

Assessing and Evaluating Research Performance of an Institution using Bibliometric Indicators

Jagdish Arora and Kruti Trivedi

The author of the article, Jagdish Arora, Director of the INFLIBNET Centre and Kruti J Trivedi Scientist, B(LS), INFLIBNET Centre explains methods and metrics for assessing and evaluating research performance of an institution using Bibliometric Indicators. The paper highlights some of the underlying principles and metrics that are used for evaluation of research performance. It also provides background information to assist the universities / institutions to conduct evaluation of research performance. Ms. Kruti J Trivedi is looking after N-LIST Programme and UGC-INFONET Digital Library Consortium. She is also involved in measuring research output of member universities. Dr. Jagdish Arora heads the institution. Both can be contacted kruti@inflibnet.ac.in and director@inflibnet.ac.in.

1. Introduction

The generation and transmission of knowledge through research has long been recognized as an essential requirement for a country's long-term growth and competitiveness as well as for creating capacity to solve social problems (World Bank, 1998). Considering the fact that the benefit of increasing research output and potentiality to increase economic growth, many governments, regulatory bodies and funding agencies have increased research funding to universities. Since funding for research to a university should logically commensurate with its research output, there is a need to identify performance indicators for qualitative evaluation of research output. These performance indicators are being used as a tool to find out research performance of a university as well as to determine funding allocations. The main focal points of research evaluation are productivity, quality, impact, and utility of research output in terms of generating technological, economical or social benefits. Qualitative parameters and performance indicators are being evolved to measure all these different aspects of research evaluation. As regards to assessment methods, the literature on research performance shows that bibliometric analysis and peer review are the main approaches for assessing the quality and impact of research. In this paper, we outline the need for university research assessment, types of performance indicators and metrics that can be used, particularly in relation to research within universities.

2. Need for Assessments of Research Output

There are a number of reasons to justify assessments of government funded research in present era. The first and foremost reason is ever growing cost of conducting planned research that includes cost of subscription to peer review literature, scientific instrumentation and infrastructure that are required to conduct research. As universities spend huge amount to set-up research facilities, it becomes necessary for the university to evaluate status of their comparative research performance. Moreover, on an average Government funding for R&D is increased 5% to 10% every year. Performance indicators are used not only to identify new areas of research and researchers for fund allocation but also to identify declining areas of research. Furthermore, evaluation of research performance is also an issue of social responsibility and accountability. Research evaluation assures Government and public that their contributions are being well spent. Universities are expected to become more efficient in their use of public resources and more accountable (Massy, 1996). In US, the Government Performance and Results Act was passed in 1993, which requires federal agencies to establish strategic planning and performance measurement. This, in turn, requires the establishment of performance goals and performance indicators to assess output, service level and outcome. Research assessment also enable research funders and policy makers to make better decisions about funding, ranking, awards, promotion, benchmarking, etc. There are three reputed

globally-recognized rankings of the world universities, namely i) Shanghai Jiao Tong University's Academic Ranking of World University; ii) Academic ranking of world universities compiled by the QS World University Ranking; and iii) Times Higher Education World University Ranking. All these three “rankings” have used different sets of indicators and parameters for ranking the universities. Apart from all other parameters, research performance is one of the major and high weightage indicators to identify ranking of World class University. It is, therefore, important for a university to evaluate their research strengths and weaknesses and prioritize research areas to formulate a strategy for future work

3. Bibliometric Indicators

Bibliometric indicators are generally used to assess the status of research performance across institutions and geographical regions/countries. It addresses the issues such as: (i) Is a University contributing more or less to research output in a particular field or sub-field in comparison to other universities? (ii) Is it performing better than others?, and (iii) Is it contributing more research output in a particular area compared to others? Two types of indicators, namely absolute and relative indicators are generally used for bibliometric analysis. Publications count, journals count, and citations count, etc. are the examples of absolute indicators. These are used frequently for bibliometric analysis. On the other hand, activity index (AI), and relative citation Index (RCI) are examples of relative indicators, which is used to assess the strengths and weaknesses of a university on relative basis. For example, activity index of a university, e.g. University of Delhi, in a particular subject field, e.g. Chemical Sciences, can be used to compare publication activity against average publication activity of the country, which is taken as the benchmark for comparing publications activity of various universities. It can also be used to identify research priorities of different countries or regions within the country. It also helps to understand the strength of the university in various subjects (NISTADS, 2006). Various indicators used for measuring the publications and citations output of the university, institute, geographical region or country across disciplines are described below..

4. Measuring Growth of Research Publication of an Institute

Number of research papers published in peer-reviewed journals is a reasonable measure of research productivity. It provides an estimate of the productivity of research and related knowledge generation. The indicator defines the level of scientific progress made by an individual, institute, group, university or country.

