BRIEF COMMUNICATION

A Proposal for a Dynamic h-Type Index

Ronald Rousseau
KHBO (Association K.U. Leuven), Industrial Sciences and Technology, Zeedijk 101, B-8400 Oostende, Belgium and K.U. Leuven, Steunpunt O&O Indicatoren, Dekenstraat 2, B-3000 Leuven, Belgium. E-mail: ronald.rousseau@khbo.be

Fred Y. Ye
Zhejiang University, Department of Information Resources Management, Hangzhou, China and Institute of Scientific and Technical Information of China, Beijing, China. E-mail: yye@zju.edu.cn

A time-dependent h-type indicator is proposed. This indicator depends on the size of the h-core, the number of citations received, and recent change in the value of the h-index. As such, it tries to combine in a dynamic way older information about the source (e.g., a scientist or research institute that is evaluated) with recent information.

Introduction

Since the introduction of the h-index in 2005 (Hirsch, 2005), many h-type indices have been proposed. In this short note, we propose yet another h-type index, which differs from other proposals in the way that it takes recent change in the value of the h-index into account. As such, we are convinced that this proposal is more useful for hiring purposes than the lifetime-achievement h-index.

For the reader's convenience, we recall the definition of the h-index. Consider a scientist's list of publications, ranked according to the number of citations received. Then this scientist's h-index is defined as the highest rank such that the first h publications received each at least h citations. It is well-known that this definition can be applied to many more source-item relations than a scientist's publications and citations. For simplicity, however, we use the original publication-citation terminology, but note that our proposal is applicable to most of these other relations. These first h articles form the h-core. The number of articles in the h-core and the list of citations received by articles in the h-core are always well-defined. As the h-index does not take the actual number of citations into account (i.e., it is just a cutoff point), several proposals have been published to make the h-index more sensitive to the real number of citations. Among these, we prefer the R-index proposed by Jin, Liang, Rousseau, and Egghe (2007), which is equal to the square root of the total number of citations received by articles belonging to the h-core (for more information about the h-index and its generalizations, refer to Bornmann & Daniel, 2007).

Received March 30, 2008; revised April 17, 2008; accepted April 23, 2008

 • Published online 11 June 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20890

Rates of Change

Burrell (2007) considered two h-rates:

- The raw h-rate at Time T is defined as $h(T)/T$, where $T = 0$ at the beginning of a scientist's career or at another appropriate time, depending on the purpose of the study.
- The least squares h-rate at Time T is defined as the slope of the least squares regression line constrained to pass through the origin for all pairs $[t, h(t)]$, $t = 0, \ldots, T$.

If the h-index increases linearly with career age, as suggested by Hirsch (2005), then these two h-rates are equal.

For further use, we define the h-increment at Time T ($T > 0$) as

$$\Delta h(T) = h(T) - h(T-1) \tag{1}$$

Assuming a continuous model $h(t)$ for the h-index, we further define the h-velocity at Time T as:

$$v_h(T) = \frac{dh}{dt}(T) = \lim_{t \to 0} \frac{h(T+t) - h(T)}{t} \tag{2}$$

If $h(t)$ is a linearly increasing function of t, say $h(t) = m \cdot t$ ($m > 0$), then Burrell's h-rate is equal to m, the increment also is m, and the h-velocity also is m. However, if $h(t)$ is another function, then Burrell's h-rate is different from the h-velocity. When using, for example, the function derived in Egghe (2007), $h(t) = ((C(t))^{\alpha-1}\ P)^{1/\alpha}$, where $C(t)$ denotes the continuous citation distribution function, P is the number of publication at $t =$ infinity, and $\alpha > 1$ is the Lotka exponent for the citation function,

$$h'(t) = \frac{\left((C(t))^{\alpha-1} \cdot P\right)^{(1-\alpha)/\alpha}}{\alpha} \times \left(P \cdot (\alpha - 1) \cdot C(t)^{\alpha-2} \cdot C'(t)\right),$$

one finds a derivative which is totally different from the corresponding Burrell h-rate.

The Rational h-Index (Ruane & Tol, 2008)

This index, denoted as h_{rat}, is defined as $(h + 1)$ minus the relative number of scores necessary for obtaining a value $h + 1$. It clearly satisfies the inequality $h \leq h_{rat} < h + 1$. More precisely, let n be the (least) number of citations necessary for obtaining an h-index 1 higher than h. This number n is divided by the highest possible n—namely, $2h + 1$. Indeed, the lowest possible situation leading to an h-index equal to h consists of h articles with h citations, followed by an article without any citation. To get an h-index equal to $h + 1$, one needs one more score for each of the first h sources, h scores in total, and $h + 1$ scores for the last one: a total of $2h + 1$. This h-index has the advantage of increasing in smaller steps than the standard h-index. It will play a role in our proposal for a dynamic h-type index.

