The Hirsch-Index and Related Impact Measures

L. Egghe

Universiteit Hasselt (UHasselt), Campus Diepenbeek, Agoralaan, B-3590 Diepenbeek, Belgium.

leo.egghe@uhasselt.be

Introduction

The Hirsch-index (or h-index), introduced in 2005 by Jorge Hirsch, is one of the most popular indicators in information science and in informetrics. Hundreds of articles have been written on the h-index or related h-type indices. To quote Ball (2007): "… the h-index does seem to be able to identify good scientists, and is becoming widely used informally, for example to rank applicants for research posts". A review on this important topic is hence not superfluous.

In the first section we give the definition of the h-index and present advantages and disadvantages of this index. To avoid some of the disadvantages of the h-index, several other, "h-type" indices (also called impact measures) have been introduced. These are discussed in the second section.

Applications of these indices are given in the third section. These applications deal with the definition and use of h-type indices to cases different from authors (e.g. journals or topics). Also an overview of case studies are given. In the fourth section we study impact measures from the point of view of their intrinsic properties they have or should have. We remark that it is not possible to exactly define what a good impact measure is. Axiomatic

characterizations of some impact measures are given. We also look at the influence of production on some impact measures.

In the fifth section we discuss some informetric models for these h-type measures (e.g. their dependence on the total number of articles and on the total number of citations). Also dynamical aspects of these indices are described: the influence of transformations on h-type indices. Also distributions of h-type indices are studied and it is noted that these are of a different nature than the distributions of the impact factor.

In the last section (before the conclusions section) we study h-type indices in function of time (e.g. the career of an author) and examine the possible shapes of these functions (several h-index sequences are studied). Case studies are presented. We close with some conclusions and open problems.

This review does not intend to describe many case studies (although an attempt is made to present a complete reference list up to the time of the manuscript's final submission) but focusses on the "philosophy" behind h-type indices as impact or performance measures and their potentialities as new informetric indicators. As always, this review contains some personal opinions.

The Hirsch-index. Definition. Advantages and disadvantages

We start with a historical note (communicated to this author by R. Rousseau) and to be found in Edwards (2005). It appears that the Hirsch index (of course not with this name) was defined some 35 years earlier by the astrophysicist sir Arthur Stanley Eddington as follows (in a communication of Eddington to the geophysicist Harold Jeffreys). In order to record his cycling prowess, Eddington records n, being the highest number of days on which he had cycled n or more miles. As will become clear below, this is nothing else than Hirsch's index

(but in cycling terminology). Of course, due to the widespread name "Hirsch index" or "h-index", we will no longer refer to Eddington or Jeffreys.

The definition of Hirsch (2005b) goes as follows: A scientist has index h if h of his/her N_p papers have at least h citations each, and the other $(N_p - h)$ papers have no more than h citations each. It must be remarked however that the first formulation by Hirsch (2005a) in the arXiv paper of August 17, 2005 had the word "fewer" instead of "no more" in the above definition. This was corrected into the definition above in the arXiv paper arXiv:physics/0508025v5 of September 29, 2005 which was then, subsequently, published as Hirsch (2005b). This correction was indeed necessary since, in the first formulation of August 17, the h-index does not always exist. Indeed, consider the following simple example: say an author has 5 papers: 4 of them have 3 citations and the fifth has 1 citation. Then $h = 3$ in the definition given above but when, in this definition, "no more" is replaced by "fewer" then the h-index does not exist (cf. also remarks on this in Glänzel (2006a) and Rousseau (2006b)). Note that the correction in Hirsch (2005b) was not noted in Anderson, Hankin and Killworth (2008) as well as in Lehman, Jackson and Lautrup (2008) and Sidiropoulos and Katsaros (2008).

Another, equivalent, definition is as follows. If we rank the papers of an author in decreasing order of the number of citations they received then this author's h-index is the highest rank $r = h$ such that the papers on ranks $1, 2, ..., h$ each have h or more citations.

Although we do not like the term, the papers on ranks $1, ..., h$ constitute the so-called h-core (introduced in Rousseau (2006b)). Note that this h-core is not really a core set of papers although they are the h most visible papers.

Table 1. Citation data of L. Egghe (October 22nd 2008, Web of Science)

rank	# citations
1	58
2	52
3	45
4	37
5	37
6	33
7	27
8	25
9	22
10	19
11	18
12	18
13	18
14	18
15	17
16	17
17	17
18	17
19	16
20	15
21	14
22	14
23	13
24	13
25	12

An example is presented based on this author's own publication and citation record (date: October 22nd 2008, using Web of Science): see Table 1: papers are ranked in decreasing order of received citations. From this it is clear that $h = 17$ (For reasons to become clear further we give the table up to rank 25).

What does the h-index measure ? In itself the h-index does not give a value to an informetric unity. This is in contrast with other informetric indicators such as the impact factor which, essentially, is an average number of citations per paper. The h-index is not an average, not a percentile, not a fraction: it is a totally new way of measuring performance,

impact, visibility, quality, … (see also the fourth section on this) of (e.g.) the career of a scientist (for other applications, see the third section). It is a simple measure without any threshold (e.g. for the Garfield-Sher impact factor one limits the citation period to two years which is – as is well-known – not optimal in some fields). This should explain the popularity of the h-index. The h-index is so popular that Scopus and Web of Science (WoS), less than two years after the introduction, have decided to present it as an indicator.

The h-index is a single simple measure that combines papers (an aspect of quantity) and citations (an aspect of quality i.e. impact). It is a robust measure in two ways: it is not influenced by a set of lowly cited papers nor by the (even severe) increase of citations to already highly cited papers. The first aspect is certainly an advantage of the h-index: a researcher should not be "punished" for writing some lowly cited papers, as long as this researcher writes highly cited papers as well. The fact that the h-index does not count the actual citations to papers is, in my view, a disadvantage of the h-index: once a paper belongs to the h-core it does not matter how many more citations this paper will receive. We feel that very highly cited papers should contribute more to an impact measure than less cited papers. Other measures have, therefore, been introduced: for this see the next section. To illustrate this insensitivity of the h-index we compared in 2006 the citation data of E. Garfield and F. Narin, see Egghe (2006b). The data can be found in Table 2.