Publication (P):

Total number of papers published by the research institutes for a given period. Number of Publications (P) can further be used to determine:

► AGR (Annual Growth Rate):

AGR provides average increase in number of publications over a period of 5, 10 or more years. This metric is a used to measure the growth of publications in different subject areas over a period of time for a given university or set of universities. AGR is calculated using the following formulae :

$$\text{AGR} = \frac{\text{Current Decade Total} - \text{Previous Decade Total}}{\text{Previous Decade Total}} * 100 / \text{Number of Years}$$

❖ For Example:

Years	No. of Publications
1971-1980	= 429
1981-1990	= 706
1991-2000	= 1031
2001-2010	= 1391

$$\text{AGR} = \frac{(706-429)}{429} * 100 / 10 = \frac{277}{429} * 100 / 10 = \frac{64.57}{10} = 6.46\%$$

$$\text{AGR} = \frac{(1031-706)}{706} * 100 / 10 = \frac{325}{706} * 100 / 10 = \frac{46.03}{10} = 4.60\%$$

$$\text{AGR} = \frac{(1391-1031)}{1031} * 100 / 10 = \frac{360}{1031} * 100 / 10 = \frac{34.92}{10} = 3.49\%$$

► **AAGR (Average Annual Growth Rate):**

Average value of Annual Growth Rate (AAGR) for different period is calculated using the formulae given below.

(Growth Rate in Period A + Growth Rate in Period B + Growth Rate in Period C +
... Growth Rate in Period X)

$$\bullet \text{ AAGR} = \frac{\text{Growth Rate in Period A} + \text{Growth Rate in Period B} + \text{Growth Rate in Period C} + \dots + \text{Growth Rate in Period X}}{\text{Number of Periods}}$$

❖ For Example:

$$\text{AAGR} = \frac{6.46\% + 4.60\% + 3.49\%}{3} = \frac{14.55}{3} = 4.85\%$$

5. Measuring Impact of Research Output

Impact of a research article can be measured by the number of citations received by it. Likewise, impact of a research institute can be measured in terms of cumulative number of citations received by all research publications authored by its faculty and researchers. The area of study that deals with evaluation and interpretation of number, frequency and pattern of citation reviewed by articles, scientist, universities, counties, etc is called citation analysis. Aim of citation analysis is to estimate impact made by publication of a scientists, institution or a country. A citation is an abbreviated alphabetic expression embedded in the body of an intellectual work that denotes an entry in the bibliographic reference section of the work for the purpose of acknowledging the relevance of the works (Wiki). The citations provide an idea about the utilization of published knowledge. It is considered that frequently cited paper have greater impact to influence subsequent research activities than a paper with no citations. Citation analysis studies that can be conducted to measure performance of an institute is as follows:

► **Citations with Self Citations:**

The number of citations of all articles of a university including self-citations (publications of a university citing themselves).

► **Real Average Citations:**

This metrics is used to assess quality of research publications across different subject areas, universities, and geographical regions, etc. The formulae for calculating Real Average Citations is:

$$\text{Real Average Citations} = \frac{\text{Total citations for the year or selected subject areas}}{\text{No. of papers having at least 1 citation}}$$

❖ For Example :

Out of 3,557 publications of a university, 2753 publications have received 24165 citations, in this case Real Average Citations is:

$$\text{Real Average Citations: } \frac{24165}{2753} = 8.78$$

► **Cited rate:**

Cited rate provides percentage of publications that have been cited at least once during pre-defined time period. The cited rate is calculated using following formulae:

$$\text{Cited Rate : } \frac{\text{Number of Papers having at least 1 Citation}}{\text{Total Number of Papers Published}} * 100$$

❖ For Example :

Out of 3,557 publications of a university, 2,753 publications have received 24,165 citations, in this case cited rate is:

$$\text{Cited rate} = \frac{2753}{3557} * 100 = 0.7739 * 100 = 7.40\%$$

► **Percentage of Publications not Cited (PNC) :** This metric provides percentage of publications not cited during a specified time period. PNC is calculated using following formulae:

$$\text{PNC} = \frac{(\text{Total Number of Papers Published} - \text{Number of Papers having at least 1 Citation})}{\text{Total Number of Papers Published}} * 100$$

❖ For Example:

Out of 3,557 publications of a university, 2,753 publications have received 24,165 citations, in this case percentage non-cited publication is:

$$\text{PNC} = \frac{(3557-2753)}{3557} * 100 = \frac{804}{3557} * 100 = 0.2260 * 100 = 22.60\%$$

➤ **H Index:**

The Hirsch index, or H-index, is a distribution-based indicator that corresponds to the number of papers at or above a given citation level equal to the value of the citation. This metric reflects the number of papers (N) in a given dataset having N or more citations (Thomson Reuters, 2008). For example, an H Index of 77, indicates that 77 papers in the given set were cited at least 77 times each. This metric attempts to reflect both productivity (number of papers) and impact (number of citations) by a single number. This metric is useful because it discounts the disproportionate weight of highly cited papers, or papers that have not yet been cited.

