Fall 9-10-2019

Inconsistency in Selecting Metrics used for Bibliometric Studies

Sankha Subhra Das

Inconsistency in Selecting Metrics used for Bibliometric Studies

Sankha Subhra Das
PhD Scholar, Department of Library & Information Science,
Manonmaniam Sundaranar University, Tirunelveli - 627012, Tamil Nadu, India
Tel: +91 7980000423, Email-ssdas91@gmail.com
ORCID- https://orcid.org/0000-0002-3480-6665

Abstract:

There are several metrics available for application in bibliometrics. Some of the metrics that have been considered here such as measuring growth of publications, citation impact, authorship pattern, h-index have been widely used to generate statistical analysis with respect to books, articles, publications. Now one must be aware of the pros and cons of each and every metrics used in the research. One has to be certain that there is no information that is getting lost when data about researchers and their institutions are squeezed into a tabular form of metrics. There are certain metrics that have been discussed to be replaced with other metrics to obtain more accurate interpretation of the research performance. If used otherwise it can create a hindrance to the real research performance when misused.

Keywords:

metrics, growth of publications, citation impact, authorship pattern, h-index, tabular form, the research performance.

1. Introduction:

Bibliometrics is considered as the statistical analysis for books, articles, or other publications. These analyses are used to track the performance of researcher based various parameters creating an output and impact on the research contribution. This further helps in attaining academic promotion and excellence during his tenure of contribution, as well as in generating funding and grants for projects in his area of interest. The h-index is an important author-level metric that quantifies research output by measuring the productivity of the author and measuring the impact created by his research. It was developed by J.E. Hirsch in 2005, defined as an index to quantify the individual's scientific research output. Journal-level metrics also measures the impact created by a journal in a particular field. These factors are calculated by measuring the number of articles published per year and the number of citations received by the articles published in that particular journal. The Article-level metrics also helps an author to track the citations received for his article. In similarity to author-level impact metrics, article-level metrics also generates promotion during the period of tenure, in attaining research grants.

2. Review of Literature:

Francisco and Julia (2019) concluded their study in which they found out that self-citations and first self-citations played a crucial role in the citation works increasing their visibility.

Bornmann & Haunschild (2018) represented a study on the journal impact and paper impact of a single researcher in one single graph. In most bibliometric studies, metrics for journal and paper impact are represented separately. But here, both the metrics were combined into a single graph.

Watman and Van (2012) argued about the behaviour of h-index. They concluded that h-index cannot be applied as an indicator to calculate the overall impact of a scientist. Instead he paid special attention to highly cited publication indicator. This indicator is more a less common to h-index but does not produce inconsistent ranking like h-index does.

3. Objectives of the Study:

1. To discuss Research Footprint as an alternative metrics to University Ranking.

2. To focus not only on Category Normalized Citation Impact (CNCI) but Impact profile also.

3. To consider the Journal Profile Page, instead of Journal Impact Factor (JIF).

4. To focus on Beam Plot and not only on H-index.

4. Methodology:

Several articles were reviewed to find out the misleading factors that were used by researchers as metrics for bibliometric studies. There are several metrics available but choosing a wrong metric leads to erroneous results. Hence several metrics have been suggested that can be used as an alternative to certain metrics. Although using certain metrics will give relevant results, but using appropriate metrics will give high precision results.

5. Analysis:

Various factors are considered for undergoing Bibliometric Studies but I have discussed some of the factors;

1. Growth of Publications

Now let us consider the study has been taken keeping into consideration for a period of ten years, data retrieved from Web of Science (WOS). Then the publication can be viewed on a table and comparisons can be made by observing the publication trend for each and every year in particular and making a comparative statement whether the publication tends to increase or decrease with each year and so on.

2. Citation Impact

This is an important factor to determine how often an article was cited by other sources. Now, citation analysis can be done for an individual article, academic journal and author. However citation rates depend on the discipline and the number of the people involved also. For example many scientists work in neuroscience as compared to neuroscientists publish more research papers than mathematician and are cited more than papers in mathematics. Similarly, review papers get more citations than research papers because they help in

summarizing results from several papers. The factor involved here is Average Citation Per Paper (ACPP).

$$\text{ACPP} = \frac{\text{Total Publication (TP)}}{\text{Total Citation (TC)}}$$

3. Authorship Pattern

Here we can separate the articles based on the number of authors example - single, double, three, four authors and so on. After doing this we can calculate the Collaboration Coefficient (CC) as suggested by Ajiferuke based on the counting of fractional productivity defined by Price and Beaver.

$$\text{CC} = 1 - \frac{\sum_{j=1}^{k} \left(\frac{1}{\text{j}}\right) \text{f}_\text{j}}{N}$$

It is given by following formula where,

f_j represents the " number of j authored research papers" ;
N represents "total number of research papers published" and
k defines the "greatest number of authors per paper".
The above formula by Ajiferuke, tends to state that CC will point towards zero when a single-authored paper dominate and counted [1 - $\frac{1}{j}$] then j authored papers being dominate.

4. H-index

This is a widely used factor to characterise a researchers publication and the citation profile is the h-index, created by physicist Jorge Hirsh (2005). It provides and reduces a list of publications and their citation counts to a single number. It is explained as a researcher (or group or country) with an h-index has published at least "h papers" each of which has been subsequently cited by several sources at least "h times".

