Annals of Library and Information Studies
Vol. 71, June 2024, pp. 221-225
DOI: 10.56042/alis.v71i2.8854

Consideration of Neglected Citations by Hirsch Index

Mohammed Shaalan Abed Fathi
University of Mosul, 41002, Iraq
Email: momeceng@gmail.com

Received: 26 February 2024; Accepted: 12 March 2024

In this research article a modified h-index (h_{mo}) was proposed. This index adopts the calculation method of the original h-index, but takes into consideration all citations of the papers that have not been included in the calculation of the Hirsch Index; specifically neglected citations in the h-core. Thought that the productivity of a scientist is still effective in the modified h-index, the weight of the neglected citations–which reflects the quality of that scientist is added according to the proposed index. The current index slowly increases with the addition of any new citations in the h-core or the h-tail.

Keywords: h-index; h_{mo}-index; h-index variants; neglected citations in the h-core.

Symbol Notification

h Hirsch Index
h_{mo} Modified h-Index.
N_p total number of papers.
$N_{cit:}$ number of citations for the i^{th} paper
NC total neglected citations
TC total citations of all papers

Introduction

h-index is a preferential metric proposed by Professor Jorge E. Hirsch in the PNAS 2005's research article: "An Index to Quantify an Individual's Scientific Research Output". The aim is to measure the scientific output of a researcher [1]References. It is based on the number of peer-reviewed publications and the number of citations received for each of which.

Simply, if a researcher has an h-index of (10), it means that he/she has published at least ten papers each of which has been cited ten times at least. A researcher can obtain an h-index of (1) if he/she just publishes one paper and this paper should be cited one time.

The main advantages of the h-index are [1-4]:

- Easy to compute manually. Mostly it is calculated automatically on many citations databases.
- Its robustness in the sense that a new publication does not contribute in increasing the h-index before it has at least (h+1) citations.
- It is beneficial in making a comparison between two researches who have similar career length or similar field of specification.
- It combines the effect of quantity (number of publications) and quality (citation rate).

On the other side, the main drawbacks of the h-index are [4-7]:

- The possibility of being inflated through self–citations.
- Ignoring the span over which the papers had been published,
- Its insensitivity to co-authors.
- It does not decrease with time.
- Ignoring of the actual citations in the h-core, especially the highly-cited papers.
- It is insensitive to the publication type.
- h-index is a natural number; it is unable to discriminate scientists with equal h-indices.
- It disadvantages junior researcher.

Literature Review

Plenty of indices were proposed as a modification of the original h-index to consider one or more of the aforementioned drawbacks. Amongst them in chronological order with the main drawback that it was considered and the method of calculation between two brackets:

v-index (the publication age) [8].

g-index (highly–cited papers, "the highest number *g* of papers that together received g^2 or more citations) [9].

$h^{(2)}$ index (highly-cited papers) [10].

A-index (actual citations in the h-core, the average of the citations in the h-core) [11].***h_I-index*** (the

number of co-authors, the index averages the authors in the h-core) [12].

R-index (actual citations in the h-core, the square root of the citations in the h-core) [13].

AR-index (the publication age) [13].

h_p-index (actual number of co-authors) [14].

m-index (actual citations in the h-core, the median of the citations in the h-core [15].

e-index (excess citations in the h-core, the square root of these citations) [16].

h_m-index (multiple co-authorship) [17]

$\overline{h}$-index (multiple co-authorship) [18].

j-index (excess citations counts) [19].

C-index (authors contribution) [20].

q-index (self–citations) [21].

Mo-index (multi-authors papers) (Mohammed Shaalan,[22].

Despite of the large numbers of papers in the literature that were proposed to complement or replace the h-index, it is still the most popular index.

The h-index has been extended from measuring the performance of researchers to evaluate the scientific impact of journals (Braun et al [23], Schubert and Glanzel [24], Bornmann et al [25], Schubert [26]), academic institutions and universities (Meyers and Quan [27], J. Molinari and A. Molinari [28], Lazaridis [29])), References and even countries (Jacsó [30], Csajbók et al[31]).