As one readily sees, both scientists had h-index $h = 27$. But Garfield has not less than 14 papers with more than 100 citations (the one on rank 1 even has 625 citations) while Narin has only one paper with more than 100 citations. This clearly illustrates the insensitivity of the h-index and this is a clear disadvantage of this index.

Table 2.   Citation data and h-index of E. Garfield and F. Narin in 2006

TC = total # of citations to the paper on rank r

Garfield E.			Narin F.	
TC	**r**		**TC**	**r**
625	1		112	1
149	2		95	2
138	3		86	3
132	4		82	4
132	5		73	5
129	6		71	6
127	7		70	7
111	8		63	8
109	9		59	9
108	10		55	10
107	11		55	11
105	12		53	12
104	13		52	13
101	14		52	14
96	15		44	15
91	16		41	16
89	17		38	17
88	18		37	18
87	19		35	19
85	20		33	20
80	21		33	21
67	22		29	22
63	23		28	23
41	24		28	24
29	25		28	25
28	26		27	26
27	27		27	27
26	28		26	28
.	.		.	.
.	.		.	.
.	.		.	.

Of course, like any citation indicator, the h-index is field dependent and hence, h-indices of authors in different fields cannot be compared. Below on can find a list of top researchers in the fields physics, chemistry and computer science (Ball (2005)).

Physics:

1. Ed Witten, Institute for Advanced Study, Princeton   h = 110

2. Marvin Cohen, University of California, Berkeley   h = 94

3. Philip Anderson, Princeton University   h = 91

4. Manuel Cardona, Max Planck Institute for Solid State

         Research, Stuttgart, Germany   h = 86

5. Frank Wilczek,   Massachusetts Institute of Technology   h = 68

Chemistry:

1. George Whitesides, Harvard University   h = 135

2. Elias James Corey, Harvard University   h = 132

3. Martin Karplus, Harvard University   h = 129

4. Alan Heeger, University of California, Santa Barbara   h = 114

5. Kurt Wüthrich, Swiss Federal Institute of Biology, Zurich   h = 113

Computer Science:

1. Hector Garcia-Molina, Stanford University   h = 70

2. Deborah Estrin, University of California, Los Angeles   h = 68

3. Ian Foster, Argonne National Laboratory, Illinois   h = 67

4. Scott Shenker, International Computer Science

         Institute, Berkeley   h = 65

4. Don Towsley, University of Massachusetts, Amherst   h = 65

4. Jeffrey D. Ullman, Stanford University   h = 65

The differences are clear: the h-indices of the shown physicists are all below the ones of the shown chemists. The h-indices of the shown computer scientists are, in turn, at the bottom of the physicists ranking. For more on this field-dependence: see the next section.

New researchers have a clear disadvantage due to the short career length and they should not be compared, by their h-index, with researchers with a long career length. On the other hand established researchers may rest on their laurels since the number of citations received may increase even if no new papers are published. In other words: the h-index can never decrease. In the next section we will review measures that take into account career length.

Typical for citation analysis is the problem of how to treat self-citations (Smith (1981), Vinkler (1986), Egghe and Rousseau (1990)). This is in no way different for the h-index – see Schreiber (2007a) and Zhivotovsky and Krutovski (2008). Also a typical problem is how to deal with multi-authored papers (citing as well as cited). Several papers dealing with this problem are discussed in the next section. Both problems (self-citations and multi-authorship) are also discussed in Burrell (2007e). Finally one should also he aware that a researcher's h-index depends on the used citation database. One can use WoS or Scopus or Google Scholar: see Bar-Ilan (2008a), Jacsó (2008a,b,c,d,e), Meho and Yang (2007) and Meho and Rogers (2008). In Meho and Yang (2007) one shows that not only h-index values can change when one goes from one database to another but that, when comparing (ranking) h-indices of a set of researchers, the ranks can significantly change. Bar-Ilan (2008a) and Jacsó (2008a,c) compare the performance of WoS, Scopus and Google Scholar in various aspects. A similar study is conducted in Meho and Rogers (2008) showing the better coverage of Scopus e.g. yielding a better distinction between researchers as far as the h-index is concerned. Note also that the website http://www.harzing.com/pop.htm is a software that retrieves and analysis citation data based on Google Scholar: not only the h-index can be calculated but also several variants of it (to be introduced in the next section). This software was used in McKercher (2008).

As any indicator, the h-index is a single number, hereby reducing a scientist's career to a one-dimensional measurement. In the last section we will add a time dimension to the h-index (and other h-type indices).

The h-index is shown to be robust in the sense of errors in citation lists (Vanclay (2007)) or missing publications (Rousseau (2007a)).

Hirsch (2007) demonstrates that his h-index is a good indicator (better than total number of citations or papers or citations per paper (impact factor)) to predict future achievements.

More discussions on (dis)advantages of the h-index can be found in Costas and Bordons (2007a), Braun, Glänzel and Schubert (2005, 2006), van Raan (2006), Wendl (2007), Kelly and Jennions (2007), Sidiropoulos and Katsaros (2008), Bar-Ilan (2006), Glänzel and Persson (2005), Bornmann and Daniel (2005, 2007b), Egghe (2006a,b,c), Popov (2005), Glänzel (2006b), Antonakis and Lalive (2008), Liu and Rousseau (2007), Jin, Liang, Rousseau and Egghe (2007), van Leeuwen (2008), the short review articles Rousseau (2006b, 2008a), Bornmann and Daniel (2009) on h-type indices and the review articles on informetrics Bar-Ilan (2008c) and Thelwall (2008).

