizational distributions of research (Glossary of Thompson…, 2008).

Tague-Sutcliffe (1992) defines Scientometrics as "the study of the quantitative aspects of science as a discipline or economic activity. It is part of the sociology of science and has application to science policy-making. It involves quantitative studies of scientific activities, including, among others, publication, and so overlaps bibliometrics to some extent".

REVIEW OF LITERATURE

In recent years, many researchers have conducted scientometric analysis in different subject fields**.** Osareh & Wilson (2002) analyzed international collaboration of Iranian scientific publications in Science Citation Index (SCI) during 1995 to 1999. One result of this study shows that Iran's publication outputs in science and technology increased dramatically in the SCI during the study period. Another result shows that Iran's main international collaborations are by authors with institutional affiliations in the U.S.A or UK. However, it is obvious that Iran is looking more and more for collaborating partners elsewhere. Collaboration with authors in Canadian and Australian institutions has increased either in absolute numbers, relative percentages or both.

Mukherjee (2008) analyzed the authorship pattern of scientific productions of the four most productive Indian academic institutions for the eight-year –period from 2000 to 2007. The results show that among four universities, the authors of Delhi University contributed the highest number of articles, followed by Banaras Hindu University. There is also an increasing tendency toward collaborative research among Indian authors as well as more frequent collaboration with international authors. Biochemistry and Molecular Biology are two of the most prolific research areas in these four Indian universities. The average rate of references per item is 28 and the citations received per item are 3.56.

Rai and Kumar (2005) investigated literature in bioinformatics using a number of bibliometric techniques including Lotka's law. Their data set was drawn from the PubMed database and analyzed to identify the core journals in the field. An exact timeframe was not clearly stated for the 16, 471 records retrieved from this database and examined, as the researchers only provide an end date of January 2005. One can assume that the data set includes all records from the founding of the PubMed database to present. Applying Bradford's law, the researchers

established that there are at least 20 core journals in this growing field. Only the first author was counted for co-authored publications accounting for a total of 39,435 authors. They calculated that 29,008 authors published only one article (73.58%) and approximately 20% of remaining authors published two to three articles.

According to their findings, authors in the bioinformatics field seem to be more on the productive side with 23% of the authors producing one article instead of the predicted 60%. Only 5% proved to be highly productive having published more than ten articles. According to the researchers a Kolmorogov-Smirnov goodness-of-fit test indicated that the authors' productivity pattern followed Lotka's law. Although the researchers assert that their calculated *c* and *n* values (0.78 and 2.69 respectively) fit the values found by Lotka's study, this assertion is somewhat exaggerated since Lotka's calculated *c* and *n* values for his data sets were 0.60 and approximately 2.0.

SOURCE JOURNAL

Indian Journal of Biotechnology (IJBT) is a quarterly publication. It is published by the National Institute of Science Communication and Research (NISCAIR)**,** New Delhi since 2002; it is published in the months of January, April, July and October. The journal covers the research, review papers and short communications. The latest developments in biotech-industry are covered under Notes and News. The major subject fields covered in IJBT include Agriculture, Animal Husbandry, Industry, Microbiology, Medicine, Bio-informatics and Socio-legal etc. The study pertains to the data collected from the journal for the period of 5 years from 2004 to 2008.

OBJECTIVES OF THE STUDY

The present study has been undertaken**:**

1. To find out volume wise distribution & average number of contributions per volume.
2. To examine the growth of contributions on Indian Journal of Biotechnology published during the period 2004-2008.
3. To study Authorship pattern of the contribution.
4. To investigate the collaborative research trend in terms of Collaborative Index (CI); Degree of Collaboration (DC) and Collaborative Co-efficient.
5. To examine the validity of Lotka's law using productivity of authorship and to undertake K-S statistics for the conformity of the results obtained by these methods.
6. To study the Geographical distributions of contributions

SCOPE OF THIS STUDY:

An attempt has been made to analyse the contributions in 20 issues of 5 volumes of the Indian Journal of Biotechnology (IJBT) spanning between 2004 and 2008, covering a period of five years.

