LOTKA'S LAW: A STUDY WITH REFERENCE TO THE LITERATURE BY FACUTLITIES OF NATIONAL INSTITUTES OF TECHNOLOGY, INDIA

Dr.N.TAMILSELVAN

Chief Librarian & Head, Rathinam Technical Campus, Coimbatore

N. SIVAKUMAR

Chief Librarian, Kalaignar Karunanidhi Institute of Technology, Coimbatore

ABSTRACT

Bibliometric analysis of the literature published in the fields of engineering and technology by the faculties of NIT's in India has been carried out. Testing of the validity of Lotka's law has been performed. It has been found that Lotka's law can be satisfactory applied to the literature brought out by the faculties of NIT's.

Keyword: Validity of Lotka's Law: A case study of NIT's in India.

INTRODUCTION

Bibliometrics method has been employed to conduct the research in library and information science as it is especially useful when studying a research intensive subject fields such as Biotechnology, Biochemistry, Microbiology, Engineering and technology where the industry is likely to have close relations with public research organizations like universities.[1] It is a quantitative scientific evaluation of any available literature. Scientific productivity is influenced by a large number of factors, including individual characteristics such as age, gender, psychological traits and educational background, as well as structural features like funding, institutional context, and the organization of the actual research[2-4]. Bibliometric analysis is also used in the science and technology policy developing arena to determine the knowledge outputs of the systems of innovation. Currently, as a developing country, India is concentrating on Engineering and Technology fields. Accordingly the existing Indian Institute of Technology (IITs) are infrastructurally developed. Further new Indian Institute of Technology are dedicated to the nation. National Institute of Technology (NITs) are also paid

an equal attention with respect to their overall growth. In this background an attempt has been made in this study to ascertain the validity of certain laws like Lotka's law concerned with bibliometrics analysis of literature pertaining to the articles published by the faculties of National Institute of Technology in India.

OBJECTIVES OF THE STUDY

The research was undertaken with the following objectives:

1. To identify and analyse the growth rate of literature among the faculties;

2. To examine the pattern of distribution of science research output of faculties;

3. To measure and calculate the relative growth rate and doubling time for publications using exponential growth model;

4. To analyse the type of co-authorship pattern and examine the extent of research collaboration among different subjects of engineering and technology.

5. To assess the institution-wise research concentration in different fields of engineering and technology;

6. To comprehend and test the applicability of Lotka's Law of scientific productivity of authors;

7. To identify the core journals of research output of faculties using the Bradford's law as an indicator for the dispersion of engineering and technology research output;

8. To identify the geographic distribution of journals from which the contributions of engineering and technology faculties get published;

9. To identify the areas of research concentration on different engineering and technology fields among National Institutes of Technology in India; and

10. To suggest a rational strategy for research and development in engineering and technology fields among National Institutes of Technology in India based on the findings of the study.

LOTKA'S LAW

The statistical analysis of scientific literature began almost 50 years before the term 'bibliometrics' was coined. In 1926, Alfred J. Lotka[5] published his pioneering study on the frequency distribution of scientific productivity determined from a decennial index (1907-1916) of Chemical Abstracts. Lotka concluded that:

"the number (of authors) making n contributions is about $1/n^2$ of those

making one; and the proportion of all contributors, that makes a

single contribution, is about 60 per cent."

This result can be considered as a rule of thumb even today, 75 years after its publication.

Lotka's law is a classical method used to test the regularity in the publication activity of authors of scientific literature or 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 range of 60 percent. This means that out of all the authors in a given field, 60 percent will have just one publication; 15 percent will have two publications ($1/2^2$ times 60); 7 percent will have three publications ($1/3^2$ times 60), and so on[6-8]. For the present study, Lotka's Inverse Power Law model that states the function describing the pattern of productivity of authors publishing in a specified subject field in a fixed time period has been applied and it is mathematically represented as:

$$y = C \times x^{-n} \tag{1}$$

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 constant C is calculated using the following formula:

$$C = \frac{1}{\sum 1/x^n} \tag{2}$$

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 follows:

$$n = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2} \qquad (3)$$

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)

In order to verify whether the observed distribution of author productivity fits the estimated distribution, Pao (1985)[9] suggests applying the non-parametric Kolmolgorov-Smirnov[10] (K-S) goodness-of- fit test. To this end the maximum difference between the real and estimated accumulated frequencies is calculated, this value then being compared with the critical value (c.v.) obtained as:

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

Year-wise Distribution of Research Output of faculties in NITs of India

The annual research output of faculties in NITs of India for the period between 2001 and 2010 has been presented in Table 1. The table indicates that the number of publications in 2001 was 401 and it rose to 1176 in 2010 which is nearly threefold increase as compared to the first year. The highest output was observed in 2010 that accounts for14.05 percent of total output over the period of study.

