

A Taxonomy of Bibliometric Performance Indicators Based on the Property of Consistency

Ludo Waltman and Nees Jan van Eck

{lwaltman, nvaneck}@ese.eur.nl

Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam (The Netherlands) and
Centre for Science and Technology Studies, Leiden University (The Netherlands)

Introduction

Many different indicators of the performance of researchers, research groups, journals etc. have been proposed in the bibliometric literature. In order to better understand the relations between various performance indicators (PIs), we propose a taxonomy of PIs. We focus on PIs that depend on publication and citation data only, and we do not consider more advanced PIs that aim to correct for, for example, differences between fields, differences between publication types (e.g., article, note, or review), or the time lag between publication and citation.

We note that a somewhat more extended version of this paper is available as a technical report (Waltman & Van Eck, 2009).

Consistency

An important role in the taxonomy that we propose is played by the property of consistency. Suppose that, according to some PI, researcher A has a higher total performance than researcher B. Suppose next that both researchers obtain one additional publication, both with the same number of citations. It is then natural to expect that researcher A's total performance remains higher than that of researcher B. A PI that is guaranteed to have this property is said to be consistent. Interestingly, the well-known *h*-index or Hirsch index (Hirsch, 2005) is not consistent. To see this, suppose that researcher A has three publications with five citations each while researcher B has four publications with four citations each. Suppose next that both researchers obtain one additional publication with five citations. Researcher A's *h*-index then increases from three to four while researcher B's *h*-index remains equal to four. This violates the property of consistency. (Notice that if both researchers obtain another publication with five citations, researcher A even outperforms researcher B.) We note that consistency is equivalent to what is called independence by Marchant (in press-a, in press-b).

The use of an inconsistent PI can lead to odd results. Suppose, for example, that the *h*-index is used to compare the total performance of two research groups, research group A and research group B. Research group A consists of five researchers. Each researcher in research group A has five publications with five citations each.

Hence, in total research group A has 25 publications with five citations each. Research group B also consists of five researchers. Each researcher in research group B has two publications with ten citations each. Hence, in total research group B has ten publications with ten citations each. It is clear that each researcher in research group A has an *h*-index of five while each researcher in research group B has an *h*-index of two. This means that each researcher in research group A outperforms each researcher in research group B. Based on this result, it is natural to expect that research group A as a whole outperforms research group B as a whole. However, this is not the case. Research group A has an *h*-index of five, while research group B has an *h*-index of ten. Hence, research group A is outperformed by research group B rather than the other way around. This demonstrates that an inconsistent PI such as the *h*-index can lead to odd results. Such results are not possible if a consistent PI is used.

Taxonomy

Let $x_A(i)$ denote the number of publications of researcher (or research group or journal) A with at least i citations, and let $\mathbf{x}_A = (x_A(0), x_A(1), \dots)$. We define a PI as a function that maps every $\mathbf{x} = (x(0), x(1), \dots)$, where $x(i) \geq x(i+1)$ for all i , to a real number. The taxonomy of PIs that we propose is shown graphically in Figure 1.

The main distinction that we make is between indicators of the average performance per publication and indicators of the total performance of a set of publications. We define an average PI as a PI f such that for all $\mathbf{x}_A$ and $\mathbf{x}_B$

$$x_A(i) / x_A(0) \geq x_B(i) / x_B(0) \text{ for all } i \Rightarrow f(\mathbf{x}_A) \geq f(\mathbf{x}_B).$$

An example of an average PI is the average number of citations per publication (e.g., the journal impact factor). We define a total PI as a PI f such that for all $\mathbf{x}_A$ and $\mathbf{x}_B$

$$x_A(i) \geq x_B(i) \text{ for all } i \Rightarrow f(\mathbf{x}_A) \geq f(\mathbf{x}_B).$$

An example of a total PI is the total number of citations of a set of publications. Another example is the *h*-index. Interestingly, some PIs proposed in the literature, such as the A-index (Jin, Liang, Rousseau & Egghe, 2007), are neither average PIs nor total PIs.