A Dynamic h-Type Index

We propose an index which depends on the h-core, the actual number of citations received by articles belonging to the h-core, and the recent increase in h. The definition contains three time-dependent elements: the size and contents of the h-core, the number of citations received, and the h-velocity. It is indeed possible that two scientists have the same h-index and the same number of citations in the h-core, but that one has no change in his h-index for a long time while the other scientist's h-index is on the rise. For hiring purposes, the second scientist is probably the better choice.

Consequently, we propose

$$R(T) \cdot v_h(T) \tag{3}$$

as a dynamic h-type index. Here $R(T)$ denotes the R-index (Jin et al., 2007), equal to the square root of the sum of all citations received by articles belonging to the h-core at Time T. In practice, we have to determine a starting point, $T = 0$, and a way of determining v_h. In our opinion, this starting point should not be the beginning of a scientist's career, but when T is "now," then $T = 0$ can be taken 10 or 5 years ago (or any other appropriate time). If one has a good-fitting continuous model for $h(t)$ over this period, then this function should be used to determine $v_h(T)$. In practice, it is probably better to find a fitting for $h_{rat}(t)$—and not for $h(t)$—as this function is more similar to a continuous function than the standard h-index. Otherwise, the increment $\Delta h_{rat}(T) = h_{rat}(T) - h_{rat}(T - 1)$ can be used (if it is not an obvious outlier). Note that when $h_{rat}(t)$ is concave, this approximation will be larger than the real derivative; when $h_{rat}(t)$ is convex, it will be smaller. When using this approximation, it is certainly appropriate to use the rational h-index as otherwise $\Delta(h)$ will often be 0 or 1, and no meaning can be attached to these values. Note that Burrell's raw h-rate $h(T)/T$ should not be used as it is equal for all scientists with the same $h(T)$, and hence, one loses the dynamic aspect. If Equation 3 is actually used for evaluating purposes, self-citations should be removed.

We claim that a period of "10 or 5 years," as mentioned earlier, is appropriate; however, determining this period is not a "scientific" choice. After decades of experience, a consensus has grown that when using journal impact factors for evaluation purposes (not of individual scientists, but for research groups), a period of 3 or 4 years is appropriate. Yet, even today the "official" Thomson Scientific impact factor uses 2 years of citations, and Thomson's Essential Science Indicators use a 10-year period. Hence, "10 or 5 years" is just a suggestion; only many years of experience can substantiate it.

Some Practical Considerations

If two scientists have the same h-index at Time T and the same number of citations, then the one whose h-function is convex (over the period [0,T]) is preferred to the one whose h-function shows a linear increase. Similarly, the one with a linearly increasing h-function is preferred to the one with a concave h-function. This is particularly true for a scientist whose h-function is given by Egghe's (2007) model, as this function is concavely increasing.

An Application

Using data from the Web of Science the first author's h-index, rational h-index and R-index from 2001 to 2008 (April 11) is determined (see Table 1). Author self-citations, but not article self-citations, were removed.

A best-fitting power function $h_{rat} = a.\ y^b$ was determined using nonlinear regression, leading to $h_{rat} = 1.67\ y^{0.801}$ ($R^2 = 0.984$). The R^2 value indicates that this is a good fit. Its derivative is $1.338\ y^{-0.199}$, and in the Year 7 (2008),

TABLE 1. Rousseau's h-index, rational h-index, and R-index over the period 2001–2008.

Year	2001	2002	2003	2004	2005	2006	2007	2008
h-index	0	1	2	3	5	5	6	7
h_{rat}	0.00	1.67	2.60	3.86	5.82	5.91	6.92	7.87
R-index	0.00	1.41	2.24	4.58	6.71	7.87	10.15	10.91

DOI: 10.1002/asi

the value of this derivative is 0.91. Hence, R.R.'s dynamic h-index in the 2008 is equal to $(10.91) \cdot (0.91) = 9.93$. Using the approximation $\Delta h_{rat}(T) = h_{rat}(T) - h_{rat}(T-1)$ for h_{rat} (7) yields $7.87 - 6.92 = 0.95$, which should be compared with 0.91.

Acknowledgment

We thank the National Natural Science Foundation of China (NSFC Grants 70773101 and 70673019) for supporting our collaboration. The first author thanks the National Institute for Innovation Management (NIIM), Zhejiang University, and in particular Profs. Jin Chen and Xiaobo Wu for promoting and financially supporting the collaboration leading to this article. Anonymous reviewers are acknowledged for some useful suggestions.

References

Bornmann, L., & Daniel, H.-D. (2007). What do we know about the h index? Journal of the American Society for Information Science and Technology, 58, 1381–1385.

Burrell, Q.L. (2007). Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. Scientometrics, 73(1), 19–28.

Egghe, L. (2007). Dynamic h-index: The Hirsch index in function of time. Journal of the American Society for Information Science and Technology, 58(3), 452–454.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, USA, 102(46), 16569–16572.

Jin, BH., Liang, LM., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin, 52, 855–863.

Ruane, F.P., & Tol, R.S.J. (2008). Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics, 75, 395–405.

DOI: 10.1002/asi