Variants of the h-index

As we mentioned already in the previous section the h-index has some clear disadvantages such as the insensitivity of the h-index to the actual number of citations to the articles in the h-core or such as the advantage that have researchers with longer careers (but it must be emphasized that it was Hirsch's explicite purpose to give an index that measures impact of "long" careers). Also multiple authorship, self-citations and field-dependence are issues that must be dealt with in connection with the h-index.

In fact, Hirsch himself introduced already in the first defining article Hirsch (2005b) a variant of the h-index: the m-quotient which is, simply, the h-index divided by the career length (= time since the publication of the first paper). A similar definition (in a time-dependent context − see also the last section) is given in Burrell (2007c) and is there called the h-rate. A variant in this "normalization" procedure in time is by not dividing the h-index by time but by the total number of papers of an author (called the normalized h-index). This is done in Sidiropoulos and Katsaros (2008). Rousseau (private communication) finds this a bad idea but thinks it is useful for journals (instead of authors) (see the next section for an application of the h-index to journals): in this way different years with different numbers of published articles can be compared, as shown in Rousseau (2006a).

Iglesias and Pecharromán (2007) claim that, if one divides the h-index by the average number of citations per paper, one obtains an index that can be used for inter-areas comparison. We are convinced that this index is more field independent than the h-index but we do not think it is useable as a completely field-independent index. Batista, Campiteli, Kinouchi and Martinez (2006) and Campiteli, Batista and Martinez (2007) seem to prove that dividing the h-index by the total number of authors in the considered h papers yields a relatively field-independent indicator. We are sceptic on the claim in Valentinuzzi, Laciar and Atrio (2007) that they have found two discipline-independent h-type indices.

Also in Radicchi, Fortunato and Castellano (2008) one tries to construct a field-independent h-index. Here one does not divide the h-index by the average number of citations per paper in the field (as in Iglesias and Pecharromán (2007)) but the actual number of citations of each paper of a researcher is divided by this average number. In addition one also divides the number of publications of a researcher by the average number of publications in the field. Using these "rescaled" numbers one applies the definition of the h-index, hereby obtaining the "generalized h-index".

The paper Qiu, Ma and Cheng (2008) also mentions the field-dependence of the h-index as well as the fact that the h-index is not very suited to evaluate the outputs of young researchers. They introduce the Paper Quality Index (PQI) which is a relative index based on the impact factor of a field and the total number of citations in a field. This index is hence based on the impact factor and is, therefore, not of h-type.

In Molinari and Molinari (2008a,b) one makes a log-log plot of the h-index versus number of papers in some journals (for countries) yielding a regression line of the form $\ln h = A - \beta \ln N$ (N = number of papers). Then h_m is defined as $h_m = e^A = \dfrac{h}{N^\beta}$ which is defined as the impact (or quality) index and which is claimed to be a correction for size. The name is not well-chosen since all indices discussed here are impact indices (see also the fourth section).

There are several papers dealing with modifications of the h-index in connection with co-authors. In Campiteli, Batista and Martinez (2007) and Batista, Campiteli, Kinouchi and Martinez (2006) one uses the h-index, divided by the average number of authors per paper in the h-core. This index is denoted h_I and equals $h / \overline{T}$ where $\overline{T} = T / h$, with T the total number of authors in the h-core, hence $h_I = h^2 / T$.

Independently, in Egghe (2008b) and Schreiber (2008a,b) the same measure h_m (in Schreiber (2008a,b)) and h_F (in Egghe (2008b)) is introduced using fractional paper counts (hereby replacing entire ranks by fractional ranks being 1 divided by the number of co-authors of the paper). Also in Egghe (2008b) the measure h_f is introduced whereby one uses fractional citation counts instead of fractional paper counts. In Egghe (2008b) a mathematical theory of h_f and h_F is presented and it is also noted that, in practise, the corrections h_f or h_F on h are not very large. Wan, Hua and Rousseau (2007) and Chai, Hua, Rousseau and Wan (2008) go even further and define the so-called "pure h-index" where one can even take

into account the scientist's relative position in the byline of an article. In this connection see also Kim and Seo (2007).

Kosmulski (2006) introduces the $h^{(2)}$-index which is a h-type index that is easier to calculate than the h-index since one needs a shorter list of papers in decreasing order of number of citations: an author has Kosmulski's index $h^{(2)}$ if $r = h^{(2)}$ is the highest rank such that all papers on ranks $1,...,h^{(2)}$ have at least $\left(h^{(2)}\right)^2$ citations. Going back to Table 1 this would mean $h^{(2)} = 5$ for this author. The numbers $h^{(2)}$ are hence much smaller than the h-indices and we feel that because of this reason, $h^{(2)}$ does not discriminate very well between authors. In addition, the calculation of the h-index is not very time-consuming and, as already mentioned, is calculated in WoS and Scopus. Notwithstanding this criticism on the $h^{(2)}$-index, we give, in the Applications section further on, an interesting application of the $h^{(2)}$-index to downloads

A generalization of the Kosmulski index is given in Levitt and Thelwall (2007). They define the Hirsch k-frequency $f(k)$ as the number of documents that are cited at least kh times. Note that $f(h) = h^{(2)}$, the Kosmulski index. The variable k enables to illustrate the distribution of the highly cited documents. They also introduce a "normalized" h index:

$$h_{norm} = 100 \frac{h^2}{T}$$ where T is equal to the number of documents of the set.

In Hu and Chen (2009) one introduces the so-called Major Contribution Index (MCI) which is the Hirsch-index of an author, based on his/her papers in which this author plays a major role (e.g. first or corresponding author). We do not see the value of this h-index variant: it is the number of citations that counts and this for any published paper of an author !

All the above variants of the h-index do not deal with the insensitivity of the h-index to the number of citations to the highly cited papers. This will be discussed now.

Note that – in our opinion – we do not need variants of the h-index dealing with the lowly cited papers. Hence we think that the v-index introduced in Riikonen and Vihinen (2008) (being h divided by the total number of papers) does not make much sense as an impact measure.

