Webology (ISSN: 1735-188X)
Volume 17, Number 1, 2020

Testing Of Bradford's Law Of Scattering In Pathogenic Research Output: A Study Based On Web Of Science

Dr. N.Amsaveni[1] , Dr. A.Rosaline Mary[2]

[1]Assistant Professor, Department of Library and Information Science, Bharathidasan University, Trichy, Tamilnadu.

[2]Librarian, Kalai Kaviri College of Fine Arts, Thiruchirappalli – 01, Tamilnadu.

Abstract

Bradford's Law of Scattering is a commonly used bibliometric principle, often applied in bibliometric research by library and information science professionals. In this particular study, the focus is on its relevance in the context of pathogenic literature. The research involves an analysis of journals published in this field during a specific timeframe. To assess the applicability of Bradford's Law of Scattering, the study utilizes a dataset comprising 140222 bibliographic entries sourced from the Web of Science database spanning from 2001 to 2019. Through ranking, it was observed that the journal "PLOS One" secured the top position with 89066 citations and 3469 articles, trailed by "Scientific Reports" with 14287 citations and 1141 articles, and "Journal of Virology" with 49858 citations and 1122 articles. The study examines various methodologies to gauge the applicability of Bradford's Law, yielding insights into its validity.

Keywords: Pathogenic research, Bradford's Law, Scattering of journal, Egghe's Model, Leimkuhler's Model, Efficiency index, h-index, immediacy index.

INTRODUCTION

The aim of the present study is to evaluate the output of research literature in the field of Pathogenic studies, focusing on both comprehensive and specific aspects. The researcher collected necessary bibliographic records for analysis from the Web of Science database, covering the period from 2001 to 2019, on February 10, 2021. The summarized table provides concise information about the total biographical details. Librarians and educators in library and information science have outlined different approaches and objectives for compiling such lists (Nisonger, 2007). Bradford's law of scattering elucidates the distribution of literature on a specific topic across journals. This principle stems from an investigation conducted in 1933 by L. Jones at the Science Museum Library in London. It was initially documented in 1934 in the 'Engineering' journal by Bradford and subsequently in the author's book titled 'Documentation' in 1948 (Bradford, 1948).

REVIEW OF LITERATURE

Dharanikumar, P., Banateppanavar, K., Girish, T. S., & Jayaraj, A. N. (2014) has focused on the number of articles, authorship patterns, distribution of articles by subject, average number of references per article, types of materials mentioned, and Bradford's Law.Findings The survey finds that journals are the most frequently cited information sources among LIS researchers (50.37% of citations). It is followed by web resources, which account for 19.26% of all citations. Together, these two sources account for 69.63% of all citations, with the remaining 30.37% coming from books, conferences, theses, reports, and gazetteers. With 31 (1.91%) citations, Library Review (UK) is in first place. Bradford's Law of Scattering was also utilized. **Chaturbhuj, S., Batcha, M. S., Chaturbhuj, S., & Sadik Batcha, M. (2020)** has analyzed the study focuses on the application of Bradford's scattering law to the field of fluid mechanics.The verbal formulation 1: n: n2 is shown to be incompatible with the fluid mechanics facts. Bradford's law was verified using the Leimkuher model, and it fits the 2:57:1462 geometric series with an error of 0.005%. The study discovered that there are tendencies in international collaboration among the leading 13 highly productive nations. The most productive author in the discipline of fluid mechanics is Tenduyar, T. E., while Chemical Engineering Science is the most preferred journal. **Amsaveni, N. (2016)** has purposed to examine how Bradford's Law of Scattering has been applied to the Indian-published research on neural networks between 2001 and 2015. The study provides a summary of research on Bradford law that has been done on it in many studies, both theoretical and practical, and it is tested here using literature on neural networks. Regarding the Bradford law's implications, this law does not hold up theoretically; nevertheless, there are alternatives, such as the Leimkuhler model, which holds up well in the literature on neural networks.