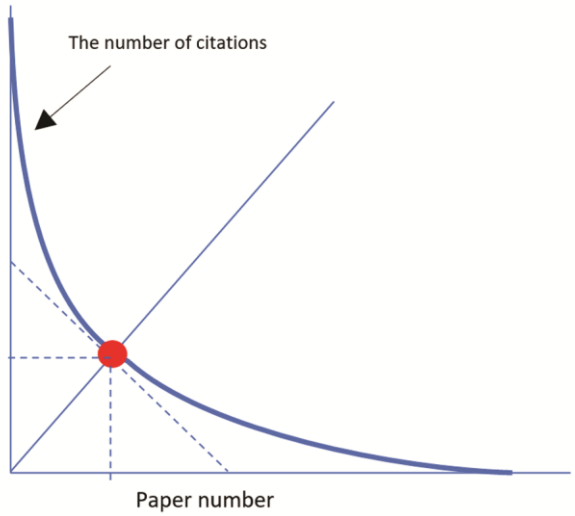

Fig. 1 — Graphically illustration of the h-index [1].

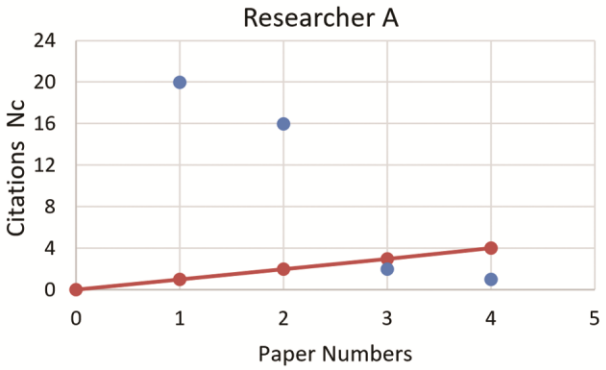

Fig. 2 — Graphical representation for the fictitious data of Researcher A.

Table 1 — Researcher A's fictitious data.

Paper	Citations
1	20
2	16
3	2
4	1

How the h-Index is calculated?

To calculate the h-index, all published papers of an author should be tabulated in a descending order of the number of citations they have received. The first column represents the paper number (**Np**) while the second column refers to the number of corresponding citations (**Nc**) of the papers. Then, by Hirsch's definition: the h-index equals the number of papers, say h, that each of them has been cited h times at least [1]. Graphically speaking (Figure 1), the h-index is the intersection point between a 45-degree line and the curve of citations vs. paper numbers[1].

Example 1:

For Researcher A with four published papers and thirty-nine of citations (Fig. 2), h-index is (2). Since the third paper has a number of citations less than three (Table 1).

Researcher B's h-index is (2), since the third paper has a number of citations less than 3. Researcher C who has four papers which have been cited sixteen times has an h-index of (4), since the fourth paper has a number of citations equal to four standing just on the threshold of the required citations to be included in the h-core (Fig. 3).

According to its database: each of Web of Science, Scopus, and Google Scholar provides updated values of an author's h-index whose papers were published in journals that are indexed in the above databases. The difference in the content indexed by these databases should interpret the discrepancies in the h-indices of the same author provided by these institutions (Table 2).

Table 2 — Hirsch's h-index and total citations on Scopus and Google Scholar.

Database	h-index	Total citation counts
Scopus	58	20225
Google Scholar	67	31393

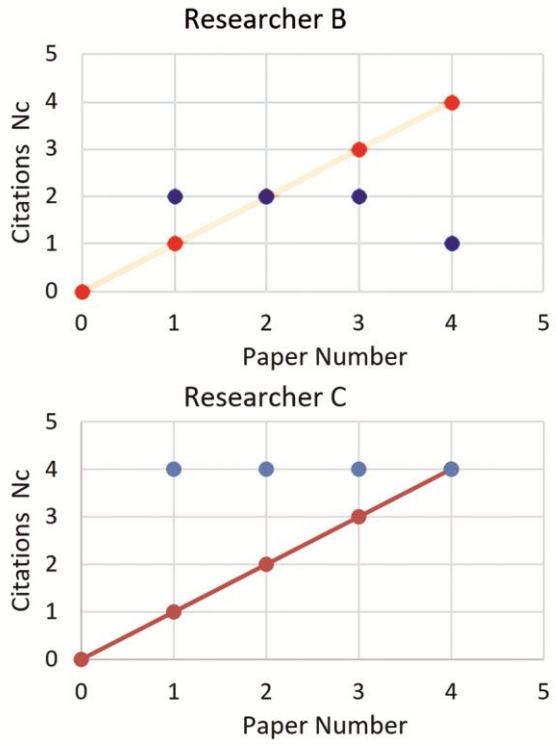

Fig. 3 — Graphical representation for the fictitious data of Researchers B and C.

