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

Ludo Waltman and Nees Jan van Eck

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2009-014-LIS
Publication	March 2009
Number of pages	7
Persistent paper URL	http://hdl.handle.net/1765/15182
Email address corresponding author	lwaltman@ese.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	We propose a taxonomy of bibliometric indicators of scientific performance. The taxonomy relies on the property of consistency. The h-index is shown not to have this important property.
Free Keywords	scientific performance, bibliometric performance indicator, consistency, h-index, taxonomy.
Availability	The ERIM Report Series is distributed through the following platforms: Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal Social Science Research Network (SSRN), SSRN ERIM Series Webpage Research Papers in Economics (REPEC), REPEC ERIM Series Webpage
Classifications	The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems: Library of Congress Classification, (LCC) LCC Webpage Journal of Economic Literature, (JEL), JEL Webpage ACM Computing Classification System CCS Webpage Inspec Classification scheme (ICS), ICS Webpage

A taxonomy of bibliometric performance indicators based on the property of consistency

Ludo Waltman Nees Jan van Eck

Econometric Institute, Erasmus School of Economics
Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
E-mail: {lwaltman,nvaneck}@ese.eur.nl

Centre for Science and Technology Studies, Leiden University
P.O. Box 905, 2300 AX Leiden, The Netherlands

Abstract

We propose a taxonomy of bibliometric indicators of scientific performance. The taxonomy relies on the property of consistency. The h-index is shown not to have this important property.

Keywords

Scientific performance, bibliometric performance indicator, consistency, h-index, taxonomy.

1 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.

Our taxonomy is based on the mathematical properties of PIs. We note that relations between PIs can also be studied empirically. For recent studies in which an empirical approach is taken, we refer to Van Raan (2006), Costas and Bordons (2007, 2008), Bornmann, Mutz, and Daniel (2008), Bornmann, Mutz, and Daniel (in press), and Leydesdorff (in press).

2 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

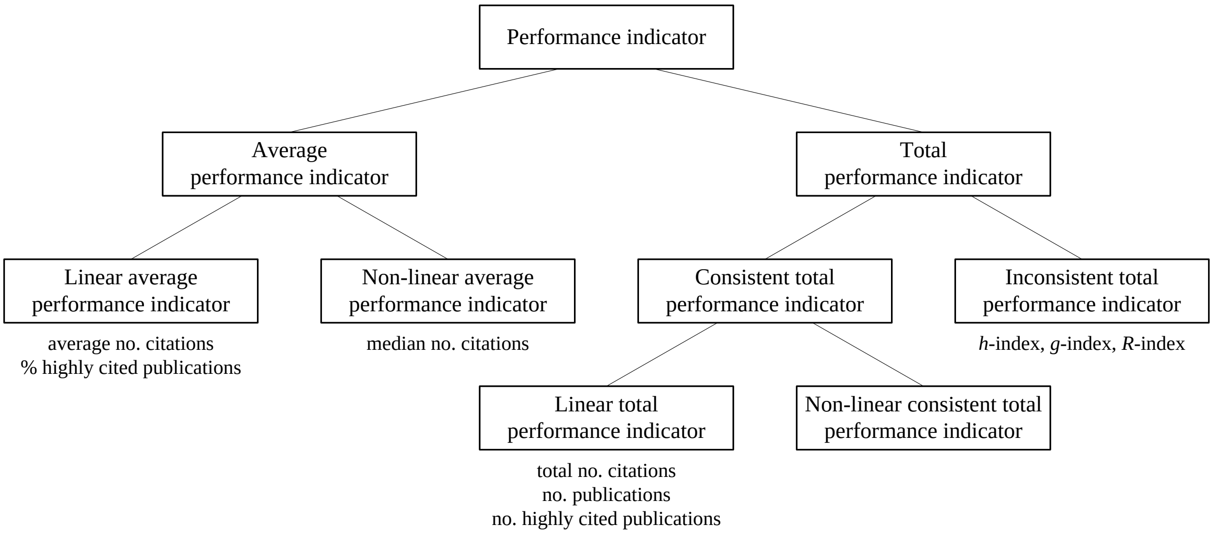

Figure 1: Taxonomy of bibliometric performance indicators.

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.