OBJECTIVES

The main purpose of this study intends to accomplish the following objectives:

- To find the core journals that holds a significant portion of the overall Pathogenic literature.
- To assess the suitability of verbal and graphical interpretations of Bradford's Law of Scattering.

METHODOLOGY

The information utilized in this study was sourced from the Web of Science database, which is maintained by Thomson Reuters in the UK. The research encompassed content present in the Web of Science Core Collection database and involved a search string used the keyword 'Pathogenic' specifically within the titles of articles. The investigation spanned the timeframe from 2001 to 2019. This search yielded results from 8,548 sources that published 140222 articles and were subjected to analysis using Bradford's law.

Publication efficiency index (PEI), in 1977 Frame developed. PEI = TNCi/TNCt divided by TNPi/TNPt; Where TNCi = total number of citations in a year, TNCt is total number of citations for all the years. TNPi is total number of publications in a year, TNPt is total number of publications for all the years. PEI value is more than 1, it means that the impacts of publication are more than the research efforts devoted it. PEI value is less than 1, it means that the impacts of publication are more than the research efforts devoted it.

Relative Quality Index (RQI) = TNCp/TNCt divided by TNCp/TNPt. RQI value is less than 1, it means that the average citation of a year is less than the overall average citations. RQI

value is greater than 1, it means that the average citation of a year is less than the overall average citations.

Collaborative index measured by the formula of dividing the total number of articles published divided by the total number of authors, i.e., $\sum = 1/f_i /N$ (Lawani, 1986). $CI = (f_1)^1+ (f_2)^2+(f_3)^3\ldots.+L+(f_k)^k / N$; where f_1, f_2, f_3 = number of authors; N is number of publication in that particular year. Degree of collaboration is measured by the Subramanian formula is $DC = N_m + N_s / N_m$.

h-index derived by Hirsch (2005) means the number of publications a scientist has produced that receive at least h citations compared to the number of publications a scientist has published. g-index means publications with widely cited citations are given a higher weighing than those with high h-index. m-index meaning is the average number of times an article is cited in the year it is published.

ANALYSIS AND INTERPRETATION

Table 1 indicated the distribution of records the largest portion of records, constituting 94.045% of the total, is attributed to the Social Science Index (SCI). This suggests that a substantial number of scholarly articles, publications, or resources relevant to the social sciences are present in this database, followed by 3.4% (5,237) from the Conference Proceedings Citation Index – Science (CPCI-S), The Current Chemical Reactions (CCR-EXPANDED), Arts & Humanities Citation Index (A&HCI), Conference Proceedings Citation Index – Social Sci. & Humanities (CPCI-SSH), and Book Citation Index – Social Sciences & Humanities (BKCI-SSH) collectively represent smaller percentages of the records. These indices cover a variety of disciplines, including chemical reactions, arts, humanities, and interdisciplinary areas.

Table 1: Sample data indexes from the Web of Science Database

Database	**Recs.**	**%**
Social Science Index (SCI)	146145	94.05
Conference Proceedings Citation Index – Science (CPCI-S)	5237	3.370
Social Sciences Citation Index (SSCI)	1753	1.128
Index Chemicus (IC)	1216	0.783
Book Citation Index – Science (BKCI-S)	867	0.558
Current Chemical Reactions (CCR-EXPANDED)	79	0.051
Arts & Humanities Citation Index (A&HCI)	69	0.044
Conf. Proceedings Citation Index – Social Sci. & Humanities (CPCI-SSH)	28	0.018
Book Citation Index – Social Sciences & Humanities (BKCI-SSH)	5	0.003
Total	**140222**	**100**

Table 2 provides an overview of the sample details extracted from the Web of Science database for Pathogenic research. The time span covered in this dataset is from 2001 to 2019, encompassing

a 20-year period. The sample consists of 140222 records that have been downloaded. The dataset involves a total of 1024098 contributing authors, 35104 distinct journals, and identifies 3785120 frequently occurring words. The data includes 30 document types during the sample time spans 30 languages.

