# The average citation ha-index

Yves Fassin, Ghent University, Belgium

Department of Marketing, Innovation and Organisation

Yves.Fassin@UGent.be or fassin@skynet.be

Reference :

**Fassin, Y. 2020. The average citation ha-index. International Society for Scientometrics and Informetrics. ISSI Newsletter 16 (4): 64-67**

https://www.issi-society.org/media/1420/newsletter64.pdf

The h-index proposed by Hirsch (2005) has been widely adopted as an indicator for assessment of scientific achievement (Ball 2005; Zhang, Thijs and Glänzel, 2011). This success is due to its easiness of calculation included in its definition, namely: *the highest number of papers, denoted as h, a scientist has that have each received at least that number of citations* (Hirsch 2005).

The application of the h-index has been extended to other datasets: from the individual researchers to teams, departments, universities, fields and also to journals (Braun, Glänzel and Schubert, 2006). Many bibliometrics scholars have pointed to imperfections and drawbacks of the h-index (Costas & Bordons, 2007; Bouyssou and Marchant, 2011; Waltman and Van Eck, 2012; Rousseau, Egghe and Guns, 2018). As a consequence, a considerable number of h-type variants have been proposed (Bormann, Mutz, Hug and Daniel, 2011), such as the g-index (Egghe, 2006) and the h²-index (Kosmulski, 2006).

The growth in the number of publications and journals since the launch of the h-index in 2005, has led to an inflation of the h-index values of all researchers and all journals. This phenomenon has diminished the significance and differentiation power of the h-index. Other criticisms refer to the need for normalization to allow fairer comparison (Harzing, Alakangas and Adams, 2014). Several attempts for normalization have been proposed, especially towards adjustments for age or career length.

Besides the h-index, Hirsch (2005) also defined the m-quotient as h/n, where n equals the number of years since the researcher's first publication. Whereas the h-index executes the ranking according to the article citations, without any normalization, the m-quotient corrects the h-index for age after ranking. In fact, the calculation includes two successive operations: the ranking and the correction for age, which is the division of the h-index by the number of years of research activity. Curiously, the alternative with the two operations performed in different order has not been proposed yet. The correction for age could precede the ranking. This could lead to an h-index variant. The principle applied to define the h-index - the ranking

of a dataset in declining order of total citations - can then be applied to the ranking of that dataset in declining order of the publications' average citation per year. An average citation h-index can be defined: the $h_a$-index of a given dataset is *the largest number of papers in this dataset that have obtained at least $h_a$ citations per year on average*

Despite numerous h-type variants, this simple alternative has not been proposed yet. It is probably because the h-index has always been presented as an integer, while the average citations per paper are not integer numbers. An average applied at the level of the individual publication confers more precision and avoid possible distorting effects of the overall average.

Table 1 presents the data and h- and $h_a$-indexes of ten scholars in entrepreneurship research with different profile and years of activity, with the ranking according to their h-, m- and $h_a$-indexes.

Table 1: Comparative data and h- and $h_a$-indexes of ten scholars in entrepreneurship research

| Author | n | TC | y | h | m | ha | *rank h* | *rank m* | *rank ha* |
|---|---|---|---|---|---|---|---|---|---|
| A1 | 125 | 24147 | 29 | 66 | 2.28 | 19 | *1* | *1* | *1* |
| A2 | 225 | 12438 | 29 | 62 | 2.14 | 19 | *2* | *3* | *1* |
| A3 | 82 | 15626 | 41 | 46 | 1.12 | 14 | *3* | *7* | *3* |
| A4 | 74 | 2812 | 13 | 29 | 2.23 | 13 | *5* | *2* | *4* |
| A5 | 33 | 2497 | 14 | 22 | 1.57 | 11 | *6* | *5* | *5* |
| A6 | 85 | 7915 | 45 | 38 | 0.84 | 10 | *4* | *10* | *6* |
| A7 | 38 | 867 | 8 | 17 | 2.13 | 6 | *7* | *4* | *7* |
| A8 | 25 | 347 | 13 | 12 | 0.92 | 5 | *8* | *9* | *8* |
| A9 | 19 | 406 | 8 | 11 | 1.38 | 5 | *9* | *6* | *8* |
| A10 | 8 | 96 | 6 | 6 | 1.00 | 3 | *10* | *8* | *10* |

*Legend: Author, n number of papers in WoS, TC total citations in WoS, y the number of years the scholar has been active (since their first publications), h-index, m-quotient (h/y), the new ha, rank corresponding ranking of the basis of h, m and ha-indexes.*

The analysis of the comparative table presents some indications with a few changes of the order, especially for the middle category, While the m-quotient heavily penalizes some older researchers, the $h_a$-index improves the position of mid-career and younger researchers compared to their h-index ranking. The $h_a$-index ranking mitigates the h-index ranking. The selection of the articles in the $h_a$-core of a dataset is different from the selection in the h-core or in the $h^2$-core. The $h_a$-core can include younger articles that are not in the $h^2$-core and vice versa. This new indicator acknowledges potential, yet also recognizes experience.