First of all we introduce the g-index, as given in Egghe (2006a,b,c). We remind the reader to the fact that the actual number of citations to the first 17 papers of this author (Table 1) is in fact of no importance (as long as this number is at least 17). The insensitivity of the h-index became also clear in Table 2 in the comparison of the careers (in 2006) of E. Garfield with F. Narin. Egghe notes that the h-index satisfies the property that the first h papers, together, have at least h^2 citations (indeed: in the h-core we have h papers each having at least h citations). Then Egghe defines the g-index as the largest rank with this property. In other words, with the same ranking of papers in decreasing order of number of citations received, we define the g-index as the highest rank such that the first g papers have at least g^2 citations together. In other words: the first g papers have at least g citations, on the average. Note that, per definition, $g \geq h$.

If we look at Table 1 we have that the total number of citations of the first 24 papers equals $580 > 24^2$ while the total number of citations to the first 25 papers equals $592 < 25^2$ so that $g = 24$ for this author (October 22nd, 2008). The same calculation is done in Egghe (2006b) on the (extended) citation data of Garfield and Narin (Table 2): despite the fact that their h-indices (in 2006) are equal: $h = 27$ we find in Egghe (2006b) that the g-index of Garfield is 59 while the one of Narin is 40, hereby showing the greater discriminatory power of the g-index in comparison with the h-index. This is also confirmed in Schreiber (2008c,d) and in Tol (2008) (the latter two articles in the connection of successive h- and g-indices: for this: see the third section).

This is also confirmed in Costas and Bordons (2008): especially for prolific scientists are the ranks based on the g-index lower than the ones based on the h-index and also, their g-index/h-index ratios are, on the average, higher.

Burrell (2007d) compares the g-index with the h-index and finds that they are both proportional to career length (see also the last section on time-dependent h-type sequences).

We refer to Rosenstreich and Wooliscroft (2009) for an application of the g-index in the ranking of management journals.

Another attempt to improve the insensitivity of the h-index to the number of citations to highly cited papers is the R-index, introduced in Jin, Liang, Rousseau and Egghe (2007). It is defined as

$$R = \sqrt{\sum_{i=1}^{h} c_i} \tag{1}$$

where documents, as usual, are ranked in decreasing order of the number of received citations, whereby c_i is the number of citations to the ith paper and where h is the h-index. So, as the g-index, also this measure takes into account the actual c_i-values in the h-core. It is an improvement of the A-index introduced by Jin (2006) (but the name was suggested by Rousseau (2006b), but see also Jin, Liang, Rousseau and Egghe (2007)) and defined as the average number of citations to papers in the h-core:

$$A = \frac{1}{h} \sum_{i=1}^{h} c_i \tag{2}$$

This measure has the undesirable property than an increase in the number of citations might lead to a decrease of the A-index (see further in the fourth section on desirable properties of impact measures). Indeed: take a first situation where four papers have 6, 5, 4, 3 citations respectively. Hence $h = 3$ and $A = \frac{1}{3}(6+5+4) = 5$. Now add one citation to the fourth paper:

6, 5, 4, 4. Now $h = 4$ and $A = \frac{1}{4}(6+5+4+4) = \frac{19}{4} < 5$, which is not a good property: more citations should not lead to a strict decrease of an impact measure. Note that the g- and R-index do not suffer from this deficiency. The square root in (1) is necessary to conform with the h-index in case all $c_i = h, \ i = 1,...,h$. Then $R = h$ as is readily seen.

In Glänzel (2008c) one compares the h-index and A-index with so-called "characteristic scores and scales " (CSS) which are mean citation rates of subsets of papers of an author. One example of a CSS is the average number of citations of those papers which have at least a number of citations larger than or equal to the overall average number of citations.

In Sezer and Gokceoglu (2009) one defines the so-called "fuzzy academic performance index" which is built around 3 indicators: total number of publications, total number of citations and academic life span and one notes a high correlation of this index with the h-index.

It can be seen from Table 1 that the R-index of this author is $R = 21.86$, the square root of the sum of the number of citations to the $h = 17$ most highly cited articles.

In order to normalize for the age of an article (which is more advantageous for attracting more citations), Jin, Liang, Rousseau and Egghe (2007) introduce the AR-index:

$$AR = \sqrt{\sum_{i=1}^{h} \frac{c_i}{a_i}} \qquad (3)$$

where h and c_i are as above and a_i is the age of article i. The AR-index is an approach to define an h-type index that can actually decrease. In this way authors cannot "rest on their laurels". The AR-index is further discussed in Rousseau and Jin (2008), Jin (2007) and Jin and Rousseau (2007) and the R- and AR-indices are commented in Glänzel (2007b). Järvelin and Persson (2008) criticize the AR approach and propose another way of taking age into account (but their proposal is not an h-type index).

In Rousseau and Jin (2008), an alternative for (3) is given: instead of the square of (3):

$$\sum_{i=1}^{h} \frac{c_i}{a_i} \qquad (4)$$

one could also define

$$AR_1^2 = \frac{\sum_{i=1}^{h} c_i}{\sum_{i=1}^{h} a_i} \qquad (5)$$

similar in (mathematical) form as the meta-journal average impact factor respectively the meta-journal global impact factor; their difference (and hence, indirectly, the difference between (4) and (5)) is studied in Egghe and Rousseau (1996a,b). However, Rousseau and Jin (2008) indicate some deficiencies of the index defined in (5).

Another way of taking the citation scores of the (highly cited) papers into account is to weight the articles according to their citations: this is the weighted h-index denoted h_w - see Egghe and Rousseau (2008) for the details. The discrete h_w-index suffers from the same problem as the A-index but it is shown in Egghe and Rousseau (2008) that the continuous model does not suffer from this aberration.