Contributions hail from 152 countries across various continents, originating from 23221 institutions worldwide. The research records have been cited a cumulative total of 99513 times in local and global references. This citation activity comprises 23,896 local citation scores and 113,695 global citation scores. Overall, the cited references for the entire dataset amount to 169,321. The provided data appears to represent a collection of statistics related to a certain academic context or dataset. It seems to be organized over a span of 20 years, from 2001 to 2019. Each year is associated with various metrics, over the years; the number of records has steadily increased, reflecting a growing collection of academic resources. The collaborative index values are 4.84 at 2001 and it was raised 7.93 at 2019. The average collaborative index value is 6.05, that means averagely 6 authors were collaborated in one article of their respective research activities of pathogenic subject.

According to table values of PEI, 2001 to 2013 years publications were earned more than 1, it means that the impacts of publication are more than the research efforts devoted it. The starting years of 2014 to 2019 publications were have value is less than 1 it means that the impacts of publication are less than the research efforts devoted it. In controversy, RQI values identified 2001 to 2013 years publications are less than 1, it means that the average citation of a year is less than the overall average citations. 2014 to 2019 publications RQI value is greater than 1, it means that the average citation of a year is less than the overall average citations.

The degree of collaboration values are 0.92 at 2001 and it was raised 0.98 at 2019. The average degree of collaboration index value is 0.95, that means only 5 percent of articles were published comes under the single authorship based and remaining 95 percent of publications were collaborated pattern of the pathogenic research output during the sample periods.

Table 2: Year wise detailed information of Pathogenic Research Productivity During 2001 to 2019

S. No	Year	Recs.	RQI	TLCS	TGCS	PEI	CR	contributing Authors	Actual authors	Sources	Countries	institutions	Word occurrence	CI	DC
1	2001	3027	0.432	244	189970	2.225	126498	14656	12978	1105	88	2429	8607	4.84	0.92
2	2002	3293	0.434	232	203626	2.192	141389	16099	14257	1170	87	2717	8631	4.89	0.92
3	2003	3583	0.451	210	216179	2.139	157890	17729	15589	1278	90	3001	9170	4.95	0.92
4	2004	3872	0.426	288	243305	2.228	165957	19645	17242	1278	92	3289	9712	5.07	0.91
5	2005	4239	0.432	421	264523	2.212	191927	22382	19224	1346	98	3417	10235	5.28	0.93
6	2006	4701	0.461	404	274521	2.070	212744	25767	22334	1511	105	3897	10849	5.48	0.93
7	2007	5209	0.235	425	268185	1.825	232918	28332	24237	1674	109	4505	11840	5.44	0.94
8	2008	5915	0.554	555	285915	1.714	269734	32764	27895	1854	113	5120	12700	5.54	0.94
9	2009	6257	0.579	705	289381	1.640	290512	36462	29996	1891	128	5281	13029	5.83	0.94
10	2010	6785	0.641	615	285323	1.491	323755	40130	33235	2082	131	5916	13971	5.92	0.95
11	2011	7622	0.701	644	291131	1.354	370174	46286	37768	2151	123	6588	14975	6.07	0.96

12	2012	8182	0.835	690	262388	1.137	377444	51346	41734	2231	126	6986	15812	6.28	0.96
13	2013	8691	0.881	729	264684	1.080	451331	55970	45095	2309	137	7578	16447	6.44	0.97
14	2014	9229	1.019	747	243191	0.934	485729	60951	48530	2350	131	8174	17034	6.60	0.97
15	2015	9837	1.190	920	221516	0.798	510062	66236	52414	2487	137	8783	17898	6.73	0.97
16	2016	10707	1.460	828	194408	0.644	569044	75299	57637	2515	149	10005	18976	7.03	0.97
17	2017	11651	1.885	990	162127	0.493	630791	84118	63212	2708	151	11076	20005	7.22	0.97
18	2018	12669	2.549	1071	123321	0.345	689295	94625	69585	2761	151	12272	20990	7.47	0.98
19	2019	14753	4.635	1247	71253	0.171	809125	116949	81736	**2981**	**158**	14454	23077	7.93	0.98
		140222 (7380)		**11965 (630)**	**4354947 (229208)**	**26.69 (1.4)**	**7006319 (368754)**	**905746 (47671)**	**714698(37616)**	**37682 (1983)**	**2304 (121)**	**12548 (6605)**	**273958(14419)**	**6.05**	**0.95**