5. Journal Impact Factor (JIF)

The Journal Impact Factor (JIF) developed by Eugene Garfield, founder of the Institute for Scientific Information Garfield (1955) who raised the idea of publication "impact" and created a "journal" impact factor (Garfield and Sher,1963) to help in selection of journals for the Science Citation Index (SCI).

for example, Nature had an Journal Impact Factor (JIF) of 41.456 in 2014.

$$\text{JIF}_{2014} = \frac{\text{Citations}_{2013} + \text{Citations}_{2012}}{\text{Publications}_{2013} + \text{Publications}_{2012}}$$

$$= \frac{29753 + 41924}{860 + 869}$$

$$= 41.456$$

6. University Ranking

Comparisons are made between similar and established, multi-faculty Universities having large medical schools, like Edinburg (29^{th}) and Manchester (57^{th}). These positions do they really mean anything.

The reality on the basis of which these institutes are ranked may differ as often. The institutes score better on some parameters and less on other whereby the position varies in comparison to others. A University ranked worldwide is a peculiar task, but it acts as a reference point. Most of the actual facts are hidden such as short-listing for students.

6. Interpretations:

1. Considering Beam Plot, instead of H-index

The Beam Plot represents a single picture of a researchers output, reflecting how it varies throughout the year and evolves over time. The percentile usage that creates citation impact, which is highly skewed, varies with discipline and time period since publication. Hence by reducing this to a single value of h-index may create a summary but signifies nothing that can be used for evaluation.

2. Opting for Journal Profile Page and not just the Journal Impact Factor (JIF)

Journal Impact Factor (JIF) creates lot of misinterpretation. It is not about the evaluation of research but solely depends on Journal Management. Hence considering Journal Impact Factor (JIF) as the single point value creates an understanding that shows that Journal Impact Factor (JIF) represents a wide range of performance at article level. Journal Impact Factor (JIF) may be considered as guide but the entire context is required for naive information outside the publishing house.

3. Consulting Impact profile, instead of isolated Category Normalized Citation Impact (CNCI)

Category Normalized Citation Impact (CNCI) values can also be misleading because the data is spread at individual and journal level, which is highly skewed and subjected to other values. The Impact skewed and subjected to other values. The Impact profile represents the data into a digestible form and shows the underlying distribution. It represents that the data for a world average and institutional average means that many articles are cited more whereas others are less often cited.

4. Analysing Research Footprint and not only on University Ranking

The University ranking tends to conceal major information than most analysis. Whereas on the other hand the Research Footprint can highlight performance on the basis of discipline and data types comparison between two institutions or countries can be done on the basis of certain criteria. But, there cannot be a sensitive way to compare two research bodies on the basis of ranking or figures.

7. Conclusion:

There are several metrics on the basis of which one can analyse the data obtained from various databases. Nowadays there are ample amount of data that can be obtained from various databases but one must also be certain about the various metrics that suits the interest of the researcher. However one should be certain that the formula selected for interpreting the obtained data is clearly defining the researchers interest. Finally, I would like to conclude by saying that all reports are potentially informative but they also suffer from widespread misinterpretation and misuse. Based on these analysis major industries, several Librarians and even policy makers undertake decisions. This type of research also helps in promotion and tenure, as well as receiving funding and grants for future projects. Hence, one must be certain and careful about the analysis of data in an appropriate manner.

8. Reference:

1. Ajiferuke I, Burrel Q and Tague J. 1988. Collaborative coefficient: A Single measure of the degree of collaboration in research. Scientometrics, Vol. 14, no.5-6: 421-33. DOI: 10.1007/BF02017100.

2. Bornmann, L & Haunschild, R. (2018). Plots for visualizing paper impact & Journal impact of Single Researchers in a Single Plot. Scientometrics, 115, 385-394. https://doi.org/10.1007/S11192-018-2658-1.

3. De Solla Price, D.J. (1963). Little Science, Big Science. Columbia University Press.

4. De Solla Price, D and Beaver, DB. 1966. Collaboration in an invisible college. American Psychologist Vol.21, no.11: 1011-18. DOI: 10.1037/h0024051.

5. Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. Science, 122, 108-111.

6. Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. American Documentation, 14, 195-201.

7. Garfield, E. (2006). The History and Meaning of the Journal Impact Factor. Journal of the American Medical Association (JAMA), 293: 90-93, January 2006.

8. Gonzala-Sala, F., Osca-Lluch, J. & Haba-Osca, J. Scientometrics (2019). https://doi.org/10.1007/s11192-019-03101-3

9. Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. PNAS, 102, 16569–72.

10. Larsen, P.O; Von Ins, M. (2010). "The rate of growth in Scientific publication and the decline in coverage provided by science citation Index". Scientometrics. 84(3): 575-603. doi:10.1007/s11192-010-0202-z.

11. "Nature". 014 Journal Citation Reports. Web of Science (Science ed.) Thomson Reuters. 2015.

12. Watman, L., & Van Eck, N.J. (2012). The inconsistency of the h-index. Journal of the American Society for Information Science and Technology, 63(2), 406-415.

About the Author

Sankha Subhra Das is a PhD Research Scholar at the Manonmaniam Sundaranar University, Dept. of Library and Information Science, Tirunelveli - 627012, Tamil Nadu, India. He is a UGC-NET JRF qualified full time Scholar in discipline of Library and Information Science. He received his MLISc (Masters in Library & Information Science) from University of Kolkata. He has published several papers in National and International journals. His research interests are Bibliometrics, Knowledge Organization, Scientometrics, User Studies, Digital Library& LIS Education.