Within the class of total PIs, we distinguish between consistent and inconsistent indicators. The

property of consistency was discussed above. Mathematically, a consistent total PI is defined as a total PI f such that for all $\mathbf{x}_A$, $\mathbf{x}_B$, and $\mathbf{x}_C$

$$f(\mathbf{x}_A) \geq f(\mathbf{x}_B) \Leftrightarrow f(\mathbf{x}_A + \mathbf{x}_C) \geq f(\mathbf{x}_B + \mathbf{x}_C).$$

We regard consistency as a very natural and highly desirable property. We therefore consider it a serious shortcoming that the h -index and its variants, such as the g -index (Egghe, 2006), the R -index (Jin et al., 2007), and the h_a - and g_a -indices (Van Eck & Waltman, 2008), do not have this property. Within the class of consistent total PIs, we make a further distinction between linear and non-linear indicators. We define a linear total PI as a total PI f such that for all $\mathbf{x}_A$ and $\mathbf{x}_B$

$$f(\mathbf{x}_A + \mathbf{x}_B) = f(\mathbf{x}_A) + f(\mathbf{x}_B).$$

The total number of citations of a set of publications is an example of a linear total PI. Linear total PIs can be interpreted as measures on a ratio scale, while non-linear total PIs can only be interpreted as measures on an ordinal (or sometimes interval) scale. Suppose, for example, that $f(\mathbf{x}_A) = 2f(\mathbf{x}_B)$. If f is a linear total PI, it makes sense to say that researcher A has performed twice as well as researcher B. If f is a non-linear total PI, this does not make sense and one can only say that researcher A has performed better than researcher B. It is easy to see that all linear total PIs are consistent. It further follows from the results of Marchant (in press-b) that every consistent total PI that satisfies a mild condition (called Archimedeaness) either is a linear total PI (called a scoring rule by Marchant) or is monotonically related to a linear total PI. This indicates that consistency and linearity are closely related properties and that there is little reason to use a total PI that is consistent but non-linear.

For every linear total PI f , there is a corresponding average PI g that is given by $g(\mathbf{x}) = f(\mathbf{x}) / x(0)$. We

call such an average PI a linear average PI. An example is the average number of citations per publication. The median number of citations of a set of publications does not have a corresponding linear total PI and is therefore an example of a non-linear average PI.

Acknowledgment

We would like to thank our colleagues at the Centre for Science and Technology Studies for their comments on this research.

References

- Egghe, L. (2006). Theory and practise of the g -index. *Scientometrics*, 69(1), 131–152.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *PNAS*, 102(46), 16569–16572.
- Jin, B., Liang, L., Rousseau, R. & Egghe, L. (2007). The R - and AR -indices: Complementing the h -index. *Chinese Science Bulletin*, 52(6), 855–863.
- Marchant, T. (in press-a). An axiomatic characterization of the ranking based on the h -index and some other bibliometric rankings of authors. *Scientometrics*.
- Marchant, T. (in press-b). Score-based bibliometric rankings of authors. *Journal of the American Society for Information Science and Technology*.
- Van Eck, N.J. & Waltman, L. (2008). Generalizing the h - and g -indices. *Journal of Informetrics*, 2(4), 263–271.
- Waltman, L. & Van Eck, N.J. (2009). A taxonomy of bibliometric performance indicators based on the property of consistency. Technical Report ERS-2009-014-LIS, Erasmus University Rotterdam, Erasmus Research Institute of Management

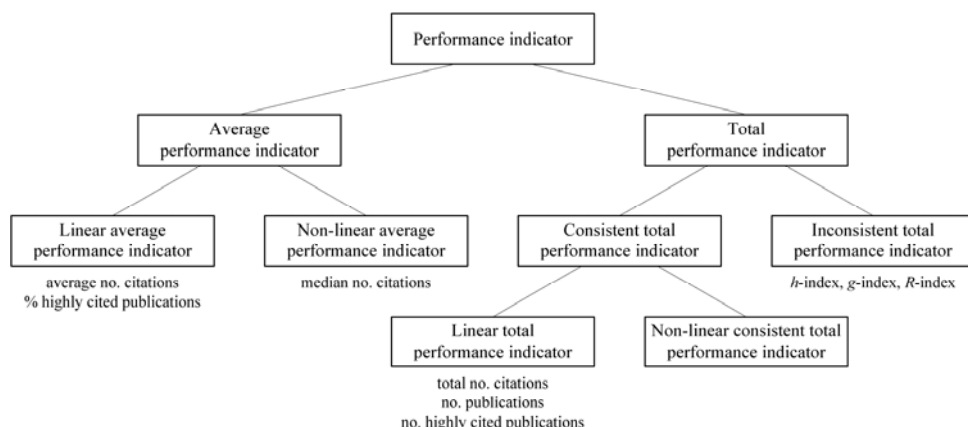


Figure 1. Taxonomy of bibliometric performance indicators.