Note: TLCS – Total Local Citation Scores; TGCS – Total Global Citation Scores;

The dataset exhibits variations in the numbers of contributing and actual authors, indicating changes in collaboration or authorship patterns. The number of sources and institutions also generally increases, potentially indicating a broader range of publications and affiliations being represented. Overall, this dataset seems to capture evolving patterns within an academic context, including changes in citation behavior, authorship trends, and publication characteristics.

Table 3: Ranking of Journals Based on SJR

Most productive journals	h_index	g_index	m_index	TC	NP	PY_start
PLOS One	468	664	51.938	**89066**	**3469**	2006
Scientific reports	137	182	24.593	14287	1141	2018
Journal of Virology	523	782	49.596	49858	1122	2001
Journal of Biological Chemistry	509	758	45.331	51451	1018	2001
Journal of Immunology	**524**	**822**	46.997	55552	978	2001
Frontiers In Microbiology	162	242	27.233	14846	915	2010
Applied and Environmental Microbiology	468	671	42.421	41517	904	2001

Table 3 reveals that the according to ranking of journals (seven journals) produced (more than 900 articles) pathogenic research. Also identified the h-index, g-index, immediacy index, total number of publication with earned total citations and started the publication year of particular journal. "PLOS One" produced highest number of articles with earned highest number of citations (89066), 468 h-index, 664 g-index started year is 2006 and 3469 articles were published by this particular journal. Followed by "Journal of Immunology" has highest h index and g index values with 978 articles its started from 2001.

BRADFORD'S LAW: A THEORETICAL INTERPRETATION

Samuel Clement Bradford, a chemist and chief librarian at the Landon Science Museum, has made a statistical analysis of two geophysics bibliographies, the Current Bibliography of Applied Geophysics (1992-1931) and the Quarterly Bibliography of Lubrication (1931-1933). Bradford's Law of scattering is used as a tool to study the output of journals. It expresses the quantitative connection between journals where the journals are arranged in descending order of productivity and divided into equal zones. He defined the first zone as "nuclear zone", which is highly productive and a small number of core journals belong to this zone. The second zone is moderately productive, while the third zone is less productive. Where the number of periodicals in the nucleus and succeeding zones will be 1: n: n^2, where n is a multiplier (Bradford, 1948). The observation of Bradford later described as linear relation by **Brookes** (1969) which is expressed $F(x)= a + \text{blog}x$ where, F(x) is the cumulative number of references contained in the first x most productive journals, and a and b are constant. Again, **Vickery** (1948) extended the verbal formulation of the Bradford law to show that its application in any number of zones of equal values.

Later on **Leimkuhler** (1967) has given a simple expression of Bradford law that is later known by his name and it is expressed R(r) = a log(1+br) Where, R(r) is the cumulative number of articles contributed by journals ranked 1 through r, a and b are parameters. Similarly, Brooke's derivation of journals productivity takes the form R(r) = a log (b/r). Further, **Egghe's** model (1985, 1986) Modifications for Calculating Bradford's Multiplier based on Leimkuhler's Model is $k = (e^y y_m)^{1/p}$ where e^y represents Euler's number (value 1.781), y_m represents the number of articles in the journal, and P represents the number of zones. Wilkinson (1972) noticed that the formulae provided by Leimkuhler and Brookes did not describe the same phenomenon. Several other mathematicians provided different models but the Brookes and Bradford laws, however, gained more acceptance than others.