Table 1 - Distribution of Contributions (Volume –wise)

MONTH	VOLUME NUMBER					Total
	3	4	5	6	7	
Jan	23	22	21	21	21	108
April	20	21	22	22	21	106
July	22	21	27	22	21	113
Oct	23	22	20	22	22	109
Total	**88**	**86**	**90**	**87**	**85**	**436**
%	**20.18**	**19.72**	**20.64**	**19.95**	**19.50**	**100.00**

As indicated in Table 1 the total number of contributions in 20 issues of 5 volumes of the journal is 436, of which 20.18% were contributed in Vol.3, 19.72% of them were published in Vol.4, 20.64% of them were published in Vol.5, 19.95% of them were published in Vol.6, and 19.50% of them were published inVol.7. A notable attribute of the study is that the Vol.5 shows the maximum number of contributions.

GROWTH RATE ANALYSIS

Relative Growth Rate

The relative growth rate is the increase in the number of publications/pages per unit of time. Here, one year is taken as the unit of time. The mean relative growth rate R (1-2) over a specified period of interval can be calculated from the following equation suggested by Mahapatra.

$$R(1\text{-}2) = W_2 - W_1 / T_2 - T_1$$

Where R-(1-2) = mean relative growth rate over the specified period of interval;
$W_1 = \log w_1$ (Natural log of initial number of publications/ pages);
$W_2 = \log w_2$ (Natural log of initial number of publications/pages);
T_2-T_1 = The unit difference between the initial time and final time.
Therefore,
R (a) = Relative growth rate per unit of publications per unit of time (year)
R (P) = Relative growth rate per unit of pages per unit of time (year)

Doubling Time

A direct equivalence exists between the relative growth rate and doubling time. If the number of publications/pages of a subject doubles during a given period, then the difference between the logarithms of the numbers at the beginning and at the end of the period must be the logarithms of the number 2. This difference has a value of 0.693. Thus, the corresponding doubling time for publication and pages can be calculated by the following formula:

Doubling time (Dt) = 0.693 / R
Therefore,
Doubling time for publications Dt (a) = 0.693 / R(a)
Doubling time for pages Dt(p) = 0.693 / R(p)

Table 2 -Relative Growth Rate and Doubling Time for Indian Journal of Biotechnology

Years	No.of Articles	Cumulative number of Articles	w1	w2	R(a)	Mean R(a)	Doubling time Dt(a)	Mean Doubling time
2004	88	88	0.00	4.48	0.00	0.37		1.34
2005	86	174	4.48	5.16	0.68		1.02	
2006	90	264	5.16	5.58	0.42		1.67	
2007	87	351	5.58	5.86	0.29	0.25	2.43	2.81
2008	85	436	5.86	6.08	0.22		3.19	
			6.08			0.31		2.08

Relative Growth Rate and Doubling Time for Indian Journal of Biotechnology

Table 2 indicates the relative growth rate of total output and also the doubling time for publications. It could be observed that the relative growth rates for all sources of output have decreased from 0.68 in 2005, 0.22 in 2008. The mean relative growth rates for the periods 2004 to 2006 and 2007 to 2008 are 0.37 and 0.25 respectively, whereas the whole study period has witnessed a mean relative growth rate of 0.31. The mean doubling time for the above two periods are 1.34 and 2.81 respectively. The whole study period has witnessed a doubling for total contribution at 2.08. In general the relative growth rate of publication output has shown a declining trend, whereas a doubling time for publication has shown increasing

Table 3: Cumulative Distribution of Authorship Patterns

Year	One	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten	Ten+	Total
2004	1	30	22	19	9	5	0	2	0	0	0	88
2005	5	30	20	17	8	2	2	0	1	1	0	86
2006	4	31	25	15	6	5	2	2	0	0	0	90
2007	5	19	22	22	10	6	1	2	0	0	0	87
2008	2	28	22	15	8	6	3	0	0	0	1	85
Total	**17**	**138**	**111**	**88**	**41**	**24**	**8**	**6**	**1**	**1**	**1**	**436**
%	**3.9**	**31.7**	**25.5**	**20.2**	**9.4**	**5.5**	**1.8**	**1.4**	**0.2**	**0.2**	**0.2**	**100.0**

The analysis of the data showed that the maximum number of authors, who contributed to the journal had a tendency to work in collaboration. From table 3 it can be see that the single

authored articles were the minimum and two authored articles the maximum with 138 (31.7%). A significant not of the study is that the majority of articles are contributed by two-authors.