Another attempt to solve the problem of the insensitivity of the h-index to the number of citations to the highest cited articles is presented in Anderson, Hankin and Killworth (2008). Although it is a non-practical solution (they have to take into account all citations to all papers) there is some theoretical interest in their proposal. Their inspiration comes from a theory of partitions, described in Andrews (1998). A table is made as follows (called a Ferrers graph): papers are vertically ranked in decreasing order of received citations. Instead of writing down the number of their received citations, dots (equal to this number) are shown horizontally, per paper. From the upper left corner, squares of side length $1, 2, 3, ...$ are made. First the square of side-length 1 (consisting of the first citation dot to the first paper) is made

and this point receives a weight of 1. The square of side-length 2 is made by adding to this square of side-length 1 the second citation (dot) to the first paper and the first two citations (dots) to the second paper. So we add 3 citations each receiving a weight of $\frac{1}{3}$ The next square with side-length 3 consists of the square of length 2 above, extended with 5 citations (each receiving a weight of $\frac{1}{5}$): The third citation (dot) to the first paper and to the second paper and the first three citations (dots) to the third paper, and so on. We can continue like this until all citations to all papers are covered. For an example, see Fig. 1.

<pre>
 1   1/3   1/5   1/7   1/9
 ●   ●   ●   ●   ●

1/3   1/3   1/5
 ●   ●   ●

1/5   1/5   1/5
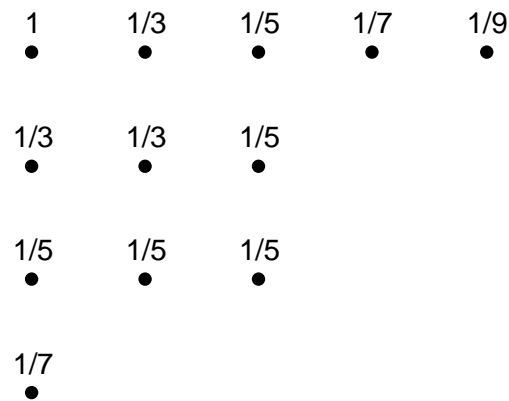 ●   ●   ●

1/7
 ●
</pre>

Fig. 1.   Ferrers graph of (5,3,3,1): 4 papers with

respectively 5, 3, 3 and 1 citation(s)

The sum of all the weights of given citations is the so-called "tapered h-index", denoted h_T and is a (theoretical) natural extension of the h-index h. Indeed, the h-index h equals the side-length of the largest completed (filled in) square of points, being also equal to the sum of all the scores in this square (called the Durfee square) (in Fig. 1: $h = 3$). Hence, clearly $h \pounds h_T$ (in Fig. 1, $h_T = 3 + \frac{2}{7} + \frac{1}{9}$) but h_T takes into account ("in a h-type way") all citations to all papers (as said, this is non-practical and hence a disadvantage of the tapered h-index h_T).

In Egghe (2008e) we show how these Ferrers graphs can be transformed so that they yield the g-index in a natural (theoretical) way. The so-called conjugate sequence of a partition (decreasing sequence of coordinates of a vector) is considered and one shows that its h-index is the same as the h-index of the original partition. Lorenz-curves (Egghe (2005)) for partitions and their conjugates are studied, both in a discrete and continuous (introduced in Egghe (2008e)) way.

In van Eck and Waltman (2008) generalizations of the h- and g-index are studied, involving an extra parameter α. The generalized h-index h_α is the highest rank such that all the papers on this rank (and lower ranks) have at least αh citations. Note that $h_1 = h$. Similarly the generalized g-index g_α is defined as the highest rank such that the first g_α papers have, together, at least αg^2 citations. Note again that $g_1 = g$. Properties for varying α are demonstrated. Independently, Levitt and Thelwall (2007) also define h_α, but in another notation: they define the Hirsch-frequency $f(k)$ as the number of documents in the collection that are cited at least kh times. Hence $f(k) = h_\alpha$ for $\alpha = k$.

Deineko and Woeginger (2008) and Woeginger (2008d) even generalize the above by replacing αh and αg^2 by general functions $s(k), k = 1, 2, ...$. For axiomatic characterizations of these indices (including the h- and g-index itself and other impact measures), see the fourth section. For informetric properties of h-type indices, see the fifth section.

Nine of these indices (h, m quotient of Hirsch, g, $h^{(2)}$, A, R, AR, h_w and a new m-index) are compared in Bornmann, Mutz and Daniel (2008). The m-index is the median number of citations received by papers in the h-core. They classify (using factor analysis) indices as

- describing the most productive core of the output of a scientist (such as h, m quotient, g and $h^{(2)}$)

- depicting the impact of the papers in the core (such as A, m-index, R, AR and h_w).

(see also Bornmann, Mutz, Daniel, Wallon and Ledin (2008)).

Furthermore, as argued in Bornmann, Mutz and Daniel (2009), the classical bibliometric indicators "number of publications" and "total citation counts" load so high on the above mentioned two factors that, in this (statistical) vision, h-type indices do not necessarily have to be used. Similar statistical conclusions are found in several papers discussed in the next section.

In Bornmann, Mutz and Daniel (2007) one introduces the b-index of a scientist, being the number of papers of this scientist that belong to the top 10% of papers in a field. Since this index uses a threshold value, we consider this as a non-h-type measure.

In Leydesdorff (2009), the h-index is statistically (using PCA = Principal Components Analysis) compared with non-h-type indices (such as PageRank, impact factor, Scimago Journal Ranking, network centrality measures,…). It is found that the h-index combines the two dimensions (size and impact).

Rational interpolations of the h-index (i.e. between h and $h+1$) are introduced in Ruane and Tol (2008), see also Tol (2008) and further studied in Guns and Rousseau (2009a) (also for the g-index). In Rousseau (2008d), real-valued h-indices and g-indices are introduced and further studied in Guns and Rousseau (2009a). They are based on the piecewise linearly interpolated citation curve.

Other h-type indices can be found in Sidiropoulos and Katsaros (2008), Rousseau, Guns and Liu (2008), Kosmulski (2007), Boell and Wilson (2008) and Ruane and Tol (2008).

The next section is devoted to applications of h-type measures (not yet in a time-dependent context – for this, see the last section) and some case studies.

Applications

It is clear from the original definition by Hirsch (2005b) that the h-index (and hence also the other h-type indices) has been defined in order to have a simple informetric indicator of the impact of a researcher's career. Worldwide the h-index (and some of its variants) is used for this purpose and we mentioned already that the h-index is produced in WoS and Scopus. Researchers in the same field (and equally long careers) can be compared with the h-index. The h-index can also be used to predict future achievements (and for deciding on tenure positions or grants) (Hirsch (2005b)).

Soon after the introduction of the h-index, Braun, Glänzel and Schubert (2005, 2006) have noticed the important fact that the h-index can also be applied to journals: the same definition applies to a ranked set of articles (in decreasing order of the number of citations they received) of the journal (e.g. in a certain period or all articles of the journal). The same goes, of course, for the other h-type indices as discussed above. This leads to the constatation that the journal's impact factor (IF) has got a serious competitor, especially since IF is dependent on the defined citing and publication period: this arbitrariness in the definition of IF is not there in the h-index and its similar types. The superiority of the h-index over IF for physics is claimed in Miller (2006) and a similar preference is expressed in Vanclay (2007) although in Vanclay (2008b), high correlation is found between journal rankings based on IF and rankings based on the h-index. The h-index for journals, based on Google Scholar, has been found in Harzing and van der Wal (2009) to be a better measure of journal impact than the IF itself. A comment on Braun, Glänzel and Schubert (2005) is given in Vanclay (2006).

In general one finds comparisons of the h-index with other bibliometric indicators such as total number of publications, total number of citations, average number of citations per publication (IF), some other indicators and even peer review for authors, journals or groups of authors in Bornmann, Wallon and Ledin (2008), van Raan (2006), Costas and Bordons

(2007a) and Bornmann and Daniel (2007a). In Wan, Hua, Rousseau and Sun (2009) the h-index for journals is compared with a new indicator, the "download immediacy index" (DII) defined as the number of downloads of a journal's articles within one publication year, divided by the number of published articles by that journal in that same year (introduced in Wan et al. (2007)). The conclusions are, in general, the same: all these rankings correlate from well to high. From this point of view, research evaluation does not need the h-index (!) although all authors agree that the h-index is a valuable new simple tool that, due to the above mentioned correlation can, to a certain extent, replace the other indicators. Bornmann, Marx and Schier (2009) reach a similar conclusion: h-type indices correlate highly with IF and hence, from a statistical point of view, are empirically redundant. Albeit not in the context of the h-index, Reedijk and Moed (2008) come to the conclusion that the "impact" of impact factors is decreasing which, indirectly leads to a higher importance of the h-index for journals. However, Honekopp and Kleber (2008) reach an apposite conclusion as far as the prediction of future success of an article is concerned.

Besides applications to authors, groups of authors (e.g. institutes) and journals there are possible applications to all source-item relations in an information production process (IPP) (cf. Egghe (2005)). An IPP is a set of sources (e.g. authors, journals,…) that produce items (e.g. articles) where one indicates which sources produce which items (see Egghe and Rousseau (1990) or Egghe (2005)) for many more examples). Hence these sources can be ranked in decreasing order of their number of items and hence the h-index (or any h-type measure) of this IPP can be calculated – see also Egghe (2008i).

In this sense the paper Banks (2006) is very interesting since it defines the h-index for (scientific) topics and compounds. In this way "hot" topics can be detected in diverse research areas. This idea is applied by Bar-Ilan on the h-index of the topic "h-index" – see Bar-Ilan (2007, 2008b), by Egghe and Rao (2008b) by The STIMULATE6 Group (2007) (including

Rousseau), in Jin and Rousseau (2008), in Rousseau (2007b) and in Minasny, Hartemink and McBratney (2007) (on soil science).

The h-index for countries (on various science fields) is studied in Csajbók, Berhidi, Vasas and Schubert (2007).

The case of the h-index for book classifications and the borrowings of the books in these classifications is an example where the h-index can be used in library management. It can be found in Liu and Rousseau (2007, 2008b).

In Guan and Gao (2009) one applies the h-index to citations to patents of companies.

Since the h-index can be applied in any IPP, Egghe and Rao (2008a) studied different h-indices for groups of authors. If we look at a group of authors (e.g. in an institute or a field) we can arrange these authors in decreasing order of their number of publications. The corresponding h-index is denoted h_P. If we arrange these authors in decreasing order of the total number of received citations, we arrive at the h-index h_C. In Rousseau and Rons (2008) a similar h-index is proposed but now using the total number of different citing papers (instead of all received citations). If we consider the group of authors as one meta-author and if we arrange the papers of this meta-author in decreasing order of their number of received citations we have the global h-index h_G, also studied earlier in van Raan (2006) and mentioned in Prathap (2006). Interrelations between these h-indices are proved in Egghe and Rao (2008a) where also the successive h-index h_2 is studied. This successive h-index was introduced – independently – by Prathap (2006) and Schubert (2007) and we will define it now.

Suppose we have a group of authors, e.g. an institute and that each author has an h-index. If we order these authors in decreasing order of their h-indices then we can define the h-index of this list (denoted h_2) as the largest rank such that the first h_2 authors have an h-

index of at least h_2. This is called the successive h-index since it was derived from the h-indices of the authors. In Schubert (2007) a slightly different example is given by replacing authors by journals (e.g. of a publisher) and whereby the h-indices of these journals (cf. their introduction given earlier and see also Braun, Glänzel and Schubert (2005, 2006)) determine the successive h-index h_2 of this publisher. One can even go to another higher aggregation level by, using the h_2-indices, determining the successive h-index h_3 of a country. In Egghe and Rao (2008a), also the successive h-index h_2 was studied in comparison with the earlier defined h_P, h_C and h_G (see above) for a group of authors (e.g. an institute). For a description of the modelling of successive h-indices we refer to the fifth section or see Egghe (2008f).

