# H-Index: The key to research output assessment

Hasan, Syed Akif and Subhani, Muhammad Imtiaz and Osman, Ms. Amber

Iqra University Research Centre (IURC), Iqra university Main Campus Karachi, Pakistan, IQRA UNIVERSITY

2012

# Published in

## Archives Des Sciences (ADS), 2012

# H-Index: The key to research output assessment

Muhammad Imtiaz Subhani[1], Syed Akif Hasan[2],

Amber Osman[3]

[1]Iqra University Research Centre-IURC , Iqra University- IU,

Defence View, Shaheed-e-Millat Road (Ext.), Karachi-75500, Pakistan

E-mail: drsubhani@yahoo.co.uk Tel: (92-21) 111-264-264 (Ext. 2010); Fax: (92-21) 35894806

[2] Office of Vice President, Iqra University- IU,

Defence View, Shaheed-e-Millat Road (Ext.), *Karachi-75500, Pakistan*

E-mail: hasandrakif@gmail.com Tel: (92-21) 111-264-264 (Ext. 1513); Fax: (92-21) 35894806

[3] Iqra University Research Centre-IURC , Iqra University- IU,

Defence View, Shaheed-e-Millat Road (Ext.) Karachi-75500, Pakistan E-mail: amber.osman@yahoo.com

Tel: (92-21) 111-264-264 (Ext. 2024); Fax: (92-21) 35894806

**Abstract**

The capacity of research has increased enormously with number of working and published research papers, research journals, research repositories, indexing and abstracting agencies and research scholars/authors proffering varied knowledge. Every field has its own way of monitoring and evaluation and similarly so does the research field. There are multiple ways of assessing the research journals, published papers and authors in order to classify research output time to time and also to acknowledge and recognize the new knowledge and information creation by the authors. Narrowing our research domain to research ranking then just few years ago, H-Index has developed its importance and use to justify the authors profile, research papers and journals in evaluating the research quality of an author or may it be a Research Journal or even universities. This paper describes the H-Index default purpose, accuracy, assessment of how it assesses the research work of author/ journals/ Universities and for which of them it matters much. The findings revealed very frankly that h-index is comparatively a better method to rank researchers only, while, interesting to foretell if any other competing ranking method is more better or a new invention it's on it way. Let's dig in more to have a brighter enlightenment than H-Index.

**Keywords:** H-index, research rankings, authors, journals, universities

## 1. Introduction

In the research arena as well, it is important to review the quality of research transmission (Subhani, Hasan & Osman, 2011). Index is a concept that is used to study and find out the researcher's output in a span of time in

reference to the research papers published globally. The index should be intelligent enough to include the significant factors to evaluate the quality of research through various aspects, which will be mentioned later in this paper.

In various states of affairs, it is essential to review the quality of research circulation. To measure the excellence of research formed by a researcher, index is needed which is able to enumerate the quality of papers created by the researcher/author. This comes as a challenge from the fact and studies that an index should be suitable in the intelligence that it should be able to integrate several other things beyond purely an authorship such as individual inputs in a mutual effort.

Another addition to the research ranking evaluation of authors and journals is H-Index. The H-index is an instrument to compute the researcher's publication impact on the basis of citations count in his/her paper.

By the Alphabet "H", Hirsch, is the originator of this index in year 2005. The sole purpose of this development is to calculate the impact of a particular researcher, which is stability between the volume of research done and the researcher's value output. The real significance of "H" has been set as the largest number therefore the h papers can obtain at least h citations each. In this manner, the assessment and impact of research papers, journals and authors can be calculated simply and is also is means of acknowledgement of the researchers/journals in their research endeavor.

The computation behind H-Index is grounded on the number of citations of a research paper and of a particular author. Now, for a particular author, all the citations counts will be taken of all published papers and H-Index only considers the highest number of citations of the papers will be selected. In this way, the papers, which have received fewer citations, will be negligible in this calculation.