COLLABORATIVE MEASURES

a) Collaborative Coefficient (CC)

Collaborative coefficient (CC) suggested by Ajiferuke and used by Karki and Garg has been used to measure the extent and strength of collaboration among the IEEE Transaction on Software Engineering. It can be expressed mathematically as:

$$CC = 1 - \sum_{j=1}^{j=k} (1/J)\, f_j/N$$

Where

- f_j is the number of J authored papers published in a discipline during a certain period of time.
- N is the total number of research papers published in a discipline during a certain period of time and
- k is the greatest number of authors per paper in a discipline

b) Degree of Collaboration (DC)

The determine the degree of collaboration in biotechnology, the following formula given by Subramanyam (1983) has been used;

$$\mathbf{DC} = \frac{\mathbf{NM}}{\mathbf{NM} + \mathbf{NS}}$$

Where DC = Degree of Collaboration
NM = Number of Multi authored papers
NS = Number of single authored papers.

c) Collaborative Index

Collaborative Index can be obtained by total number of authors divided by total number of published articles.

$$CI = \frac{\text{Total number of Authors}}{\text{Total number of Articles}}$$

Where,
CI = It is the number of authors per paper**.**

Table -4 Authorship pattern and collaborative Measures

Year	TP	TA	CI	CC	DC
2004	88	294	0.30	0.65	0.99
2005	86	278	0.31	0.61	0.94
2006	90	291	0.31	0.62	0.96
2007	87	306	0.28	0.63	0.94
2008	85	295	0.29	0.64	0.98
Total	436	1464	0.30	0.63	0.96

TP=Total Papers; TA=Total Authors; CI=Collaborative Index; CC=Collaborative Co-efficient and DC=Degree of Collaboration.

Table 4 shows that Authorship pattern and collaborative measures. The collaborative Index for the year 2004-2008 was 0.30 which show that collaborative; the value of Collaborative Co-efficient (CC) have shown higher (CC) value with more than 0.50, which show greater probability of multiple- authorship and the degree of collaboration (DC) for the year 2000-2005 was 0.96.

AUTHOR PRODUCTIVITY AND LOTKA'S LAW

Author productivity

The term author productivity, scientific productivity, publication productivity and trends of publications are used synonymously. Regarding the author productivity one can say that, author productivity means authors productiveness or author‘s efficiency in publication. In other words author productivity can be explained as the effectiveness of productive efforts to produce fruitful publication.

Table 5- Distribution of Authorship Productivity

No. of Articles	No of authors Observed	Percentage of Authors
1	1002	84.49
2	136	11.47
3	25	2.11
4	14	1.18
5	4	0.34
6	2	0.17
7	1	0.08
8	1	0.08
12	1	0.08
Total	1186	100.00

Table 5 shows frequency distribution of author productivity in the field of Indian journal of Biotechnology. Of the 1186 unique authors names, 1002 (84.49%) produced one article, 136 (11.47 %) produced two articles and so forth. The number of authors who produced more than 12 articles is quite small (only 0.08 %).

LOTKA'S LAW

Lotka's law is a classical method used to test the regularity in the publication activity of authors of scientific literature. It describes the frequency of publication by authors in a given field. It states that the number of authors making *n* contributions is about $1/n^2$ of those making one; and the proportion of all contributors that make a single contribution is in the region of 60 per cent. This means that out of all the authors in a given field, 60 per cent will have just one publication; 15 percent will have two publications ($1/2^2$ times 60); 7 per cent will have three publications ($1/3^2$ times 60), and so on.

This law can be expressed as:

$$y = C \times x^{-n} \qquad (1)$$

Where *x* is the number of publications of interest (1, 2, etc.); *n* is an exponent that is constant for a given set of data; *y* is the expected percentage of authors with frequency *x* of publications, and *C* is a constant. The productivity corresponds not to the number of articles published by an author but to its logarithm; it seems that a multiplicative, rather than simply additive, model provides a better fit to this measure or counting method.