An application of successive h-indices to the field of economics in the Republic of Ireland is given in Ruane and Tol (2008). In Tol (2008) the successive g-index is introduced and applied to the same topic. As mentioned above, Tol (2008) finds that "the successive g-index has greater discriminatory power than the successive h-index". Arencibia-Jorge and Rousseau (2008) apply successive h-indices to Cuban institutions. A similar application is given in Arencibia-Jorge, Barrios-Almaguer, Fernández-Hernández and Carvajal-Espino (2008). Rousseau, Guns and Liu (2008) study the (successive) h-index in the generalized framework of a conglomerate (which is a slight extension of an IPP). In da Silva et al. (2009) one uses successive h-indices in the construction of a relative h-index for groups (amongst a set of groups). They combine this relative h-index for groups with the Gini index (see Gini (1909)) of the group (based on its members' h-indices) to yield the α-index. The goal of this α-index is to measure the quality of a group (and this α-index is independent of the size of the group).

The problem of author names recognition for h-index calculations has been studied in Jacsó (2007).

h-indices have lots of applications. They can not only be used with citations as ingredients but also with downloads. This has been applied (as a fringe topic) in O'Leary (2008). More inspirational is the paper Hua, Rousseau, Sun and Wan (2009) where one explores the same idea but where one notices that, in general, downloads yield higher numbers than citations and, therefore, one does not use the h-index but the Kosmulski $h^{(2)}$-index (described above). We can consider this as an interesting use and advantage of the $h^{(2)}$-index.

One could determine h-indices of web sites (based on their in- or out-links). We even feel (although there are no references yet - to the best of my knowledge) that the h-index (and h-type indices) could be used in econometrics and other -metrics field as a new assessment indicator, e.g. on wealth or income.

In general we can talk, nowadays of a "3-tier" evaluation system (peer review, citations, downloads) in a "2-tier" communication system (pre-publication in a repository (e-print server), publication in a (e-)journal). The British government has announced that, after 2008, it will base funding assessments for universities based on h-type indicators (Ball (2007)).

We mentioned already the methodological applications of the h-index to researchers, journals, institutes and topics. Further applications (case studies) are mentioned now, classified per broad topic.

Authors, Career assessment: Kelly and Jennions (2006, 2007), Sidiropoulos and Katsaros (2008), Saad (2006), Oppenheim (2007), Vanclay (2007), Sanderson (2008), Bornmann, Wallon and Ledin (2008), Bar-Ilan (2006), Glänzel and Persson (2005), Vinkler (2007), Minasny, Hartemink and McBratney (2007), Batista, Campiteli, Kinouchi and Martinez (2008), Lovegrove and Johnson (2008), Sangam and Girji (2008), Rodriguez, Janssens, Debackere and De Moor (2007, 2008), de Araujo (2008), Mugnaini, Packer and Meneghini (2008), Jones (2008), Engqvist and Frommen (2008), Machacek and Kolcunova (2008),

Aguayo-Albasini and Campillo-Soto (2008), Heaney (2007), Hermes-Lima, Alencastro, Santos, Navas and Beleboni (2007), Imperial and Rodriguez-Navarro (2007), Rau (2007), Salgado and Paez (2007), Sidiropoulos, Katsaros and Manolopoulos (2007), Ursprung and Zimmer (2007), Schreiber (2007b), Torro-Alves, Herculano, Tercariol, Filho and Graeff (2007), Harnad (2007), Romero, Cortés, Escudero, López and Moreno (2007), Rao (2007), Costas and Bordons (2007b), van Haselen (2007), Ferrand (2007), Cronin and Meho (2006), Brahler and Decker (2005), Frangopol (2005), McKercher (2008), Purvis (2006) (a 20-line note), Timofieiev, Snasel and Dvorsky (2008), Dorta-Contreras, Arencibia-Jorge, Marti-Lahera and Araujo-Ruiz (2008), van Leeuwen (2008), Lufrano and Staiti (2008), Dodson (2008), Levitt and Thelwall (2009), Sebire (2008), Kellner (2008), Kellner and Ponciano (2008), Humberto Romero, Garcia and Kiwi (2009).

Journals: Saad (2006), Vanclay (2007, 2008a,b), Rousseau (2006a), Liu, Rao and Rousseau (2008), Barendse (2007), Satyanarayana and Sharma (2008), Jeang (2007), Sidiropoulos, Katsaros and Manolopoulos (2007), Olden (2007), Rousseau and Small (2005), Orbay and Karamustafaoğlu (2008), Krauskopf and Krauskopf (2008), Levitt and Thelwall (2008), Sebire (2008), Gracza and Somoskovi (2008), Harzing and van der Wal (2009).

Research groups and institutions: van Raan (2006), Molinari and Molinari (2008a,b); the methodological aspects where already described in the previous section, Torro-Alves, Herculano, Tercariol, Filho and Graeff (2007), Mugnaini, Packer and Meneghini (2008), de Araujo (2008), Kinney (2007), Grothkopf and Stevens-Rayburn (2007), Pirez da Luz et al. (2008).

Countries: Mahbuba and Rousseau (2008), Lufrano and Staiti (2008), Pilc (2008).

Topics: Minasny, Hartemink and McBratney (2007), Rousseau (2007b), The STIMULATE6 Group (2007), Marx and Barth (2008).

Library management: Liu and Rousseau (2007, 2008b).

It is clear that the majority of case studies on h-type indices is on authors (researchers) and/or career assessment, which is in line with the original purpose of the h-index as defined by Hirsch (2005a,b).

We end this section with an intriguing question: if we consider "all" publications of all authors as if they were written by one meta-author M, what is then the h-index of M ? This h-index could then be considered as the h-index of "humanity". I know that – in essence – the answer to the above problem has no real value e.g. because we mix all scientific disciplines which is not allowed when dealing with citation analysis (and hence when dealing with the h-index). So the above question should be merely considered as an intriguing challenge.