Similarly, just like the considering the only highest citation count, it doesn't attain the number of authors of a paper and their individual input in the research paper and the one selected author is assessed as the lone author of h papers, which is having the least h citations (Hirsch, 2007; Abbas, 2010).

Other renowned softwares for ranking the research by authors, papers, journals are done by ISI Web of Science database, Scopus Hub and multiple ranking softwares by these ranking agencies. Repec has also different ranking methods using their own research repository. For instance, ISI announces its own Citation Report based on their ranking software by the sum number of papers Np, sum number of citations Nc, Citations per paper nc – Nc/Np and the h index (Mingers, 2009; Geurin, 2010).

It has been inference that, the *h* index trails just about a linear conduct with time, and the sum of citations is generally quadratic with point in time.

To ensure the probability that assigning a higher weightage to greatly cited papers could possibly improve the predictive power of the *h* index, this expression has been assessed:

$$h_a \equiv \sqrt{h^2 + \alpha N_c}$$

This equation has been rewritten using the relationship $N_c = ah^2$,

$$h_a = h\sqrt{1 + \alpha a}$$

Cronin and Meho (2006) mentioned that h index can be calculated directly from the cumulative distribution function $F_c(x) = \Pr\{X > x\}$ of $X$, with two different methods.

1. **Expected value of h-index or single-moment method (Method 1)**

This method gives an approximation of the mean value of h-index:

$$F_c(Y_k) = P_r\{X > Y_k\} \approx K/N_T, \ (k = h)$$

2. **Probability distribution of h-index or all-moments method (Method 2)**

This method provides the probability distribution of the h-index:

$$P_r\{h=n,\varepsilon\} = \binom{N}{n} [F_c(n)]^n [1-F_c(n+\varepsilon)]^{(N_T-n)}, \; n= \{1, 2,\ldots., N_T\}, \; \forall \varepsilon \in [0,1)$$

(Bletsas & Sahalos, 2009)

## 2. Literature Review

Previously, the function of h-index was specifically to gauge author's published research work's productiveness. At the surface as mentioned earlier, this index basically checks all the research papers of an author and gives the result not less than h citations for each paper. The calculation is based on the publication/citation information i.e. papers with 0,1,2,3 etc. citations.

It is already known that the h-index, alongside with other bibliometric indicators often has an impact on critical judgments. However, it is very much possible that authors may use the identical procedure for rating departments and researchers, given that the bibliometric ratings of departments almost hold comparable structure (Bouyssou & Marchant, 2011).

In order to rule out the inconsistencies of h-index, suitable ascending of the h-index is proposed, based onto its probability distribution, which is estimated for any implicit citation distribution (Bletsas & Sahalos, 2009). Further, if or not h-index may be used to predict the future achievements, a particular contribution is noticed while assessing the individual scientific achievements of researcher using h-index as a predictive tool (Hirsch, 2007). For a ranking to be consistent, it is required that it be independent as well. Consistency is a very forcing condition. Any ranking violating consistency would be very suspicious. Independence, however, is relatively somewhat forcing but not towards the equivalent extent. Under totality, consistency suggests independence, thereby strengthening the beauty of independence (Bouyssou & Marchant, 2011).

*2.1 Functionality and Advantages of H-index*

The h-index is a practical index of technical worth that can be valuably utilized to contribute academic appointment actions and to disperse research study resources (Hirsch, 2007). This index was initially and formerly introduced in the physical sciences and then it was later on extended to the social sciences. Now we can see that the h-index is being used a productive method to rank individuals, journals and research topics (Guerin, 2010).

It's fascinating that the h-index is able to foresee the productivity of the papers published. It is considered that the superiority of h-index equated with total citations as a forecasting tool is because of co-author ship concern (Hirsch, 2007).

It has been confirmed that amount of publications and mean number of citations per publication area is required for significant h-index rankings, exception as of the assessed h-index of every researcher. If totality and consistency are supported, under few additional reasonable conditions, the only probable rankings of researchers and departments are averaging or scoring rules (Bouyssou & Marchant, 2011).