Some tools such as Publish or Perish Software by Anne-Wil Harzing, Scholar h-index Calculator (Google Scholar) can be helpful in obtaining the h-index and some other bibliometric indicators (Fig. 4). As stated earlier self-citations affect the value of the h-index. Both Scopus and Web of Science databases offer the possibility of removing the self-cited papers of an author, which is not the case in Google Scholar.

The Current Modification

The current paper suggests a modified h-index (h_{mo}-index) to distinguish between researches who have the same h-index but some of them have neglected citations ***for being more than the required threshold in the calculation of h-index***. The modified h-index has two decimal points that gives weight for the neglected citations (NC) as a percentage of the total citations (TC). The neglected citations in the calculation of the h-index can be expressed as:

$$\mathrm{NC} = \sum_{i=1}^{i=h}\left((N_{cit})_i - h\right) + \sum_{i=(h+1)}^{i=N_p}(N_{cit})_i \qquad \ldots (1)$$

The first term in Eq. 1 represents **the excess citations in the h-core**, while the second term is **the h-tail citations** (if any). The total number of citations (TC) is:

$$\text{Total Citatioons (TC)} = \sum_{i=1}^{i=N_p}(N_{cit})_i \qquad \ldots (2)$$

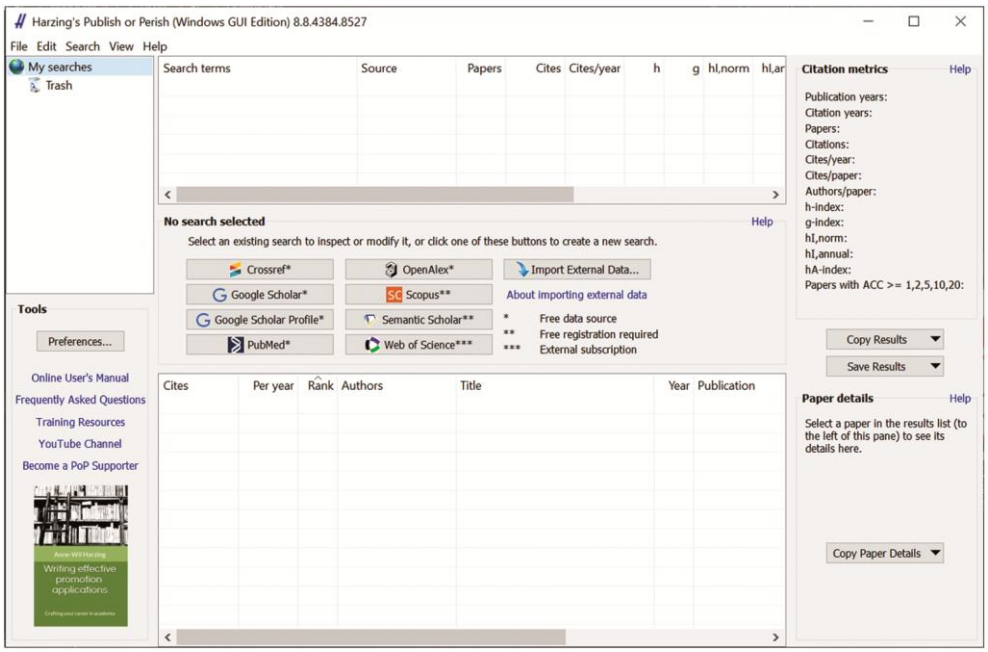

Fig. 4 — The interface of Publish or Perish software.

Since the useful number of citations in the h-core to calculate h-index is h^2, Eq. (1) for the NC can be re-written as:

$$NC = TC - h^2 = \sum_{i=1}^{i=N_p}(N_{cit})_i - h^2 \quad \ldots (3)$$

By the definition of the h_{mo}-index above, it can be mathematically formed as:

$$h_{mo} = h + \frac{NC}{TC} \quad \ldots (4)$$

Substituting Eq. 3 in Eq. 4 yields:

$$h_{mo} = h + \frac{TC - h^2}{TC} \quad \ldots (5)$$

$$h_{mo} = h + \left(1 - \frac{h^2}{TC}\right) \quad \ldots (6)$$

It is clear that $h_{mo} \geq$ h-index and the part of equation [6] that

Example 2:

In the preceding example, Researcher A has published four papers and has *thirty-six citations* for his two papers which have been used to calculate his h-index. At the same time, Researcher B has published two papers with only *four citations* for his two papers. Yet, Researcher B has the same h-index as Research A. If we calculate (h_{mo}) for both researchers using Eq. 6 :