❖ For Example:

Number of Publications	Number of Citations Received (Highest to Lowest)
Paper -1	105
Paper -2	89
Paper -3	50
Paper -4	20
Paper -5	12
Paper -6	10
Paper -7	7
Paper -8	4
Paper -9	3

Data Sorted by Number of Citations Received in Decreasing Order

H Index

6 Analyzing Active and Passive Areas of Research

➤ **Relative Specialization Index:**

The Relative Specialization Index (RSI) compares the shares of subject disciplines in university's total publications to the overall shares of each subject discipline in world's total publications. The RSI is a relative indicator which is based upon the Activity Index (AI). The formulae for calculating Activity Index is:

$$AI = \frac{\text{The share of given field in the publications of the given university}}{\text{The share of the given field in the world total of publications}}$$

The RSI is then calculated using Activity Index as follows:

$$RSI = \frac{AI - 1}{AI + 1}$$

❖ For Example:

If, Total Number of Publications of the World in Engineering and Technology = 53,68,118 and
Total number of Publication of a University in Engineering and Technology = 3557

$$AI = \frac{3557}{53,68,118} = 0.0007$$

Now,

$$RSI = \frac{0.0007-1}{0.0007+1} = \frac{-0.9993}{1.0007} = -0.9987$$

RSI takes its values in the range -1 to < 1. The value indicates whether a university has a higher-than-average activity in the world in a scientific field (RSI >0) or a lower-than-average activity (RSI <0). RSI = 0 reflects a completely balanced situation. It also reflects a certain internal balance among the fields at the given university; positive RSI values must always be balanced by negative ones, as no university can have only positive RSI values.

7 Measuring Collaboration

Research collaboration can be defined as working together of researchers to achieve the common goal of producing new scientific knowledge. Research collaboration requires work in partnership with other research investigators including government departments and agencies, universities and colleges, and industry. It extends across disciplines and organizational boundaries. Metrics of this indicator are used to measure share of collaborative activity of university and compare it with other institutes, country and world.

Collaboration Index is a relative measure of the collaborative activity of an institution. It is measured using following formulae:

$$\text{Collaboration Index} = \frac{\text{(Share of collaborated papers by an institution - in its total publications output)}}{\text{(Share of total collaborated papers by the country - in its total publications output)}}$$

❖ For Example:

If, university collaborates 1260 papers out of 5,286 total publications in area of medical and health sciences and India collaborates 8569 papers out of 16085 total publications, in that case Collaboration Index is:

$$\text{Collaboration Index} = \frac{23.83}{53.27} = 0.44$$

International Collaboration Index is a relative measure of the collaborative activity of an institution at international level. It is calculated using following formulae:

$$\text{Int. Collaboration Index} = \frac{\text{(Share of the int. collaborative papers by institution in its total publications output)}}{\text{(Share of the int. collaborative output by country in its total publications output)}}$$

❖ For Example:

If, university collaborates 750 papers out of 5,286 total publications in area of medical and health sciences and India collaborates 4587 papers out of 16085 total publications, in that case Int. Collaboration Index is:

$$\text{Int. Collaboration Index} = \frac{14.19}{28.52} = 0.49$$

8. Measuring Correlation

The correlation analysis, a statistical technique, is deployed for measuring the strength of association between two random variables. Two sets of data are said to be correlated when any change in one set of data accompany by corresponding change in other. The relative degree of intensity of association between two variables is measured by correlation coefficient. The correlation coefficient is a ratio that expresses the extent to which the two variables are accompanied by the change in concerned variable. The value of correlation coefficient ranges from -1 to +1. When there is a perfect positive correlation, the value of correlation coefficient is equal to +1 and it is -1 in case of perfect negative correlation.

There are a number of methods to calculate correlation coefficients can be calculated. The two most commonly used are Pearson's correlation coefficient and Spearman's rank correlation coefficient. Pearson's correlation coefficient requires both variables to be measured on an interval or ratio scale, and the calculation is based on the actual values. Spearman's rank correlation coefficient requires data that are at least ordinal and the calculation, which is the same as for Pearson's correlation, is carried out on the ranks of the data. Each variable is ranked separately by putting the values of the variable in order and numbering them: the lowest value is given rank 1, the next lowest is given rank 2 and so on. If two data values for the variable are the same they are given averaged ranks.