TABLE 1: YEAR-WISE DISTRIBUTION OF RESEARCH OUTPUT OF FACULTIES IN NITS OF INDIA

Year	Output	Percentage	Cumulative Percentage
2001	401	4.79	4.79
2002	523	6.25	11.04
2003	718	8.58	19.61
2004	792	9.46	29.07
2005	693	8.28	37.35
2006	1048	12.52	49.87
2007	924	11.04	60.91
2008	1022	12.21	73.11
2009	1075	12.84	85.95
2010	1176	14.05	100.00
Total	**8372**	**100.00**	

It was observed that a steady growth in terms of productivity was maintained throughout the period of study. However, there was a sudden increase in 2006 and 2008 while there was a declining trend observed in 2005 and 2007.

Authorship Pattern of NITs Literature

The objective of the study of authorship pattern is to bring out research pattern in a discipline. It is a well known fact that now-a-days research is carried out by group of researchers rather than by a single researcher. Through collaboration, researchers share and exchange knowledge and techniques, that bring in a mixture of positive scientific thoughts and decrease cost at the same time. Bibliometricians have paid due attention to these phenomena ever since the beginning. Intensifying co-author relationship has been reported for all fields and practically at all levels of aggregations, for instance, by Glanzel[11] for the

macro level, by Gomez[12]   for the meso level, and Ding[13]   and Glanzel[14] for the micro level. As a result, multi- authorship necessarily increases productivity and always results in high citation impact.

In the light of the above fact, an attempt was made to identify the nature of authorship pattern of literature produced by the NIT faculties in India. This analysis not only enables one to identify the magnitude of research interests among the scientists but also explains their contribution of the number of papers in any given year of the study period.

Table 2 projects the overall analysis of the pattern of authorship and Table 3 indicates the year-wise authorship pattern and its percentage in contributing to research. The authors are classified according to their contribution that they have published. The total contribution published by NIT faculties in India comes to 8372 records over the period of study.

TABLE 2: DISTRIBUTION OF AUTHORSHIP PATTERN

Authorship Pattern	No. of Contribution	Percentage	Cumulative Percentage
1	523	6.25	6.25
2	3122	37.29	43.54
3	2188	26.13	69.68
4	1251	14.94	84.62
5	579	6.92	91.53
6	295	3.52	95.06
7	165	1.97	97.03
8	97	1.16	98.19
9	64	0.76	98.95
>10	88	1.05	100.00
Total	**8372**	**100.00**	

TABLE 3: YEAR-WISE DISTRIBUTION OF AUTHORSHIP PATTERN OF NIT'S

Authorship Pattern	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total	%	Mean	SD	C.V.
1	31	36	52	54	44	46	53	67	76	64	523	6.25	52.3	13.93	26.63
2	141	192	321	289	241	457	313	362	382	424	3122	37.29	312.2	100.15	32.08
3	112	155	164	205	185	256	259	268	283	301	2188	26.13	218.8	63.34	28.95
4	51	80	95	112	110	131	163	169	164	176	1251	14.94	125.1	42.60	34.05
5	27	33	38	67	58	69	59	65	67	96	579	6.92	57.9	20.45	35.32
6	18	7	22	26	18	39	22	46	52	45	295	3.52	29.5	14.92	50.59
7	11	8	12	11	14	20	21	20	22	26	165	1.97	16.5	6.00	36.39
8	5	4	7	8	6	11	12	12	12	20	97	1.16	9.7	4.74	48.85
9	2	6	4	10	7	8	10	2	5	10	64	0.76	6.4	3.13	48.97
>10	3	2	3	10	10	11	12	11	12	14	88	1.05	8.8	4.39	49.91
Total	**401**	**523**	**718**	**792**	**693**	**1048**	**924**	**1022**	**1075**	**1176**	**8372**	**100.00**	**837.2**	**254.62**	**30.41**
DC	**0.92**	**0.93**	**0.93**	**0.93**	**0.94**	**0.96**	**0.94**	**0.93**	**0.93**	**0.95**	**0.94**				

DC=Degree of Collaboration

It could be noted that two authored papers rank first in order sharing 37.29 percent of the total research output.   The year-wise analysis shows that the performance of two authored papers is better in almost all the years, except for 2001 and 2002. The three authored papers follow second in order taking 26.13 percent of the total research contributions. The year-wise analysis reveals that the three authored contributions have shown a considerable trend from 2006 onwards.

Four authored contributions take the third position in order sharing 14.94 percent of the total research output during the study period. The performance of research output was found to be less in the years 2001 – 2003. All the remaining years have recorded more than 100 contributions.

Five authored papers rank next reporting 579 contributions that result in 6.92 percent of total research output.   The year-wise analysis in its output performance was notable in the years from 2004 onwards.

It is interesting to note that single authored papers retain fifth place witnessing 523 contributions which represent only 6.25 percent of total research output.   The single authored contribution shows a gradual growth through the years of study.

Productivity of Authors based on Lotka's Law

It is appropriate to examine and analyse the implications of Lotka's law in relation to author productivity on research publications by NIT faculties of India.

To validate Lotka's law, a calculation was done using the equations (1–3) to identify the values of n and C to test whether the concept of Lotka's law fits into the data of the present study or not. Thus, based on the data presented in Table 4, the calculated values of n and C are 1.89 and 0.59 respectively. The calculated critical value using equation (4) is 0.24 and the value of maximum difference (D) between the real and estimated accumulated frequencies is 0.015. Therefore, it is clear that D value 0.015 (Table 4) is less than the critical value 0.24 which resulted in convincing application of Lotka's law to the data of literature produced by NIT faculties.

TABLE 4: DISTRIBUTION OF AUTHOR PRODUCTIVITY BASED ON LOTKA'S LAW

x	y	X	Y	X^2	XY	$y_x/\sum y_x$	$\sum(y_x/\sum y_x)$	$1/x^n$	$f=C(1/x^n)$	$\sum$	D
1	7463	0	3.873	0.000	0.000	0.605	0.605	1	0.590	0.590	**0.015**
2	1656	0.301	3.219	0.091	0.969	0.134	0.739	0.270	0.159	0.749	-0.010
3	891	0.477	2.950	0.228	1.407	0.072	0.811	0.125	0.074	0.823	-0.012
4	513	0.602	2.710	0.362	1.632	0.042	0.853	0.073	0.043	0.866	-0.013
5	279	0.699	2.446	0.489	1.710	0.023	0.876	0.048	0.028	0.894	-0.019
6	257	0.778	2.410	0.606	1.875	0.021	0.897	0.034	0.020	0.914	-0.018
7	167	0.845	2.222	0.714	1.878	0.014	0.910	0.025	0.015	0.929	-0.019
8	126	0.903	2.100	0.816	1.897	0.010	0.920	0.020	0.012	0.941	-0.020
9	105	0.954	2.021	0.911	1.929	0.009	0.929	0.016	0.009	0.950	-0.021
10	121	1.000	2.082	1.000	2.082	0.010	0.939	0.013	0.008	0.958	-0.019
11	79	1.041	1.896	1.084	1.975	0.006	0.945	0.011	0.006	0.964	-0.019
12	49	1.079	1.686	1.165	1.820	0.004	0.949	0.009	0.005	0.969	-0.020
13	55	1.114	1.741	1.241	1.940	0.004	0.953	0.008	0.005	0.974	-0.021
14	113	1.146	2.052	1.314	2.352	0.009	0.963	0.007	0.004	0.978	-0.015
15	39	1.176	1.595	1.383	1.876	0.003	0.966	0.006	0.004	0.982	-0.016
16	28	1.204	1.440	1.450	1.734	0.002	0.968	0.005	0.003	0.985	-0.017
17	22	1.230	1.348	1.514	1.659	0.002	0.970	0.005	0.003	0.987	-0.018
18	14	1.255	1.159	1.576	1.455	0.001	0.971	0.004	0.003	0.990	-0.019
19	17	1.279	1.232	1.635	1.575	0.001	0.972	0.004	0.002	0.992	-0.020
20	9	1.301	0.963	1.693	1.253	0.001	0.973	0.003	0.002	0.994	-0.021
21	14	1.322	1.159	1.748	1.533	0.001	0.974	0.003	0.002	0.996	-0.022
22	17	1.342	1.232	1.802	1.654	0.001	0.976	0.003	0.002	0.998	-0.022
23	13	1.362	1.118	1.854	1.522	0.001	0.977	0.003	0.002	0.999	-0.023
24	9	1.380	0.963	1.905	1.329	0.001	0.977	0.002	0.001	1.001	-0.023
25	277	1.398	2.442	1.954	3.414	0.022	1.000	0.002	0.001	1.002	-0.002
	12333	25.191	48.060	28.533	42.468	1.000	23.012	1.699	1.002		

C=0.59;   n=1.89;   c.v.=0.24;   D=0.015

RANKING OF INDIVIDUAL JOURNALS ACCORDING TO OUTPUT

Bibliometric studies assess a field usually beginning with the determination of some core set of periodicals and articles by a statistical analysis and use these to generate links to show their relationship in that field. In the present study, the journals are ranked on the basis of their published papers on Engineering and Technology output. Table 5 indicates the journals in which the NIT faculties preferred to publish their articles. Further, it was analysed to find out the key journals in the field of Engineering and Technology in which more number of publications were made by the faculties.