3 Taxonomy

Let $x_{\mathrm{A}}(i)$ denote the number of publications of researcher (or research group or journal) A with at least i citations, and let $\mathbf{x}_{\mathrm{A}} = (x_{\mathrm{A}}(0), x_{\mathrm{A}}(1), \ldots)$. We define a PI as a function that maps every $\mathbf{x} = (x(0), x(1), \ldots)$, 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 $\mathbf{x}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{B}}$

$$\frac{x_{\mathrm{A}}(i)}{x_{\mathrm{A}}(0)} \geq \frac{x_{\mathrm{B}}(i)}{x_{\mathrm{B}}(0)} \text{ for all } i \Rightarrow f(\mathbf{x}_{\mathrm{A}}) \geq f(\mathbf{x}_{\mathrm{B}}). \tag{1}$$

We note that the idea underlying (1) is similar to the idea underlying the decision-theoretic concept of first-order stochastic dominance. An example of an average PI is the average number of citations per publication (e.g., the journal impact factor). The percentage of uncited publications (Van Raan, 2004; Van Leeuwen & Moed, 2005) can also be seen as a kind of average PI. However, in the case of this indicator, a higher value indicates a lower rather than a higher performance. We define a total PI as a PI f such that for all $\mathbf{x}_\mathrm{A}$ and $\mathbf{x}_\mathrm{B}$

$$x_\mathrm{A}(i) \geq x_\mathrm{B}(i) \text{ for all } i \Rightarrow f(\mathbf{x}_\mathrm{A}) \geq f(\mathbf{x}_\mathrm{B}). \tag{2}$$

We note that a similar idea is discussed by Egghe and Rousseau (2008). 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), the m-index (Bornmann et al., 2008), and the h_w-index (Egghe & Rousseau, 2008), are neither average PIs nor total PIs. A PI can also be both an average PI and a total PI. A trivial example is given by $f(\mathbf{x}) = 1$ for all $\mathbf{x}$. Another example is the number of citations of the most cited publication (e.g., Lehmann, Jackson, & Lautrup, 2008; Marchant, in press-a).

Within the class of total PIs, we distinguish between consistent and inconsistent indicators. The property of consistency was discussed in Section 2. Mathematically, a consistent total PI is defined as a total PI f such that for all $\mathbf{x}_\mathrm{A}$, $\mathbf{x}_\mathrm{B}$, and $\mathbf{x}_\mathrm{C}$

$$f(\mathbf{x}_\mathrm{A}) \geq f(\mathbf{x}_\mathrm{B}) \Leftrightarrow f(\mathbf{x}_\mathrm{A} + \mathbf{x}_\mathrm{C}) \geq f(\mathbf{x}_\mathrm{B} + \mathbf{x}_\mathrm{C}). \tag{3}$$

We regard consistency as a very natural and highly desirable property. As demonstrated in Section 2, the use of an inconsistent total PI can lead to odd results. The h-index and its variants, such as the g-index (Egghe, 2006a, 2006b), the $h(2)$-index (Kosmulski, 2006), the R-index (Jin et al., 2007), the h_T-index (Anderson, Hankin, & Killworth, 2008), and the h_α- and g_α-indices (Van Eck & Waltman, 2008), are inconsistent. We consider this a serious shortcoming of these indicators.

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 $\mathbf{x}_\mathrm{A}$ and $\mathbf{x}_\mathrm{B}$

$$f(\mathbf{x}_\mathrm{A} + \mathbf{x}_\mathrm{B}) = f(\mathbf{x}_\mathrm{A}) + f(\mathbf{x}_\mathrm{B}). \tag{4}$$

The total number of citations of a set of publications is an example of a linear total PI. Other examples are the number of publications and the number of highly cited publications (e.g., Marchant, in press-a; Van Raan, 2004). 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 $f(\mathbf{x}_\text{A}) = 2f(\mathbf{x}_\text{B})$. 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 Archimedeanness) 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. An example of a non-linear consistent total PI is the square root of the total number of citations of a set of publications (e.g., Hirsch, 2007).