APPLICATION OF BRADFORD'S LAW

Table-4: Scattering of journals in Bradford zone of Pathogenic research output

Zone	No. of journals	%	No. of Pubs.	Multiplier factor
1	132	1.54	51929 (33.4)	-
2	657	7.69	51822 (33.3)	4.98
3	7759	90.76	61648 (33.2)	11.81
	8548	**100**	**140222**	**8.393**

Based on the table - 4, Bradford's law is applied verbally to collected sample data. The identified zones arranged in the geometric series in the form of 1: n: n2 as given by Bradford. We found that the relationship of each zone in the present study is 132:657:7759. Here, 132 journals found in the nucleus zone and the mean value of Bradford's multiplier is n=8.4.

Therefore, 132: (132 x 8.4) : (42 x 8.4^2):: 1 : n : n^2
132 : 1108.8 : 9313.92 = 110555 – 8548 / 8548*100;
Since the percentage of error is positive here, the data is not fits the Bradford's law.

APPLICATION OF LEIMKUHLER MODEL TO VERIFY BRADFORD LAW

The Bradford law's of scattering has been widely used to study the distribution of literature around the world. The law states that there are a small number of journals in which there is a large amount of literature, forming a nucleus of core journals. There are many variants of Bradford's law, which were discussed by Leimkuhler. For the purpose of this study, we tested both Bradford's law as well as Leimkuhler's law to explain the scattering of research output in Pathogenic.

A = 140222 (Total number of articles)
y_m = 3861 (Number of items in most productive source);
T = 8548 (Total number of sources)
P = 3 (Number of zones in which the data has to be divided)
y_0= A/p = 140222/3 = 46740.67 (approx.)

In this study explains the Leimkuhler model expressed in the form of verbal formulation of Bradford law as:

$R(r)= a \log(1+br)$, r = 1, 2, 3 ----------(1)

Egghe explained Leimkuhler model as $a = Y_0/\log k$---------- (2)

$b = k-1/r_0$ ----------(3)

Where, r_0 is the number of sources in the first Bradford zone, Y_0 is the number of items in each Bradford zone and k is Bradford multiplier. R(r) is cumulative number of items produced by the sources of ranks 1,2,…..….r and a and b are the constant appearing in Leimkuhler model. In forming Bradford's group, it is shown that the number of groups p is a parameter that can be chosen freely. For calculation of the Bradford Multiplier, Egghe has given a mathematical expression as:

$k=(e^x y_m)^{1/p}$ ---------- (4)

where, g is Euler's number (e^x = 1.781); y_m indicates the number of articles published by that journal of rank one, whereas p indicates is number of zones or Bradford's groups. p = 3, e^r= 1.781, y_m = 3861.

His expression is based on, if the sources are ranked in decreasing order of productivity, then expression y_m is the number of items in the most productive source. Then, Y_m and r_0 are expressed as $Y_0 = y_m^2 \log k$ ---------(5) and $r_0 = (k-1)y_m$ -------(6)

Once one gets the value of p in Eqn. (4), the value of k can be calculated by using the following Eqn.

$k= (1.781*y_m)^{1/p}$ ----------(7)

and $Y_0 = A/p$, where A is the total number of articles. $k= (1.781*8548)^{1/3}$

Again, let T represents the total number of journals in Bradford zone, then there are r0ki-1 sources would constitute the group (i = 1, 2, 3,….p)

$T = r_0 + r_0k+r_0k^2 + \ldots\ldots\ldots\ldots + r_0k^{p-2}$ ----------(8)

So, $r_0 = T/1+k+k^2+...+k^{p-1} = T(k-1)/(k^p-1)$ ---------- (9)

From this, one is able to derive the value of A and T, r_0 and y_0 can be calculated by expression using Eqn. (7) which gives the value of p. This whole expression is tried here to justify the application of Bradford law and Leimkuhler model, Egghe's Model to test their fitness of good for the Pathogenic research output.

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: 8548; k = Bradford's constant: 19.02; P = No. of zones: 3;

$r_0= 8548(19.02-1)/(19.02^3-1) = 154034.96/6879.682808 = 22.39$ (approx.)

Thus, $k=(1.781*Y_m)^{1/p}$; $k= (e^r y_m)^{1/p} = (1.781 \times 3861)^{1/3} = (6876.441)^{1/3} = 19.02$ (approx.)

let denote the total number of journals in Bradfords' Group, there are r_0k^{i-1} sources (i = 1,2,3, ……p).

$T = r_0 + r_0k+r_0k^2+ \ldots.+r_0k^{p-2}$

$r_0 = r_0 \times 1 = 22.39 * 1 = 22.39$; $r_1 = r_0 * k = 22.39 * 19.02 = 425.8578$;
$r_2 = r_0 * k^2 = 22.39 * 19.02^2 = 8099.8154$
i.e. $1:n:n^2 = 22.39 : 425.8578 : 8099.8154 = 22 : 426 : 8100$

Percentage of error = 8548.06316 –8548/8548 *100 = 0.0037.

$a = y_0/\log k = 46740.67/\log 18.46 = 46740.67/1.27 = 36803.677$ (approx.)
$b = k\text{-}1/r_0 = 17.46/23.73 = 0.7357$

As you can see, the error percentage is negligible. Therefore, Bradford's law does not apply to the present set of data. It can be observed that the percentage error in this case is negligible, and the number of journals publishing literature has increased by 8.4, the new Bradford multiplier.

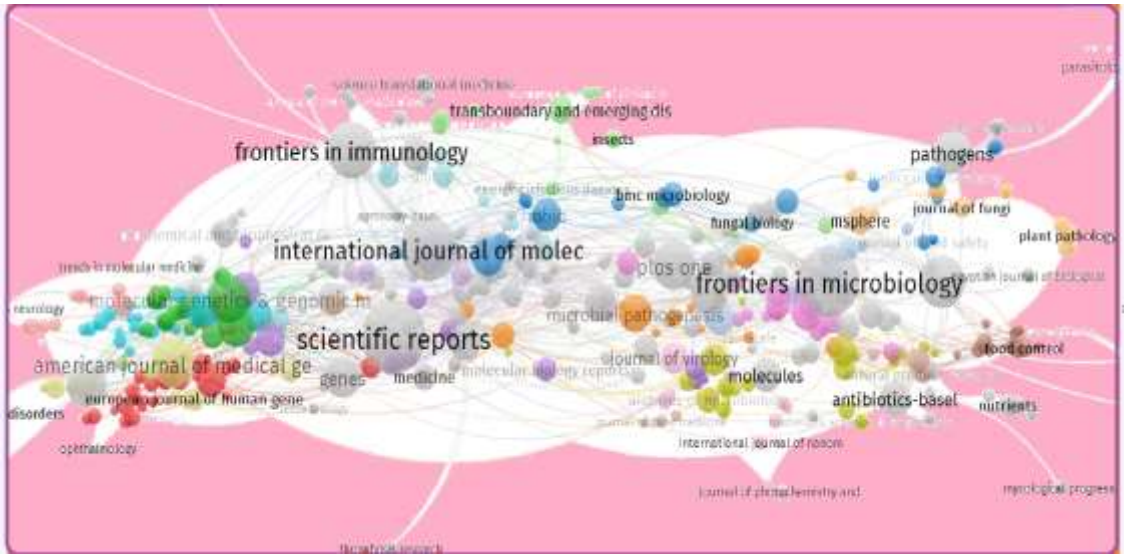
Table-5: Application of Egghe;s model for Bradford law of Pathogenic research output

Zone	No. of journals	%	Multiplier factor
1	22	1.54	-
2	426	7.69	19.4
3	8100	90.76	19.01
	8548	**100**	**19.21**

This theoretical distribution of Bradford's law can be used to evaluate whether it would be suitable for the study in question since the calculated values were 22:426:8100 (mean value of multiplier factor is 19.21) while the observed values were 132:657:7759. In the words of 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 132:657:7759 (8.39 is the mean value of multipliers. It means that 132 journals constitute the nucleus zone i.e., core journals in the field under study. Thus the percentage error is positive hence the Egghe's model also not fit for this study.