Researcher A:
1. h-index =2
2. h^2=4
3. Total number of citations (TC) = 39
4. Applying Eq. 6 to obtain h_{mo}-index.

$$h_{mo} = h + \left(1 - \frac{h^2}{TC}\right)$$
$$= 2 + \left(1 - \frac{4}{39}\right) = 2.90$$

Researcher B:
1. h-index =2
2. h^2=4
3. Total number of citations (TC) = 7
4. Applying Eq. 6 to obtain h_{mo}-index.

$$h_{mo} = h + \left(1 - \frac{h^2}{TC}\right)$$
$$= 2 + \left(1 - \frac{4}{7}\right) = 2.43$$

Table 3 — h-index and total citations of hypothetical researchers D, E, and F.

Researcher	N_p	N_{cit}	Current h-index
Researcher D	1	1	1
Researcher E	1	5	1
Researcher F	1	20	1

Table 4 — h_{mo}-index and total citations of hypothetical researchers D, E, and F.

Researcher	N_p	N_{cit}	h_{mo}
Researcher D	1	1	h_{mo} = h = 1.00
Researcher E	1	5	h_{mo} = 1.80
Researcher F	1	20	h_{mo} = 1.95

Using h_{mo}, it is clear that Researcher A has been given a better stand than Researcher B for having a better-quality paper (based on their citations) than Research B, for the same number of the published papers (quantity).

Example 3:

Assume three researchers D, E, and F with the fictitious data in Table 3.

All of the researchers have the same h-index of (1). But, Researcher E and Researcher F have more four and nineteen citations respectively (Table 4) than Researcher D. If h_{mo} is calculated for all of the researchers using Eq. 6:

Before publishing a second paper with at least two citations, Researcher F will never have an h_{mo} equal to or more than (2). As soon as he/she does, he/she will have an h_{mo} of (2.82) rather than an h-index of (2). If Researcher E published a second paper and it would be cited at least two times, he/she will have an h_{mo} of (2.43) rather than an h-index of (2). Obviously, the large number of citations for Researcher F's first paper is still effective to give him/her a lead to Researcher E.

Conclusions

The scientific development is powered by the scientific research. The more the articles the continuous the development. On the other hand, the number of citations any scientific article receives should be an indication of its impact in the research community. To some extent, Hirsch Index correlates between the quantity and the quality of a scientist or a scholar. Yet, it ignores all citations beyond his/her current deserving h-index. On that, a modification takes care of those neglected citations was suggested in this article.

The h_{mo}-index still depends on the original h-index and stands on the limit of the quantity before

reflecting the effect of the quality each time the quantity increases. The outcome is that h_{mo} converts all ignored citations into a two–decimal–digit addition to the original h-index.

Some of the advantages of the h_{mo}-index are:

- The easiness with which it can be computed comparing with many other indices. All it takes to calculate this index is an author's *total number of citations* and *his/her h-index*. Such resources can be obtained from any public citation database.
- It considers excess citations that are ignored by the h-index.
- It slowly increases with the addition of any new citations in the h-core or the h-tail.

References

1 J. E. Hirsch , An index to quantify an individual's scientific research output, Proceedings in the National Academy of Sciences of The United States of America, 102,(46)((2005), \16569-16572, doi: 10.1073/pnas.0507655102.

2 University of Waterloo's Library, at https://subjectguides. uwaterloo.ca.

3 Rodrigo Costas, Mar´ıa Bordons, The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level", *Journal of Informetrics*, 1(3) (2007) 193–203, doi: 10.1016/j.joi.2007.02.001.

4 S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera , h-Index: A review focused in its variants, computation and standardization for different scientific fields, *Journal of Informetrics*, 3(4) (2009) 273–289, doi: 10.1016/j.joi.2009.04.001.

5 L. A. Zhivotovsky and K. V. Krutovsky (2008), "Self-citation can inflate h-index," Scientometrics, 77 (2) (2008) 373–375, doi: 10.1007/s11192-006-1716-2.

6 W. Dinkel, The h-index: Definition, calculation, limitations and benefits, Part 1, ESSS, 10.1016/j.joi.2007.02.001.

7 R. S. J. Tol (2008) , A rational, successive g-index applied to economics departments in Ireland, *Journal of Informetrics*, 2 (2) 149–155, doi: 10.1016/j.joi.2008.01.001.