The exponent *n* is often fixed at 2, in which case the law is known as the *inverse square law of scientific productivity.* However, given that the exponent *n* predicts the relative number of authors at each productivity level it would seem useful to calculate it. In the present study, least square method has been used. It can be expressed as:

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} \qquad (2)$$

Where N is the number of data pairs considered
X is the logarithm of x (x=number of articles) and
Y is the logarithm of y (y=number of authors)
The constant C is calculated using formula:

$$C = \frac{1}{\sum 1/x^n} \qquad (3)$$

To verify that the observed distribution of author productivity fits the estimated distribution, Pao (1985) suggested applying the non-parametric Kolmolgorov-Smirnov (K-S) goodness-of fit test. To this end, the maximum difference between the real and estimated accumulated frequencies was calculated and these values were then compared with the critical values (c.v) obtained from the following equation:

$$c.v \frac{1.63}{\left(\sum yx + \left(\sum yx/10 \right)^{1/2} \right)^{1/2}} \quad (4)$$

Table 6 Author productivity Using Lotka's law

x	y	X	Y	x^2	xy	Yx/∑yx	∑(yx/∑yx)	$1/x^n$	$fe=C(1/x^n)$	∑fe	D
1	1002	0	3.001	0	0	0.845	0.845	1	0.855	0.855	**0.01**
2	136	0.301	2.134	0.091	0.642	0.115	0.96	0.113	0.096	0.951	0.008
3	25	0.477	1.398	0.228	0.667	0.021	0.981	0.031	0.027	0.978	0.002
4	14	0.602	1.146	0.362	0.69	0.012	0.992	0.013	0.011	0.989	0.003
5	4	0.699	0.602	0.489	0.421	0.003	0.996	0.006	0.005	0.994	0.001
6	2	0.778	0.301	0.606	0.234	0.002	0.997	0.004	0.003	0.997	0
7	1	0.845	0	0.714	0	0.001	0.998	0.002	0.002	0.999	0.001
8	1	0.903	0	0.816	0	0.001	0.999	0.001	0.001	1.001	0.001
12	1	1.079	0	1.165	0	0.001	1	0	0	1.001	0.001
	1186	5.685	8.582	4.469	2.654	1	7.768	1.17			

n= 3.15, c=0.855, D = 0.010

To validate the Lotka's law a calculation was done using Eqns (1-4), (Table-6), to know the values of n and C to test whether application of Lotka's law fits or not. Thus, based on the data presented in Table 6, the calculated values of n and C are 3.15 and 0.855, respectively. In order to verify that the observed distribution of the productivity of the authors fits the theoretical distribution, we subjected the data to the non-parametric Kolmolgorov-Smirnov test. To this end, we used the data in the last column of Table 6 (Dmax), obtained as the absolute value of the difference between columns 8 and 11 of the same table. The greatest value of this column (Dmax) will be taken as reference for comparison with the "critical value" (c.v), obtained by the asymptotic formula (5)

The critical value is 0.047 and the value of Maximum Difference (D) between the real and estimated accumulated frequencies is 0.010, which is less than the critical value 0.047. This resulted fitting the application of Lotka's law to the data of Indian Journal of Biotechnology.

Table-7 Geographical Distribution (Indian and Foreign Contributions)

Sl.No	Name of State	Contributions	Cumulative no of Contributions	% of Contributions	Cumulative % of Contribution
1	Maharashtra	61	61	13.99	13.99
2	Tamil Nadu	52	113	11.93	25.9
3	Andhra Pradesh	33	146	7.57	33.5
4	Uttar Pradesh	32	178	7.34	40.8
5	Karnataka	25	203	5.73	46.6
6	Rajasthan	23	226	5.28	51.8
7	West Bengal	22	248	5.05	56.9
8	New Delhi	21	269	4.82	61.7
9	Haryana	19	288	4.36	66.1
10	Kerala	16	304	3.67	69.7
11	Uttarakhand	13	317	2.98	72.7
12	Madhya Pradesh	13	330	2.98	75.7
13	Himachal Pradesh	12	342	2.75	78.4
14	Gujarat	11	353	2.52	81
15	Punjab	10	363	2.29	83.3
16	Assam	8	371	1.83	85.1
17	Andhaman & Nicobar	5	376	1.15	86.2
18	Meghalaya	5	381	1.15	87.4
19	Arunachal Pradesh	3	384	0.69	88.1
20	Orissa	3	387	0.69	88.8
21	Nagaland	2	389	0.46	89.2
22	Pondicherry	2	391	0.46	89.7
23	Goa	2	393	0.46	90.1
24	Jammu & Kashmir	2	395	0.46	90.6
25	Manipur	2	397	0.46	91.1
26	Manipal	1	398	0.23	91.3
27	Mizoram	1	399	0.23	91.5
28	Sikkim	1	400	0.23	91.7
29	Foreign	36	436	8.26	100

The geographical distribution of contributions (National and International) to the journal is presented in Tables 6 Out of the total 436 contributions in the five volumes of the journals under study, 400 contributions have been made by Indian states. Maharashtra has a maximum contribution of 61 (15.25%) which is followed by Tamil Nadu, Andhra Pradesh and Uttar Pradesh. From the above analysis India has highest number of contributions compared to foreign.