The average citation $h_a$-index has other advantages compared to the classic h-index. It renders a better selectivity and more stability. The $h_a$-index is lower in value than the high levels of the classic h-index. The number of articles in the $h_a$-core lies in the same range as the $h^2$-index, maybe somewhat higher. Only those publications that sustain a high growth rate over a longer period of time will maintain their presence in the $h_a$-core. In this way, the $h_a$-index is a truly dynamic index. The increase of citations by one unit per year of existence is indeed a severe criterion. The stricter measure also prevents manipulation of the h-index through self-citations.

The greater stability of the ha-index makes the application and comparison of ha-indexes even more useful for academic journals, or other datasets such as scientific fields with much larger numbers of articles and a smoother citation distribution curve. Table 2 displays the data and h-type indexes for a few journals in library and information sciences and in research policy. For academic journals also, the ha-index ranking mitigates the h-index ranking.

Table 2: data and h-type indexes of library and information sciences journals

| | n | avg cit | h | h² | ha |
|---|---|---|---|---|---|
| JASIS - JASIST | 6797 | 21.9 | 144 | 21 | 27 |
| Scientometrics | 6318 | 18.1 | 117 | 18 | 21 |
| Journal of Informetrics | 1038 | 22.0 | 66 | 14 | 18 |
| Information Processing and Management | 3694 | 15.6 | 93 | 15 | 16 |
| Journal of Documentation | 4141 | 8.4 | 75 | 16 | 13 |
| Journal of Information Science | 2349 | 11.2 | 32 | 12 | 11 |
| Journal of Data and Information Science | 91 | 3.6 | 8 | 3 | 3 |
| Malaysian Journal of Library and Informatior | 275 | 4.4 | 15 | 5 | 3 |
| Journal of Scientometric Research | 199 | 0.6 | 4 | 2 | 2 |
| Research Policy | 3702 | 68.5 | 231 | 27 | 42 |
| Technological Forecasting and Social Change | 5687 | 16.3 | 106 | 16 | 26 |
| Research Evaluation | 642 | 14.8 | 45 | 10 | 10 |
| ° retrieved from WoS on 12th November 2020 | | | | | |

The evolution of the h- and ha-index over the years presents evidence of a greater stability. Whereas the h-index tends to progress linearly, the ha-index manifests smaller increases in a parabolic form to reach a plateau in the phase of maturity. The ha-index of *Scientometrics* advanced from 15 in 2010 towards 21 by the end of 2020, while its h-index nearly doubled from 60 to 117. Its h-index continues to rise by more than 5% a year, or 5 to 9 units a year. The rise of the ha-index is somewhat slighter, but in absolute values it increases only by one every one and a half year.

The ha-index offers more stability over time and provides a response to the inflation of levels of h-indexes. It has the same ease of calculation as the h-index and can easily be incorporated in databases.

**References**


Ball, P. (2005). Index aims for fair ranking of scientists. *Nature*, 436, 900.

Bouyssou, D., & Marchant, T. (2011). Ranking scientists and departments in a consistent manner. *Journal of the American Society for Information Science and Technology*, 62(9), 1761-1769.

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the h-index and 37 different h-index variants. *Journal of Informetrics*, 5(3), 346-359.

Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, 69(1), 169-173.

Costas, R. & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1, 193–203.

Egghe, L. (2006). An improvement of the h-index: The g-index. *ISSI newsletter*, 2(1), 8-9.

Harzing, A., Alakangas, S., & Adams, D. (2014) hIa: An individual annual h-index to accommodate disciplinary and career length differences, *Scientometrics*, 99 (3), 811-821.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences USA* 102, 16569-16572.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI Newsletter*, 2(3), 4-6.

Rousseau, R., Egghe, L., & Guns, R. (2018). *Becoming metric-wise: A bibliometric guide for researchers*. Chandos Publishing.

Waltman, L. and Van Eck, NJ. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*. 63(2), 406–415.

Zhang, L., Thijs, B., & Glänzel, W. (2011). The diffusion of H-related literature. *Journal of Informetrics*, 5(4), 583-593.


*Annex*

Table 3 exhibits the analysis of the evolution of the h- and $h_a$-index over the years for the journal *Scientometrics*.

Table 3: The evolution of various h-type indexes of *Scientometrics* over the years

| | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020° | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 297 | 626 | 1032 | 1530 | 2041 | 2886 | 4369 | 6318 | 4369 | 4748 | 5144 | 5542 | 5874 | 6318 |
| tot cit | 620 | 2000 | 3500 | 5500 | 8500 | 25000 | 55000 | 120000 | | | | | | |
| h | 10 | 17 | 23 | 28 | 36 | 60 | 87 | 117 | 87 | 92 | 96 | 102 | 108 | 117 |
| $h^2$ | 3 | 5 | 6 | 7 | 8 | 10 | 14 | 18 | 14 | 15 | 16 | 17 | 17 | 18 |
| ha | 4 | 4 | 4 | 4 | 6 | 15 | 17 | 21 | 17 | 18 | 19 | 19 | 20 | 21 |
| ° retrieved 11th November 2020 | | | | | | | | | | | | | | |