Expected value and standard deviation of an h-index can be used as a basis for ascending and grading procedures, which can be gathered through the probability distribution of the h-index (Bletsas & Sahalos, 2009).

H-index favors authors with a series of influential papers i.e. the output of a research via number of publications and citation counts in a combination, giving a huge impact of the cumulative research contributions (Bornmann, Mutz, & Daniel, 2008).

H-index is a simple tool and doesn't require any procedural skills. The authors/researchers with greatest citations in their publications would have a higher h-index, which means it basically base the citations as the calculation of a researcher's publication (Norris & Oppenheim, 2010).

In some cases, there is the similarity between the software features of reference-elaborated databases and the h-index. As H-index output relies on the citation counts, regulating to increase the h-index, one can start focusing on self-citations. This is possible by keeping an eye on one's own citations of the published papers and self-cite in another accepted paper accordingly. It is also possible that a researcher handling stray references. This will lead the researcher towards a considerably lower h-index then what he/she actually deserves. The h-index clearly stresses on the completeness of the research papers. If the papers are properly written and structured and consists of self-citations, then the h-index is bound to improve (Jacso´, 2008).

The Hirsch's h-index is most likely to be considered as a alternate to other bibliometric indicators, such as the average number of citations, the number of publications and the sum of all citations is defined that has been defined as "An author has index $h$ if $h$ of his or her $n_p$ papers have at least $h$ citations each and the other $(n_p - h)$ papers have $\leq h$ citations each" (Hirsch, 2007).

For e.g an h-index of 5 shows an author has published five papers and each of them has at least 5 citations. If there is an h-index of 0 then it does not positively give out that the author is idle and inactive, rather it shows that the author have published some papers but none of them have been cited and thus the h-index tends to be zero.

Another technique of computing the h-index is by retrieving and recuperating all items sources of a particular journal from a given year and then they will be sorted by the number of times cited. Hence, the h-index cannot be higher than the number of research papers that are published at a particular time. The journals, which publish merely a small amount of highly cited papers should not be integrated in the ranking list which is based on the h-index. This phenomenon is specifically concerned with the journals that publish reviews.

Another method of calculating the h-index is by using this formula h=c(h), where h is the number of papers that the author have published and c(h) is the number of citations papers have received. When both the sides are equal, we derive the h-index.

Another method of calculating h-index is given below.

*h-index = no. of papers = citations per paper*

For instance, there are two authors: author A and author B. Author A has published four research papers, each of which has been cited four times. Author B has published three papers, each of which has been cited six times.

**Author A**

| No. of Papers | Citations per Paper |
|---------------|---------------------|
| 1 | 4 |
| 2 | 4 |
| 3 | 4 |
| 4 | 4 |

Table 1: Author A citations counts per paper

**Author B**

| No. of Papers | Citations per Paper |
|---------------|---------------------|
| 1 | 6 |
| 2 | 6 |
| 3 | 6 |

Table 2: Author B citations counts per paper

After looking at the formula, we can say that the h-index of author A is 4, since four of the total four research papers have been cited at least four times each, while the h-index of author B is 3 because three of the total three papers have been cited at least three times each. Though author B has six citations per paper, but the number of papers is three. Thus author B has an h-index of 3.

Now, let us consider that both the authors have published one more paper which has been cited six times.

**Author A**

| No. of Papers | Citations per Paper |
|---------------|---------------------|
| 1 | 4 |
| 2 | 4 |
| 3 | 4 |
| 4 | 4 |
| 5 | 6 |

Table 3: Author A citations counts per paper

**Author B**

| No. of Papers | Citations per Paper |
|---|---|
| 1 | 6 |
| 2 | 6 |
| 3 | 6 |
| 4 | 6 |

Table 4: Author B citations counts per paper

The new h-index of author A remains 4, since four of the total five papers have been cited at least four times. The h-index of author B, however, has increased to 4 because four of the total four papers have been cited four times each.