It is worth mentioning that every linear total PI f can be written as

$$f(\mathbf{x}) = \sum_{i=0}^{\infty} w(i)(x(i) - x(i+1)), \tag{5}$$

where the function $w : \{0, 1, \ldots\} \to [0, \infty)$ is non-decreasing. We call $w(i)$ the weight of a publication with i citations. It follows from (5) that every linear total PI is completely characterized by its weight function w. For example, the total number of citations of a set of publications is characterized by the weight function $w(i) = i$ and the number of publications is characterized by the weight function $w(i) = 1$.

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. Another example is the percentage of highly cited publications (e.g., Costas & Bordons, 2007, 2008). The median number of citations of a set of publications (e.g., Lehmann, Jackson, & Lautrup, 2006; Lehmann et al., 2008) 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

Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics*, *76*(3), 577–588.

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.

Bornmann, L., Mutz, R., & Daniel, H.-D. (in press). Do we need the h index and its variants besides standard bibliometric measures? *Journal of the American Society for Information Science and Technology*.

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*(3), 193–203.

Costas, R., & Bordons, M. (2008). Is g-index better than h-index? An exploratory study at the individual level. *Scientometrics*, *77*(2), 267–288.

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

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

Egghe, L., & Rousseau, R. (2008). An h-index weighted by citation impact. *Information Processing and Management*, *44*(2), 770–780.

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

Hirsch, J. E. (2007). Does the h index have predictive power? *Proceedings of the National Academy of Sciences*, *104*(49), 19193–19198.

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.

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.

Lehmann, S., Jackson, A. D., & Lautrup, B. E. (2006). Measures for measures. *Nature*, *444*, 1003–1004.

Lehmann, S., Jackson, A. D., & Lautrup, B. E. (2008). A quantitative analysis of indicators of scientific performance. *Scientometrics*, *76*(2), 369–390.

Leydesdorff, L. (in press). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology*.

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.

Van Leeuwen, T. N., & Moed, H. F. (2005). Characteristics of Journal Impact Factors: The effects of uncitedness and citation distribution on the understanding of journal impact factors. *Scientometrics*, *63*(2), 357–371.

Van Raan, A. F. J. (2004). Measuring science: Capita selecta of current main issues. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19–50). Kluwer Academic Publishers.

Van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.

Publications in the Report Series Research* in Management

ERIM Research Program: "Business Processes, Logistics and Information Systems"

2009

How to Normalize Co-Occurrence Data? An Analysis of Some Well-Known Similarity Measures
Nees Jan van Eck and Ludo Waltman
ERS-2009-001-LIS
http://hdl.handle.net/1765/14528

Spare Parts Logistics and Installed Base Information
Muhammad N. Jalil, Rob A. Zuidwijk, Moritz Fleischmann, and Jo A.E.E. van Nunen
ERS-2009-002-LIS
http://hdl.handle.net/1765/14529

Open Location Management in Automated Warehousing Systems
Yugang YU and René B.M. de Koster
ERS-2009-004-LIS
http://hdl.handle.net/1765/14615

VOSviewer: A Computer Program for Bibliometric Mapping
Nees Jan van Eck and Ludo Waltman
ERS-2009-005-LIS
http://hdl.handle.net/1765/14841

Nash Game Model for Optimizing Market Strategies, Configuration of Platform Products in a Vendor Managed Inventory (VMI) Supply Chain for a Product Family
Yugang Yu and George Q. Huang
ERS-2009-009-LIS
http://hdl.handle.net/1765/15029

A Mathematical Analysis of the Long-run Behavior of Genetic Algorithms for Social Modeling
Ludo Waltman and Nees Jan van Eck
ERS-2009-011-LIS
http://hdl.handle.net/1765/15181

A Taxonomy of Bibliometric Performance Indicators Based on the Property of Consistency
Ludo Waltman and Nees Jan van Eck
ERS-2009-014-LIS
http://hdl.handle.net/1765/15182

A Stochastic Dynamic Programming Approach to Revenue Management in a Make-to-Stock Production System
Rainer Quante, Moritz Fleischmann, and Herbert Meyr
ERS-2009-015-LIS
http://hdl.handle.net/1765/15183

Some Comments on Egghe's Derivation of the Impact Factor Distribution
Ludo Waltman and Nees Jan van Eck
ERS-2009-016-LIS
http://hdl.handle.net/1765/15184

---

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship