3-10-2021

Application of Bradford's Law of Scattering to Synthetic Biology Literature

Naheem KT Mr.
Dept. of Library & Information Science, Annamalai University& Central Library, Pondicherry University, naheemkt@gmail.com

Sivaraman P. Dr.
Dept. of Library & Information Science, Annamalai University, psraman.p@gmail.com

Saravanan G Dr
French Institute of Pondicherry, saravanan.g@ifpindia.org

Application of Bradford's Law of Scattering to Synthetic Biology Literature

Naheem K.T.,[1]Research Scholar, Dept. of Library & Information Science, Annamalai University, Annamalai Nagar – 608002.; [2]Professional Assistant, Central Library, Pondicherry University, Kalapet, Puducherry-605014. Email: naheemkt@gmail.com;

Dr P. Sivaraman[3]., Professor and Coordinator, Library & Information Science Wing, DDE, Annamalai University, Annamalai Nagar – 608002. Email: psraman.p@gmail.com

Dr G. Saravanan[4]., Librarian, Library and Publication Division, French Institute of Pondicherry. Email: saravanan.g@ifpindia.org

Abstract

This study aimed to examine Bradford's law of scattering on the global synthetic biology literature output published between 2005 to 2019. The authors used the publications data obtained from Web of Science database for the present study. The total data set consists of 8411 journal articles and 1453 journals. The authors employed both verbal and graphical formulations of Bradford's law. The author's calculated the value of Bradford's multiplier using Egge's formula and found it as 10.52. This analysis shows that a three-zone Bradford's distribution fits with the present data west with computed value Bardford's multiplier 10.52. A list of core journals covering around twelve journals also prepared.

Keywords: Synthetic Biology, Bradford's law of scattering, Journal articles, Core journals, Web of science

1 Introduction

Bradford's Law[1] of scattering is one of the prominent laws of bibliometrics. This law is a proven methodology for the identification of core journals in various fields. Samuel Clement Bradford formulated this law in 1934. He conducted a study in geophysics based on the publications that appeared between 1931 to 1933. This study found a regularity, namely an "inverse relationship" between the number of articles published in a subject area and the number of journals in which the items appear. According to Bradford, "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 groups or zones containing the same number of articles as the nucleus, when the numbers of periodicals in the

nucleus and succeeding zones will be as 1: n: n2[...][2]". Which means that, in a certain filed, few journals credited with a considerable amount of the total publications in that area. In contrast, more numbers of journals publish less articles in that area. In other words, only few journals make available the core of papers or articles on a given subject. The purpose of the present study is to examine Bradford's Law of scattering on literature output published in Synthetic Biology during 2005- 2019.

2. Review of literature

So far, many authors applied and tested the conformity of Bradford's law of scattering in various subjects. Gopakumar and Sudhier[3] examined the applicability of Bradford's law of scattering in the Indian biotechnology literature. The data for this study collected from WOS database. The authors employed verbal and graphical formulations to test the applicability of Bradford's law to biotechnology literature. Results show that the analyzed data set fit well with Bradford's law with the calculated value of k 8.10. In another study, Sudhier[4] reported that Bradford's law does not found a fit with Physics literature. In this also both verbal and Leimkuhler models were applied for testing Bradford's law.

Similarly, Savanur[5] applied Bradford's law f scattering to the literature in Economics. And found that the journal distribution pattern of the economics literature fits with Bradford's law. Venable[6] et al., examined the applicability of Bradford's law in "Pediatric neurosurgery". Egghe' s formulation and the verbal formulation of Bradford's law were used in this study. Hiremath[7] et al., examined Bradford's law in Material science literature using the data obtained from Web of science database. In this paper authors tested the theoretical aspects of Bradforsd's law and found that the data do not fit with sample data. Ram and Plaiwal[8] tested the application of Bradford's law to Psoriasis literature with the help of data obtained from PubMed from 1960 to 2009. The authors applied the theoretical and practical aspects of Bradford's law and found that the law does not fit the theoretical distribution. Similarly, Tunga[9] applied Bradford's law of scattering to the horticulture literature and found that this subject's journal distribution pattern does not fit with Bradford's law of scattering.

3. Objectives

Following are the main objectives of the study.

1. To apply and test Bradford's law of scattering to Synthetic Biology literature.
2. To find out the core journals in Synthetic Biology.

4.Methodology

The authors retrieved data for this study from Web of Science database of Clarivate Analytics. Bibliographical details pertain to the document category " Journal Article" were identified using an advanced search strategy developed with the help of combination key terms. The authors took the identified records consisting of 8411 journal articles and 1453 journals publishing these articles were taken for analysis. The descriptive statistical analysis was done with the help of Micro soft Excel 2010. The following methodology is adopted to apply and test Bradford's Law of scattering to this data set.

There is two widely recognized formulation of the Bradford's Law, the verbal formulation and the graphical formulation. In the verbal formulation, the authors used the following steps:

1. Identification of articles published in the field of study.
2. List the source, i.e., journals, that publish the articles in rank order, beginning with the source journal that produces most items.
3.Calculation of the value of Bradford's constant k
4.Calculation of the number of journals that will belong to Bradford's core zone (r0) and subsequent zones.
5. Distribution of the number of journals in each zone along with the number of articles in a table and calculate the value of k.

Graphical Formulation

The graphical formulation is done with the help of a logarithmic plot (common logarithm) of the cumulative number of journals on the horizontal (X) axis and the cumulative number of articles on the vertical (Y) axis.

5. Analysis and findings

Analysis of the data for the distribution of synthetic biology literature output indicates that from 2005 to 2019, a total of 8411 SB related literature (i.e., journal articles) scattered in 1453 journals. Data presented in Table 1 used to test the verbal formulation of Bradford's law. The rank order, number of journals, cumulative numbers, articles, cumulative frequencies of articles and the log value of cumulative articles also provided. The numbers of journals are arranged by decreasing order of frequencies (i.e., number of articles).

Table 1 Dispersion of scientific literature in Synthetic Biology area

Rank	No. of journals	Cum. No. of Journals	No. of articles	Total No. of articles	Cumulative No. of articles	Log N (Cumulative journals)
1	1	1	655	655	655	0.0000
2	1	2	275	275	930	0.6931
3	1	3	272	272	1202	1.0986
4	1	4	242	242	1444	1.3863
5	1	5	165	165	1609	1.6094
6	1	6	158	158	1767	1.7918
7	1	7	154	154	1921	1.9459
8	1	8	114	114	2035	2.0794
9	1	9	113	113	2148	2.1972
10	1	10	100	100	2248	2.3026
11	2	12	98	196	2444	2.4849
12	1	13	82	82	2526	2.5649
13	2	15	67	134	2660	2.7081
14	1	16	63	63	2723	2.7726
15	1	17	58	58	2781	2.8332
16	1	18	56	56	2837	2.8904
17	1	19	55	55	2892	2.9444
18	1	20	52	52	2944	2.9957
19	1	21	49	49	2993	3.0445
20	2	23	48	96	3089	3.1355
21	1	24	47	47	3136	3.1781
22	2	26	46	92	3228	3.2581
23	2	28	45	90	3318	3.3322
24	1	29	44	44	3362	3.3673
25	1	30	42	42	3404	3.4012
26	1	31	41	41	3445	3.4340

27	1	32	37	37	3482	3.4657
28	1	33	36	36	3518	3.4965
29	1	34	35	35	3553	3.5264
30	3	37	34	102	3655	3.6109
31	3	40	33	99	3754	3.6889
32	2	42	31	62	3816	3.7377
33	2	44	30	60	3876	3.7842
34	2	46	29	58	3934	3.8286
35	3	49	28	84	4018	3.8918
36	2	51	27	54	4072	3.9318
37	1	52	26	26	4098	3.9512
38	2	54	25	50	4148	3.9890
39	1	55	24	24	4172	4.0073
40	5	60	23	115	4287	4.0943
41	4	64	22	88	4375	4.1589
42	2	66	21	42	4417	4.1897
43	2	68	20	40	4457	4.2195
44	6	74	19	114	4571	4.3041
45	5	79	18	90	4661	4.3694
46	7	86	17	119	4780	4.4543
47	10	96	16	160	4940	4.5643
48	11	107	15	165	5105	4.6728
49	11	118	14	154	5259	4.7707
50	5	123	13	65	5324	4.8122
51	12	135	12	144	5468	4.9053
52	5	140	11	55	5523	4.9416
53	19	159	10	190	5713	5.0689
54	13	172	9	117	5830	5.1475
55	18	190	8	144	5974	5.2470
56	26	216	7	182	6156	5.3753
57	32	248	6	192	6348	5.5134
58	52	300	5	260	6608	5.7038
59	64	364	4	256	6864	5.8972
60	116	480	3	348	7212	6.1738
61	226	706	2	452	7664	6.5596
62	747	1453	1	747	8411	7.2814

It is found that a single journal publishes the highest numbers of articles (655), and 747 journals together contribute a total of 747articles. It is obvious to note that the top 10 journals contribute a notable portion of literature during the study period.

5.1 Application of Bradford's Law

The steps presented in the methodology part of this paper are used to apply verbal formulation of Bradford's law. To find out the value of multiplier k of Bradford's law, the following equation formulated by Egge[10] and Egghe and Rousseau[11] were used:

$k = (e^y * y_m)^{1/p}$

where,

e^y is the Euler's number (value 1.781),

y_m is the number of articles in the journal, and

P is the number of zones.

It can be seen from the above Table 1 that a total of 8411 articles published in 1453 journals. And the most productive journal published 655 articles.

Thus; y_m: 655, $e^{y: 1.781,}$ *P*:3

Hence, $k = (1.781 \times 655)^{1/3}$
$= (1166.56)^{1/3}$
Taking log of both sides
$\text{Log}k = (1166.56)^{1/3} = (3.066907)^{1/3} = 1.022302$
k = antilog of 1.022302= 10.5269= 10.52
k = 10.52

The number of journals (core journals) in Bradford's first zone, i.e., r_o, can be calculated utilizing the following formula:

$r_0 = T(k-1)/(k^P-1)$

T = Total number of journals : 1453
k = Bradford's constant : 10.52
P = Number of Bradford's zones: 3

$r_0 = 1453 (10.52 – 1) / (10.52^3 - 1)$
$= 1453(9.52) / (1164.252 - 1)$
$= 13832.56 / 1163.252$
$r_0 = 11.89$

It can then be concluded that twelve journals will be the most productive (core journals. Knowing the values of both *k* and *r0* we can obtain the theoretical distribution of journals across the reaming Bradford's zones (corresponding to r_1 *and* r_2). The expected number of journals in different zones in this present study will be :

r0 = *r0* x 1 = 11.89 = 12
r1 = *r0* x k = 11.89 x 10.52 = 125.082 = 125
r2 = *r0* x k^2 = 11.89 x 10.52^2 = 1315.87 = 1316

Table 2 Bradford's zones in Synthetic Biology

Zones	**No. of journals**	**No. of articles**	***k***
Core zone	12	2444	-
Zone 1	125	3046	10.42
Zone 2	1316	2921	10.53
Total	1453	8411	10.47

From this theoretical distribution, it is possible to test Bradford's law's exact fit for the present study. The number of articles in each zone's can be counted, as shown in Table 2. The exact value of multiplier k is also calculated by dividing the number of journals of a zone by the number of journals to its previous zone.

According to S.C Bradford "zones will form about a geometric pattern in the form of 1:n:n^2. Here the relationship among each zone is 12:125:1316. It means that 12 journals constitute the nucleus zone i.e., core journals in the field under study. And 10.47 is the mean value of multipliers.

Thus 1 : n: n^2

:: 12 : 12*10.47 : 12*(10.47)^2

:: 12 : 125.64 : 1315.45

>>> 1453.09

Percentage error is = ((1453-1453.09) / 1453)*100 = -0.0061%

Since the percentage of error is meager and k values are very similar and similar to the calculated value by the k formula (k = 10.52), it can conclude that the data fit a three-zone Bradford's distribution.

5.2 Graphical Formulation of Bradford's Law

Brooks[12] developed the graphical approach to verify the verbal formulation of Bradford's law. The same has used for plotting the visual formulation in the present study. This graphical formulation is just the experimental verification of the verbal formulation, which observes certain regularity in distributing scientific publications[3]. The same has been attempted and presented in Figure 1. This graph is a logarithmic plot (common logarithm) of the cumulative number of journals on the horizontal (X) axis and the cumulative number of articles on the vertical (Y) axis. Figure 1 showcase Bradford's curve with a gross droop. It can be observed from the figure that the core journals are those lie along the initial curved part of the Bradford plot before it becomes a straight line.

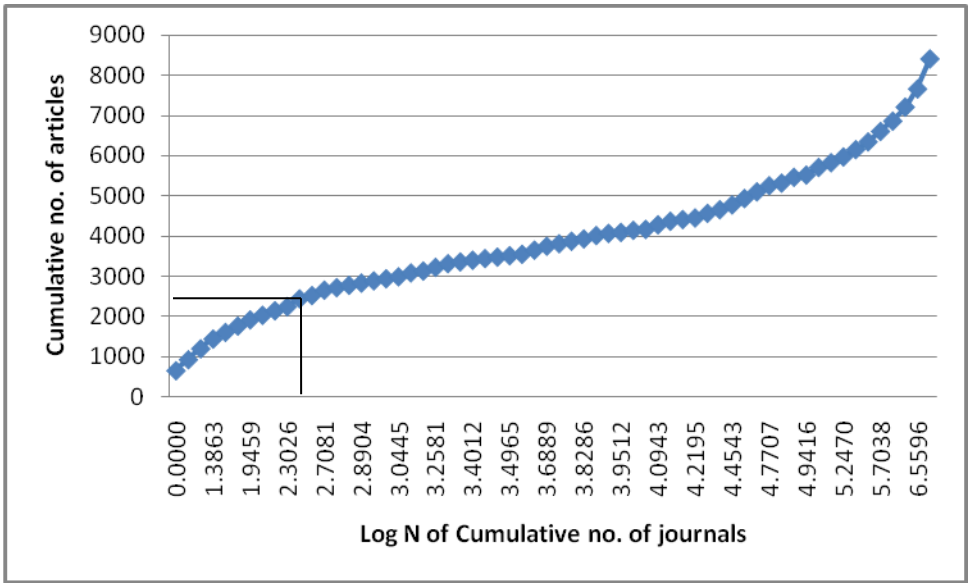

Figure 1: Bradford Bibliograph

Table 3 shows the list of journals that are coming under the core zone. This table gives the journal's name with the corresponding number of articles, percentage share of articles, citations, citations per paper (CPP), and impact factor (JCR 2019) of each journal. These twelve journals together contributed a total of 2444 articles, which is about 29 per cent of the total items. It is also shows that the articles together received a total of 77967 (34%) citations during the study period.

Table 3: Core journals in Synthetic Biology

Sl.no	Name of the journal	No. of articles	% of 8411 articles	Citations	Citations Per Paper	IF (2019)
1	ACS Synthetic Biology	655	7.79	12068	18.42	4.411
2	Nucleic Acids Research	275	3.27	11448	41.63	11.502
3	PLOS One	272	3.23	5631	20.70	2.74
4	Proceedings of the National Academy of Sciences of the United States of America	242	2.88	19332	79.88	9.412
5	Scientific Reports	165	1.96	2303	13.96	3.998
6	Nature Communications	158	1.88	5704	36.10	12.121
7	Metabolic Engineering	154	1.83	6204	40.29	7.263
8	Biotechnology and Bioengineering	114	1.36	4120	36.14	4.002
9	Journal of the American Chemical Society	113	1.34	4570	40.44	14.612
10	Microbial Cell Factories	100	1.19	2097	20.97	4.187
11	Angewandte Chemie-International Edition	98	1.17	3087	31.50	12.959
12	ChemBioChem	98	1.17	1403	14.32	2.576
	Total of core journals	2444	29.06	77967	31.90	
	Total no. of articles	8411		228950	27.22	
	%Share of core journals	29.06		34.05		

It is found from the Table 3, that the core twelve journals have 98 to 655 articles. Of which journal "ACS Synthetic Biology" is the top journal with 655 articles followed by "Nucleic

Acids Research" with 275, "PLOS One" with 272, "Proceedings of the National Academy of Sciences of the United States of America" with 242 articles and so on.

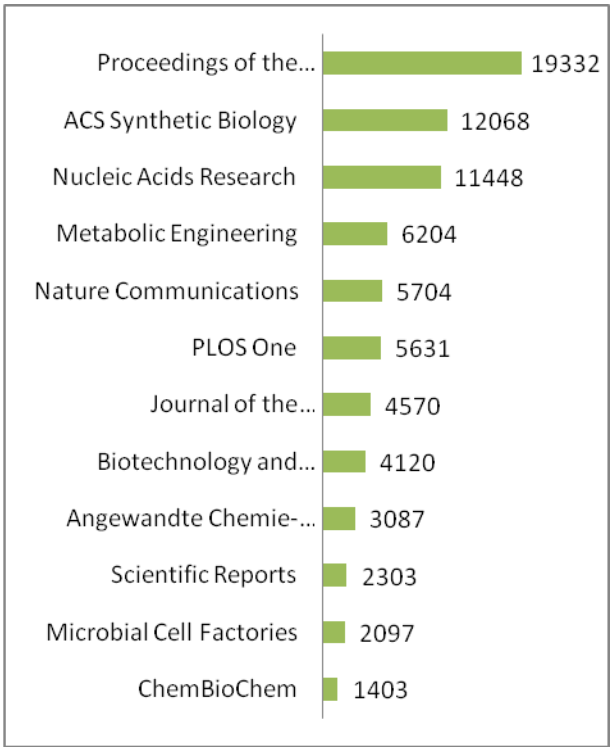

Figure 2a. Citation based ranking of core journals

With respect to citations, these core journals received ciations from 1403 to 19332 citaions during the study period. The journal "Proceedings of the National Academy of Sciences of the United States of America" topped the list with 19332 citations followed by "ACS Synthetic Biology" with 12068 citations, "Nucleic Acids Research" with 11448 citations, "Metabolic Engineering" with 6204 citations and so on (Fig.2a).

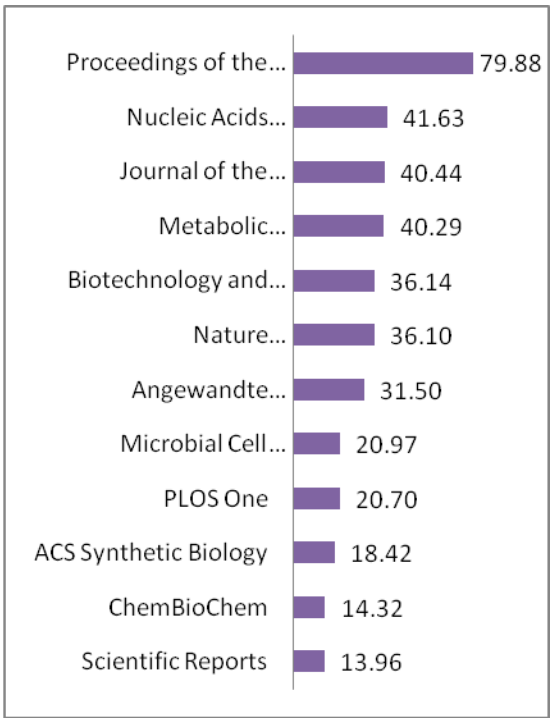

Figure 2b. CPP based ranking of core journals

With regards to CPP the journal "Proceedings of the National Academy of Sciences of the United States of America" leads with 79.88 per cent followed by "Nucleic Acids Research" with 41.68 per cent, "Journal of the American Chemical Society" with 40.44, "Metabolic Engineering" with 40.29 per cent and so on (Fig.2b).