8 Vaidya, J. S. V-index: A fairer index to quantify an individual's research output capacity, *BMJ*. (2005)

9 Egghe L,\ Theory and practice of the g-index ,\ *Scientometrics*, 69(1) (2006)\ 131–152.

10 M. Kosmulski, A new Hirsch-type index saves time and works equally well as the original h-index, *International Society for Scientometrics and Informetrics* (ISSI),(2006) 4-6.

11 Jin, B. h-Index: An evaluation indicator proposed by scientist, *Science Focus*, 1(1)(2006)).8–9.

12 P. D. Batista, M. G. Campiteli, and O. Kinouchi, "Is it possible to compare researchers with different scientific interests, *Scientometrics*, 68 (1) (2006), 179–189, doi: 10.1007/s11192-006-0090-4 .

13 B. Jin, L. Liang, R. Rousseau, and L. Egghe, The R- and AR-indices: Complementing the h-index," *Chinese Society Bulletin*, 52(6)(2007) 855–863, doi: 10.1007/s11434-007-0145-9.

14 J. Wan, P. Hua, and R. Rousseau, The pure h-index: calculating an author's h - index by taking co-authors into account, *Collnet Journal of Scientometrics and Information Management*, 1(2) (2007) 1–5, doi:10.1080/09737766. 2007.10700824.

15 L. Bornmann, R. Mutz, and H.-D. Daniel , 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) (2008) 830–837, doi: 10.1002/asi.20806.

16 C.-T. Zhang , The e-Index, Complementing the h-Index for Excess Citations," *PLOS One*, 4(5) (2009) doi: 10.1371/journal.pone.0005429.

17 M. Schreiber, A case study of the modified Hirsch index hm accounting for multiple coauthors, *Journal of the American Society for Information Science and Technology*, 60, (6) (2009) 1274–1282, doi: 10.1002/asi.21057.

18 J. E. Hirsch , An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship, *Scientometrics*, 85, (3) (2010) 741–754 , doi: 10.1007/s11192-010-0193-9.

19 R. Todeschini, The j-index: a new bibliometric index and multivariate comparisons between other common indices, *Scientometrics*, 87, (3) (2011) 621–639, doi: 10.1007/s11192-011-0346-5.

20 A. Post et al, c-index and Subindices of the h-index: New Variants of the h-index to Account for Variations in Author Contribution, *Cureus*, 10, (5) (2006), doi: 10.7759/cureus.2629.

21 C. Bartneck and S. Kokkelmans, Detecting h-index manipulation through self-citation analysis, *Scientometrics*, 87(1) (2011) 85–98, doi: 10.1007/s11192-010-0306-5.

22 Fathi, M. S. A., Mo-Index for multi-authors papers, *Annals of Library and Information Studies* (ALIS), 69(4) (2022) 323-326.

23 T. Braun, W. Glänzel, and A. Schubert, A Hirsch-type index for journals,\ *Scientometrics*, 69, (1) (2006) 169–173, doi: 10.1007/s11192-006-0147-4.

24 A. Schubert and W. Glanzel , A systematic analysis of Hirsch-type indices for journals," *Journal of Informetrics*, 1(3) (2007) 179–184, doi: 10.1016/j.joi.2006.12.002.

25 L. Bornmann, W. Marx, and H. Schier , Hirsch-Type Index Values for Organic Chemistry Journals: A Comparison of New Metrics with the Journal Impact Factor, *European Journal of Organic Chemistry*, 2009 (10) (2009) 1471–1476, doi: 10.1002/ejoc.200801243.

26 András Schubert, Successive h-indices, *Scientometrics*, 70(1) (2007) 201-205.

27 M. A. Meyers and H. Quan , The use of the h-index to evaluate and rank academic departments, *Journal of Materials Research and Technology*, 6(4) (2017) 304–311, doi: 10.1016/j.jmrt.2017.09.004.

28 J.-F. Molinari and A. Molinari , A new methodology for ranking scientific institutions, *Scientometrics*, 75 (1) (2008) 163–174, doi: 10.1007/s11192-007-1853-2.

29 T. Lazaridis , Ranking university departments using the mean h-index," *Scientometrics*, 82 (2) (2010) 211–216, doi: 10.1007/s11192-009-0048-4.

30 P. Jacsó , The h-index for countries in Web of Science and Scopus, *Online Information Review*, 33 (4) (2009) 831–837, doi: 10.1108/14684520910985756.

31 E. Csajbók, A. Berhidi, L. Vasas, and A. Schubert , Hirsch-index for countries based on Essential Science Indicators data, *Scientometrics*, 73 (1) (2007) 91–117, doi: 10.1007/s11192-007-1859-9