8.1 Spearman's Rank

Formulae :

Year	Faculty Member (X)	Number of Articles (Y)	Rank x	Rank y	d	d ²
2006	3506	1164	1	3	-2	4
2007	3570	1130	2	2	0	0
2005	3628	1093	3	1	2	4
2008	3722	1330	4	4	0	0
2009	3769	1553	5	5	0	0

$$r = 1 - \frac{6 \sum d^2}{n(n^2-1)}$$

$$r = 1 - \frac{48}{5(25-1)}$$

$$r = 1 - \frac{48}{5(24)}$$

$$r = 1 - 0.4$$

$$r = 0.60$$

8.2 Person's Rank

❖ Formulae :

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

❖ For Example

Year	Number of Downloads (X)	Total Publications (Y)	(X - \bar{X})	(Y - \bar{Y})	(X - \bar{X})*(Y - \bar{Y})	(X - \bar{X}) ²	(Y - \bar{Y}) ²
2005	11189	126	-25030.2	-82.1667	2056645	626509243.4	6751.361
2006	17268	156	-18951.2	-52.1667	988619.2	359146718	2721.361
2007	23797	174	-12422.2	-34.1667	424424	154310224.7	1167.361
2008	44281	248	8061.833	39.83333	321129.7	64993156.69	1586.694
2009	60999	273	24779.83	64.83333	1606559	614040140	4203.361
2010	59781	272	23561.83	63.83333	1504030	555159990	4074.694
Total	217315	1249			6901408	2374159473	20504.83

$$r = \frac{6901408}{\sqrt{2374159473 * 20504.83}} = \frac{6901408}{6977230.417} = 0.99$$

$$r = 0.99$$

9. Summary

Evaluation and measurement of research performance is an important input for policy maker that require transparent and objective evaluation tools and techniques. The objective of this paper is to provide background information to assist the universities / institutions to conduct evaluation of research performance. The paper highlights some of the underlying principles and metrics that are used for evaluation of research performance. First part of the paper indicates needs of assessment of an institute using using of bibliometric indicators for measuring research performance of institute. Paper proposes various formulae and examples related to these indicators. While, this paper restricts itself to enlisting bibliometric indicators for measuring research productivity, there are a number of other parameters that are using for evaluating research performance of an individual or an institute. These indicators include i) institutional strengthening, ii) funds allocation iii) research infrastructure iv) quality assurance v) honours and awards received, etc. A separate paper will be published highlighting all these performance indicators used for performance evaluation of a university in forthcoming issue of this newsletter.

10. References

Aalto University Bibliometric Report 2003 – 2007, 37p.

Available at: www.aalto.fi/.../aalto_university_bibliometric_report_2003-2007.pdf

Arora, J. And Trivedi, Kruti (2012). Measuring Research Output and Its Impact Using Bibliometric Indicators: Profiles of 50 selected Universities. Ahmedabad: INFLIBNET, p. 357.

Geuna, Aldo and Martin, Ben R. (2001). University Research Evaluation and Funding: An International Comparison, Science and Technology Policy Research. Brighton: Science and Technology Policy Research, University of Sussex, p. 45. Available at: www.sussex.ac.uk/Units/spru/publications/imprint/.../sewp71.pdf

Gupta, B.M. and Dhawan, S.M. (2006). Measures of Progress of Science in India: An Analysis of the Publication Output in Science and Technology. New Delhi: Institute of Science, Technology and Development Studies, p.196. Available at: psa.gov.in/writereaddata/11913286541_MPSI.pdf

Martin, B. R. (1996). The Use of Multiple Indicators in the Assessment of Basic Research. *Scientometrics*, Vol. 36, No. 3, p. 343-362.

Meek, V. Lynn and Van der Lee, Jeannet J. (2005). Performance Indicators for Assessing and Benchmarking Research Capacities in Universities: Background Paper prepared for the Global University Network for Innovation - Asia and the Pacific. Australia: Apeid, Centre for Higher Education Management and Policy, University of New England, p.49. Available at: unesdoc.unesco.org/images/0015/001529/152960e.pdf

Noyons, ECM (2008). Bibliometric Study of Lund University, 2002 – 2007. The Netherlands, Centre for Science and Technology Studies, p.40. Available at: www.lu.se/upload/.../Lund_2008_report_bibliometric_study.pdf

Thomson Reuters (2008). Using Bibliometrics: A Guide to Evaluating Research Performance with Citation Data. White Paper, p. 12. Available at: www.wrclib.noaa.gov/lib/files/Bibliom_whitepaper.pdf

Wikipedia (Accessed 28th November 2012)
Available at: <http://en.wikipedia.org/wiki/Citation>