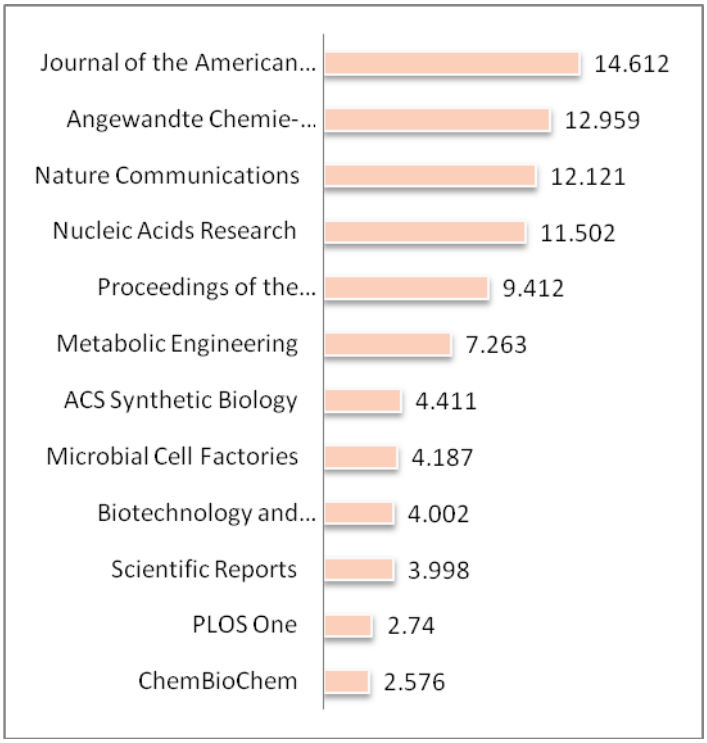

Figure 2c. IF based ranking of core journals

In the case of IF, these core journals have an impact factor ranging from 2.576 to 14.612. The journal "Journal of the American Chemical Society" occupies the first spot with 14.612 IF (JCR2019), followed by "Angewandte Chemie-International Edition" with IF 12.959, "Nature Communications" with IF 12.121 and so on(Fig.2c).

6. Conclusion

Bradford's law of scattering is an indispensable procedure to find out the core journals in a research field. The present analysis of the journal productivity pattern in Synthetic Biology literature was based on the data retrieved from WOS database for a period of fifteen years, discloses the core or most preferred journals in this research filed. Both verbal and graphical formulation of Bradford's law was applied and found that the journal distribution pattern of synthetic biology literature fits well with Bradford's law. The calculated value of Bradford's multiplier was 10.52. The present study, which based on the publication data of global synthetic biology research output, will help librarians select the core journals in this field using the list of core journals prepared by applying Bradford's scattering law.

References

1. Bradford, S. C. (1934), Sources of information on specific subjects. *Engineering:An Illustrated Weekly*, 137: 3550, 85–86.
2. Bradford, S.C. (1950),Documentation, Public Affairs Press, Washington, DC.
3. Gopakumar, V. and Sudhier, KG (2020). Testing Bradford's law of scattering and identifying core journals of Indian biotechnology research. IASLIC Bulletin, *65*(2), 115-127.
4. Sudhier, K. G. (2010). Application of Bradford's Law of Scattering to the Physics Literature: A Study of Doctoral Theses Citations at the Indian Institute of Science. *DESIDOC Journal of Library & Information Technology*, *30*(2), 3-14.
5. Savanur, K. P. (2019). Application of Bradford's Law of Scattering to the Economics Literature of India and China: A Comparative Study. *Asian Journal of Information Science & Technology (AJIST)*, *9*(1).
6. Venable, G. T., Shepherd, B. A., Roberts, M. L., Taylor, D. R., Khan, N. R., & Klimo, P. (2014). An application of Bradford's law: identification of the core journals of pediatric neurosurgery and a regional comparison of citation density. *Child's Nervous System*, *30*(10), 1717-1727.
7. Hiremath, M. R., Gourikeremath, M. G. N., Hadagali, G. S., & Kumbar, B. D. (2016). Application of Bradford's Law of Scattering to the Materials Science Literature: A study based on Web of Science Database. *International Journal of Library and Information Science*, *6*(4), 157-172.
8. Ram, S., & Paliwal, N. (2014). Assessment of Bradford Law's of Scattering to Psoriasis Literature through Bibliometric Snapshot. *DESIDOC Journal of Library & Information Technology*, *34*(1), 46-56.
9. Tunga, S. K. (2013). Application of Bradford's Law of Scattering to the Horticulture Literature: a Citation Study of Doctoral Dissertations 1991-2010. *SRELS Journal of Information Management*, *50*(3), 305-316.
10. Egghe, L. (1990). Applications of the theory of Bradford's law to the calculation of Leimkuhler's law and to the completion of bibliographies. *Journal of the American Society for information Science*, *41*(7), 469-492.

11. Egghe, L., & Rousseau, R. (1990). Introduction to informetrics: Quantitative methods in library, documentation and information science. Elsevier Science Publishers.

12. Brookes, B. C. (1969).The complete Bradford-Zipf 'Bibliograph'. Journal of Documentation. 25(1), 58-60