There were 1508 journals observed in which the NIT faculties have published their articles over the period of study. There have been 275 contributions published by a single journal 'American Journal of Food Technology' from USA and it is ranked at the first position. The second position is taken by 'Nature Biotechnology' which is accounted to 194 publications. The other two journals namely 'Nature' and 'Applied Composite Materials' stand at the next two ranks publishing 142 articles each. A detailed list of journals along with their related ranks has been provided in Table 5.

TABLE 5: RANKING OF JOURNALS ACCORDING TO THEIR NUMBER OF PUBLICATIONS

Sl. No.	Name of the Journal	No. of Articles	Journal's Rank
1	American Journal of Food Technology	275	1
2	Nature Biotechnology	194	2
3	Nature	142	4
4	Applied Composite Materials	142	
5	Automation and Remote Control	113	6
6	Journal of Computational Electronics	113	
7	Natural Language Engineering	97	9
8	Biotechnology and Bioengineering	97	
9	Heat and Mass Transfer	97	
10	Biotechnology: an Indian Journal	89	11
11	Journal of Computational Methods in Sciences and Engineering	89	

AUTHOR PRODUCTIVITY OF NITS RESEARCH OUTPUT

The productivity of scientific researchers, which normally referred as scientific productivity, is measured in terms of their scientific outputs that are published through different forms viz. papers, reports, books and technical output. These measures of scientific productivity have been generally accepted and used since the beginning of the twentieth century.

Moreover, it is a well known fact that the principal means of communication used by authors is a publication. It enables an author to obtain critical response to his/her work, to verify the reliability of information, to acquire a sense of relative importance of a contribution. Also, it is through publication that an author receives professional recognition and esteem, as well as promotion, advancement, and funding for further research. Publication is so important to an author's productivity as the research carried out by him/her becomes a 'work' when it is published that can be accessed, assessed, and acknowledged by the scientific community.

Keeping in view the above in mind, the author productivity is determined in the present study on the basis of number of papers contributed by NIT faculties.

In the present study, a total of 12333 authors contributed 8372 papers to Engineering and Technology literature were analysed. Table 6 shows the distribution of the number of articles published by each one of the author. The study reveals that around 60.51 percent of the authors have contributed only one paper in the field of Engineering and Technology. Hence it ranks first in order with respect to the total number of contribution during the period of study. The number of authors contribute two papers are 13.42 percent of the total authors and it is placed next. The contributors of three papers are found to be 7.22 percent and it is ranked at third place. The results of analysis reflected in Table 7 brings out a fact that when the number of contributions increases, the number of authors decreases. In the light of the above, a greater level of research performance is shown by a few NIT faculties. Out of various levels of contributions, 123 papers are the highest which is the productivity of one individual faculty member next to which 114 papers, 110 papers and 103 papers, 94 papers and so on.

It is noteworthy that the authors contributing from one paper to nine papers constitutes 92.89 percent of the total output. Further, it portraits the range of authors who have contributed more than 10 papers constitutes the remaining 7.11 percent.

It could be deduced from the above discussion that when the number of published papers increases, the number of contributing authors decreases. More number of publications by a researcher in any field requires high degree of inquisitiveness, competencies, efficiency, presences and exposure to literatures. That is the reason, in the present study, only a few authors have contributed more number of papers. Moreover author productivity is influenced by the nature of institution in which the faculties are working, their area of specialization and availability of infrastructure and fund for research.

TABLE 6: DISTRIBUTION OF AUTHOR PRODUCTIVITY OF ENGINEERING AND TECHNOLOGY

No. of Contribution	No. of Authors	Percentage	Cumulative Percentage
1	7463	60.51	60.51
2	1656	13.42	73.93
3	891	7.22	81.16
4	513	4.16	85.32
5	279	2.27	87.58
6	257	2.08	89.67
7	167	1.35	91.02
8	126	1.02	92.04
9	105	0.85	92.89
10	121	0.98	93.87
11	79	0.64	94.51
12	49	0.39	94.90
13	55	0.45	95.35
14	113	0.91	96.26
15	39	0.32	96.58
16	28	0.22	96.80
17	22	0.18	96.98
18	14	0.12	97.10
19	17	0.14	97.24
20	9	0.07	97.31
21	14	0.12	97.43