CITATION ANALYSIS BY SOURCE WISE

In this analysis is to find the highly citations sources using through VOSviewer software for citation analysis by the different sources of pathogenic research output. Out of 2883 (VOSVierwer counted) different sources, only 776 sources (journals) were meet the threshold minimum number of documents of individual sources has earned at least 5 citations. For each of the 776 sources, the total strength of the citation links with other sources will be calculated. The sources with the greatest total link strength will be selected. The all 776 sources were not connected to each other of this network. The largest set of connected items consists of 555 sources.



CONCLUSION

The journal distribution pattern of Pathogenic research literature published by global level from 2001 to 2019 follows the Bradford distribution pattern. When Egghe's model was applied to verify Bradford's law, it confirmed that the data sets not fit to Bradford's distribution pattern, although with an uneven distribution of articles across three zones. The data also indicated that the majority of Pathogenic research publications are spread across various journals with a multidisciplinary focus, primarily in social sciences and medical sciences. In essence, Pathogenic publications are not heavily concentrated in specialized Pathogenic journals; instead, they are dispersed across a diverse range of journals.

REFERENCES

1) Amsaveni, N. (2016). Assessment of Bradford law's of scattering to Neural Network literature through Bibliometric study. International Journal of Next Generation Library and Technologies, 2(3), 1-15.

2) Bradford, S.C. (1934). Sources of information on specific subjects. Engineering, 137, (3550), 85-86.

3) Brookes, B. C. (1969). Bradford's law and the bibliography of science. Nature, 224(5223), 953-56.

4) Brookes, B.C. 1977. "Theory of Bradford's Law". Journal of Documentation 33(3): 180-209.

5) Chaturbhuj, S., Batcha, M. S., Chaturbhuj, S., & Sadik Batcha, M. (2020). Application of Bradford's law of scattering and Leimkuhler model on fluid mechanics literature published during 2001-2019. Library Philosophy and Practice (e-journal), 4524.

6) Dharanikumar, P., Banateppanavar, K., Girish, T. S., & Jayaraj, A. N. (2014). Bradford's Zone to LIS Publications Published in Library Management Journal from 2010-2012: A Citation Study. Library Philosophy & Practice.

7) Egghe, L. (1986). The dual of Bradford's law. Journal of the American Society for Science and Technology, 37, 246-255

8) Egghe, L. (1990a). Applications of the theory of Bradford's law to the calculation of Leimkuhler's law and to the completion of bibliographies. Journal of the Association for Information Science and Technology, 41(7), 469-492

9) Leimkuhler, F. F. (1967). The Bradford Distribution. Journal of Documentation, 23 (3), 197-207.

10) Nicolaisen & Hjorland. 2007. "Practical Potentials of Bradford's: a critical examination of the received view." Journal of Documentation 63(3): 359- 77.

11) Rao, I.K.R. (1998). An analysis of Bradford multipliers and a model to explain law of scattering. Scientometrics, 41(1-2), 93-100.

12) Savanur, K P & Hullolli, P. (2018). Application of Bradford's law of scattering to the Solar Physics world literature: a scientometric study, Journal of Library Development (JLD), 4(1), 19-30.

13) Sudhier KG (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 J. Library Information Technol. 30(2): 3-14.

14) Vickery, B. C. (1948). Bradford law of Scattering. Journal Documentation, 4(3), 198-203.