FINDINGS

The following are the major findings of the study:

1. The maximum number of papers 90 were published in Volume No 5 and minimum of 85 in Volume No 7.

2. The value of an average RGR of article (Rt (P)) gradually decreasing and the values of Doubling time of the articles Dt(P) gradually increasing.

3. Authorship pattern of Indian Journal of Biotechnology (IJBT) for the year 2004-2008, the majority of articles are contributed by two authors. 138 (31.7 %) .

4. The collaborative Index for the year 2004-2008 was 0.30,; the value of Collaborative Co-efficient (CC) with more than 0.50, which show greater probability of multiple-authorship and the degree of collaboration (DC) for the year 2000-2005 was 0.96.

5. Author Productivity was calculated and that 1002 (84.49%) authors contributed one article, and 12 (0.08%) authors contributed more than 12 articles.

6. The productivity of the authors does fit a lotka's distribution (Dmax=0.010 and critical value = 0.047).

7. The national contributions are maximum in the journal. Among the Indian states Maharashtra 61(13.99 %) tops the list of contributors and among International contribution Indians are the top runner researchers in the present day scenario.

CONCLUSION

The publishing trend totally depends on the productivity of contributors, pattern of contributions and the quality of research. The Indian journal of Biotechnology is one such reputed journal, which is published in India since 2002. The study presents a detailed Scientrometric analysis of papers published in the IJB from 2004-2008. The study reveals that the average number of contribution per volume has come around more than 80. The number of contributions from Maharashtra at the National level and from India at the International level is significants. When the author productivity was calculated it was found that 84.49 % authors contribute single article. The study reveals that the value of an average RGR of article (Rt (P)) gradually decreasing and the values of Doubling time of the articles Dt(P) gradually increasing. Lotka's law was tested and found to fit the data.

REFERENCE

1. Ajuferuke, I., Burell, Q. and Tague, J. (1988): Collaborative coefficients: A single measure of the degree of collaboration in research, Scientometrics, 14(5-6): 421-433.

2. Glossary of Thompson scientific terminology. (2008). The Thompson Corporation. Retrieved from http://science.thomsonreuters.com/support/patents/patinf/terms/

3. Lolis, S. F. & et. al. (2009). Scientometric analysis of energetic ecology: Primary production of aquatic macrophytes. *Maringá, 31* (4), 363-369.

4. Lotka,A.J. (1926) The frequency distribution of scientific productivity. *Journal of the Washington Academy of Science,* 16 , 317-23.

5. Mahapatra, M. (1985) On the Validity of the Theory of Exponential Growth of Scientific Literature: Proceedings of the 15th IASLIC Conference, Bangalore, p.61-70.

6. Mukherjee, B. (2008). Scholarly literature from selected universities of Delhi and Uttar Pradesh: A pilot study. LIBRES, *18* (1). Retrieved from http://libres.curtin.edu.au/ libres18n1/Mukherjee_Final_rev.pdf

7. Osareh, F., & Wilson, C.S. (2002). Collaboration in Iranian scientific publications. *Libri, 52*, 88-98.

8. Rai, L. P., & Kumar, N. (2005) A rationale for Lotka's law: An examination of empirical data. *Annals of Library and Information Studies, 52* , 103-107.

9. Sangam, S. L., Kumbar, B. D., Naik, R. R., Keshava, Hadagali, G. S. And Krishnamurthy, C. (2010) : Webmetrics, Informetrics and Scientometrics – Measuring Scientific and technological Progress of India, National Seminar Papers and Proceedings, 21-22 December, 2009, pp. 1-239.

10. Subranmanyam, K. (1983): Bibliometric studies of research in collaboration: A review, Journal of Information Science, 6(37).

11. Tague-Sutcliffe, J. M. (1992). An introduction to Informetrics. *Information Processing & Management*, *28*, 1–3.