Suppose that both the authors publish another paper which receives six citations.

**Author A**

| No. of Papers | Citations per Paper |
|---|---|
| 1 | 4 |
| 2 | 4 |
| 3 | 4 |
| 4 | 4 |
| 5 | 6 |
| 6 | 6 |

Table 5: Author A citations counts per paper

**Author B**

| No. of Papers | Citations per Paper |
|---|---|
| 1 | 6 |
| 2 | 6 |
| 3 | 6 |
| 4 | 6 |
| 5 | 6 |

Table 6: Author B citations counts per paper

Here also the h-index of author A remains 4, since four of the total six papers have been cited at least four times. But since five of author B's total five papers have received at least five citations each, their h-index increases to 5.

*2.2 Limitations of H-index*

There are some limitations of each ranking method, for instance the endogenous coverage of co authorship or of multiple affiliations. To analyze this, the problem has been addressed previously and that it has been determined to share publications with some authors amongst them as asserted by some scheme can be a remedy (Bouyssou & Marchant, 2011).
The h-index is said to have certain flaws such as it is unresponsive to the changes seen in performance. Also, the h-index is only inadequately susceptible to the amount of citations acknowledged and it never decreases. Apart from this, one can observe that when the h-index is in its original setting then it can easily put all the newcomers in a serious disadvantage as the observed citation rates and the publication output, both will be quite low. Also the number of co-authors present can and will also and influence the amount of citations received. For accuracy and better results, it's practical to use various indicators to measure and evaluate research performance in order to generate and acquire authentic and relevant results (Bornmann, Mutz, & Daniel, 2008).

More limitations of h-index is that the calculation of h-index depends critically on the database in use and due to this reason like common names and different spellings, flaws are visible. Another issue is that some research materials are in high regards of acceptance that they do not need to be cited and due to that h-index does not reflect results proportionate to contributions all the times. Also, by being a member of a large prolific group can increase your h-index because h-index counts number of citations and do not consider where your place is in the author list. H-index ignores the un-cited work done by authors in their articles (Norris & Oppenheim, 2010).

*2.3. A, G,H-index and bibliometric indicators*

A-index is given more superiority and value over the h-index. In the study about the research performance of different universities in a city, it was found that there was an association between g-index, h-index and a-index but still a-index was far capable enough to locate out the difference of the performance of universities. Also, a-index and g-index can prevail over the shortcomings of h-index by differentiating where h-index does not and a-index served as a relatively better evaluative tool for distinguishing the performance of universities.
Association is strong between these indices is high as they are derived from the same concept (Jasco, 2011).
Whereas, g-index covers the weaknesses of h-index by highlighting highly cited papers and which have high impact (Jasco, 2011).

Google has introduced Google scholar, which is dedicated to the filtered view of research papers, journals and related research content only. Recently, it has also induced the bibliometric indicators also h-index for calculating the research output/ranking of a particular author. The results are not much reliable. This is due to the fact of the data stored and its handling. Scopus has a grand database and doesn't include any such issues. They are more eminent in research evaluative and ranking. Google is reliable but certain issues due arise due to the afore-mentioned negligence of accurate database of research (Jasco, 2011).

**3. Methodology**

*3.1 Description of data and sample*

The data of h- index and the ranks of top 1000 Authors, Journals and top 1000 Universities for last 5 years (2007-2011) were taken from the various sources which include, ISI Thomson Reuters, Scopus, IDEAS databases, while The data bases of Times Higher Education (THE), formerly Times Higher Education Supplement (THES) and QS World University Rankings were also used to get the data of outlined variables for Top 1000 Universities in particular.

*3.2 Descriptions of Econometrics applied*

The h-index of authors, journals, and universities in relevance to their impacts upon their rankings have been comparatively interrogated via deploying the simple linear regression.