22	17	0.14	97.57
23	13	0.11	97.68
24	9	0.07	97.75
25	12	0.10	97.85
26	26	0.21	98.06
27	24	0.19	98.25
28	12	0.10	98.35
29	12	0.10	98.44
30	9	0.07	98.52
31	8	0.06	98.58
32	10	0.09	98.67
33	42	0.34	99.01
34	41	0.33	99.34
35	7	0.05	99.39
36	5	0.04	99.43
37	4	0.03	99.46
38	4	0.03	99.50
39	3	0.02	99.52
40	4	0.03	99.55
42	3	0.02	99.57
44	3	0.02	99.59
45	1	0.01	99.60
46	1	0.01	99.61
47	1	0.01	99.62

48	7	0.05	99.68
49	3	0.02	99.70
51	1	0.01	99.71
52	3	0.02	99.73
55	1	0.01	99.74
56	4	0.03	99.77
57	1	0.01	99.78
58	1	0.01	99.79
60	1	0.01	99.80
63	3	0.02	99.82
64	3	0.02	99.85
65	1	0.01	99.86
72	1	0.01	99.87
78	1	0.01	99.88
84	1	0.01	99.89
86	3	0.02	99.91
87	1	0.01	99.92
93	3	0.02	99.94
94	1	0.01	99.95
103	1	0.01	99.96
110	1	0.01	99.97
114	1	0.01	99.98
123	1	0.01	100.00
Total	**12333**	**100.00**	

CONCLUSION

Lotka's Law pertaining to author productivity is considered as one of the important classical law's of bibliometrics. This study clearly indicates that Lotka's generalized inverse square law holds good to Engineering and Technology literature published by the faculty of NIT's in India during the study period 2001 – 2010. In this study it has been found n=1.89 and c.v.. = 0.24 and c=0.59 for overall data using least square method.

This preliminary study on the authorship distribution in the field of engineering and technology research carried out in NIT's may initiate more such research to evaluate Engineering and Technology in other government funded Institutions. Future research should be focused towards understanding authorship distributions within various sub-fields patterns in monographs and other publication, types collaborative author affiliation etc. No doubt, such studies would be useful for a clear understanding of the development of Engineering and Technology in India

REFERENCES

1. Godin, Benoît (2006). On theOrigins of Bibliometrics.Scientometrics*, 68*(1): 109–133.

2. Glänzel, W. (2003). Bibliometrics as a Research Field: A Course on Theory and Application of Bibliometric Indicators.

3. Gupta, B.M., Kumar, S., Aggarwal, B. S. (1999). A Comparison of Productivity of Male and Female Scientists: of CSIR. Scientometrics, 45:269-289.

4. Prpic,K.(2000).ThePublication Productivity of Young Scientists: An Empirical Study. Scientometrics, 49: 453–490.

5. Lotka, A.J. (1926). The Frequency Distribution of Scientific Productivity. Journal of Washington Academy of Sciences, 16: 317-323.

6. Rowlands, Ian (2005). Emerald Authorship data, Lotka's Law and Research Productivity. *Aslib Proceedings: New Information Perspectives*, 57(1):5-10.

7. Yueh, Ming, Jou, Shiow-Jem and Ma, Sheau-Shin (2000). Bibliometric Study of Semiconductor Literature, 1978-1997. *Scientometrics,* 49: 491.509.

8. Kawamura, M… [et al] (1999). Lotka's Law and the Pattern of Scientific Productivity in the Dental Science Literature. *Medical Informatics and the Internet in Medicine*, 24(4):309-315.

9. Pao, M. L. (1985). Lotka's law: a testing procedure. Information Processing & Management, 21(4), 305–320.

10. Kolmogorov A (1933). "Sulla determinazione empirica di una legge di distribuzione". *G. Inst. Ital. Attuari* 4: 83.

11. Glanzel, W. (2001). National Characteristics in International Scientific Co-authorship. Scientometrics, 51(1): 69 – 115.

12. Gomez, I, Fernandez, M.T. and Mendez, A. (1995). Collaboration Patterns of Spanish scientific publications in different research areas and disciplines. In: *Proceedings of the Biennial Conference of the International Society for Scientometrics and Informetrics* (ed. By M.E.D. Koeing and A.Bookstein), 187 – 196.

13. Ding, Y., Foo, S., and Choudhury, G. (1999). A bibliometric analysis of collaboration in the field of information retrieval. *The International Information and Library Review*, 30, 367 – 376.

14. Glanzel, W. (2002). Co-authorship patterns and trends in the sciences (1980 – 1998): A bibliometric study with implications for database indexing and search strategies, *Library Trends,* 50 (3): 461 – 473.