**4. Findings and Results**

Table 1: H-Index Analysis

| Explained Variables / Predictors | Ranks of Authors/ Researchers | Ranks of Journals | Ranks of Universities |
|---|---|---|---|
| | Beta (T-Stats) | Beta (T-Stats) | Beta (T-Stats) |
| H-Index for Authors | 0.2145 (5.561) | | |
| H-Index for Journals | | 0.1907 (2.932) | |
| H-Index for Universities | | | 0.1038 |

| | | | (2.087) |
|---|---|---|---|
| Intercepts | 0.2965 (3.708) | 0.2704 (2.056) | 0.0838 (1.573) |
| | | | |
| R- Squared | 0.5341 | 0.3405 | 0.1788 |
| F- Statistics | 11.834 | 8.357 | 4.324 |

The findings of this paper as explained in Table 1, reveal that h-index of authors, journals and universities they all contribute for the research rakings of authors, journals and universities. While, h- index of authors matters more for their rankings as beta coefficient was found significant and relatively strong (i.e. beta= 0.2145, at t-stats= 5.561> 1.5) in addition to it the model which is comprised upon the intercept and the predictor (h-index of authors) for gauging the variations in explained variable (Authors Ranks) is also found significant as F= 11.834 > 3.84. Whereas, h-index for journals and university also significantly affects their ranking as their betas are significant at t-stats > 1.5. Interestingly, h-index of universities matters significantly but very meagerly for the research ranking of universities (as beta= 0.1038 at t > 1.5) as highlighted in the Table 1.

## 4. Conclusion

One can say that h-index is by far a good tool for assessing the ability of researchers, journals and universities' research output in terms of citations in the publications and scientific contributions (Mingers, 2009). But, the h-index works more for rankings of researchers then the research journals and universities. Of course, it is not an ideal but an acceptable measure so far because it doesn't include the calculation of co-authors. Another major factor in contribution to research output other than citations of an author, research journals and universities is the total number of publications have been produced by them, which also globally makes an author/ journals/ Universities ranks higher lacks in the h-index.

## References

Abbas, A. M. (2010).Weighted Indices for Evaluating the Quality of Research with Multiple Authorship. *Scientometrics*, 88(1): 107-131.

Bletsas, A., & Sahalos, J. N. (2009). Hirsch Index Rankings Require Scaling and Higher Moment. *Journal of the American Society for Information Science and Technology*, 60(12): 2577–2586.

Bornmann, L., Mutz, R.,& Daniel, H. D. (2008). Are there better indices for evaluation purposes than the h index?A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, 59(5), 830–837.

Bouyssou, D., & Marchant., T. (2011). Bibliometric rankings of journals based on impact factors: An axiomatic approach. *Journal of Informetrics*, 5(1):75-86.

Cronin, B. & Meho, L. I. (2006). Using the H-index to Rank Influential Information Scientists. *Journal of the American Society for Information Science and Technology*, 57(9): 1275-1278.

Guerin, M. T. (2010). Ranking the Scholarly Impact of New Jersey Colleges and Universities in the Field Of Psychology. *TCNJ Journal of Student Scholarship*, 12: 1-11.

Hirsch, J. E. (2007). Does the h index have predictive power? *PNAS*, 104(49): 19193–19198.

Jacso, P. (2011). The h-index, h-core citation rate and the bibliometric profile of the Scopus database. *Online Information Review*, 35(3): 492-501.

Jacso, P. (2008). The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced databases. *Online Information Review*, 32(2): 266-283.

Mingers, J. (2009). Measuring the research contribution of management academics using the Hirsch-index. *Journal of the Operational Research Society*, 60(9):1143 –1153.

Norris, M. & Oppenheim, C. (2010). The h-index: a broad review of a new bibliometric indicator. *Journal of Documentation*, 66(5): 681-705.

Subhani, M. I., Hasan, S. A., & Osman, A. (2011). Research methodologies for management sciences & interdisciplinary research in contemporary world. *European Journal of Scientific Research*, 63(4), 543–547.