Of course one has to define what it means "all" publications. We have limited ourselves to WoS. Henry Small (2007) (personal communication in May 2007) reported to us (and this information was published in Egghe and Rao (2008b) with Small's permission) that, dependent on the used time period (up to 1955 or up to 1972) the h-index of WoS (standing for "all" publications) is between 1,500 and 2,000. Note that in May 2007, WoS (up to 1972) contained a bit more than 33.10^6 documents of which about 80% have received at least one citation (Small (2007), personal communication).

Theories of impact and performance measures

In Glänzel (1996), supported by Rousseau (2002), Glänzel explains the need for standards in informetric research. This is certainly lacking when introducing a new topic like "h-type indices".

In this section we will examine theoretical properties of h-type indices so that we can learn to "understand" the real nature of these indices. But first of all: "what do we want ?" In other words what do we expect from an impact or performance measure (also names should be standardized, so I suggest, from now on, only to use the term "impact measure").

Let us compare this with concentration measures (i.e. measures capable of measuring concentration or inequality) (Egghe (2005), Chapter 4). Its introduction was in econometrics, a bit more than 100 years ago. Concentration measures have little to do with impact measures but act, like impact measures, on a decreasing sequence $x_1, x_2, ..., x_N$ of positive numbers. Also, further in this section, we will show that the g-index has a concentration property. So we will briefly introduce the notion of concentration measure and we will see that we will be able to define, in an exact way, what we expect from a concentration measure. Then we will go back to impact measures and we will investigate whether such an exact "wish list" can also be given for impact measures.

Concentration theory was invented in econometrics by Lorenz in 1905 (Lorenz (1905)). It has also found its way into informetrics since – as in econometrics – many phenomena in informetrics are very concentrated in the sense that (in general terminology) few sources have many items and many sources have few items (e.g. few papers receive many citations and many papers receive few (or no) citations (Egghe (2005)), Chapter 1 or see also the next section).

Lorenz concentration theory goes as follows. Let $X = (x_1, x_2, ..., x_N)$ be a decreasing vector with positive coordinates x_i, $i = 1, ..., N$. The Lorenz curve $L(X)$ of X is the polygonal curve connecting $(0,0)$ with the points $\left(\dfrac{i}{N}, \displaystyle\sum_{j=1}^{i} a_j \right)$, $i = 1, ..., N$, where

$$a_i = \frac{x_i}{\displaystyle\sum_{j=1}^{N} x_j} \tag{6}$$

Note that for $i = N$ we have $(1,1)$ as end point of $L(X)$. Let X and $Y = (y_1, y_2, ..., y_N)$ be two such vectors. We say that X is more concentrated than Y if $L(X) > L(Y)$, i.e. if the

Lorenz curve of X is strictly above the one of Y (except, of course, in the common points $(0,0)$ and $(1,1)$). We also say that the coordinates of X are more unequal than the ones of Y.

That Lorenz curves are the right tool to measure concentration (or inequality as a synonym) is seen by the result of Muirhead (1903) stating[1] $L(X) > L(Y)$ if and only if X is constructed, starting from Y, by a finite number of applications of elementary transfers. An elementary transfer (e.g. on $Y = (y_1, ..., y_N)$) changes Y into the vector

$$\left(y_1, ..., y_i + a, ..., y_j - a, ..., y_N \right) \tag{7}$$

where $1 \leq i < j \leq N$ and $a > 0$. Since Y is decreasing, this means, in econometric terms that "we take away $(a > 0)$ from the poor (j) and give it to the rich (i)" which indeed yields a more unequal (concentrated) situation, which is applied repeatedly to yield X out of Y.

Now we define a function f, acting on N-dimensional vectors $X = (x_1, ..., x_N)$ with values $f(X) = f(x_1, ..., x_N) \geq 0$ to be a good concentration measure if

$$L(X) > L(Y) \text{ implies } f(X) > f(Y) \tag{8}$$

Well-known good concentration measures are the variation coefficient (i.e. the standard deviation of X divided by its average) or the well-known Gini-index (Gini (1909) or see Egghe (2005), Chapter 4).

So we were able to define in a mathematically exact and unique way what we expect from a concentration measure. We are, at the time being, unable to do the same for impact measures. Indeed, as will be seen in the sequel, the different impact measures we have defined so far each have different impact properties and it is proved (see later) that there does not exist a measure that satisfies all of these impact properties. Let us start with the work of

---

[1] Muirhead's theorem was published in 1903, two years before Lorenz introduced the Lorenz curve (Lorenz (1905)). Muirhead's theorem hence did not use the Lorenz terminology but a combinatorial variant of it. Here we present the Lorenz variant of Muirhead's theorem. Muirhead's theorem can also be found in Hardy, Littlewood and Pólya (1952) and in Egghe and Rousseau (1991).

Woeginger on this. In Woeginger (2008a) the following definition of an impact measure (there called a scientific impact index) is given: it is a function f, acting on vectors (throughout the sequel we assume that the vectors X, Y, … are decreasing (i.e. their coordinates are in decreasing order)) $X = (x_1,...,x_N)$ and $Y = (y_1,...,y_M)$ (note that now M can be different from N) such that

    (i) $f(0,...,0) = 0$ (in any dimension, including 0)

    (ii) $X \leq Y$ implies $f(X) \leq f(Y)$. Here $X \leq Y$ is defined as follows: $X \leq Y$ if

        $N \leq M$ and $x_i \leq y_i$ for $i = 1,...,N$.

It is clear that (i) and (ii) are necessary conditions for an impact measure but that they are far from sufficient: a zero vector has no impact (in other words: if an author has no papers with citations or has no papers at all, the impact of this author is zero and if on every common coordinate (say rank) of two authors X and Y the paper (on that rank) of Y has more $(\geq)$ citations than the paper (on that rank) of X, then Y has more impact than X. These requests for an impact measure are non-controversial but are not enough to call f a good impact measure; in fact I had preferred that Woeginger had chosen another name for f satisfying (i) and (ii) (cf. the requested standardization in informetrics mentioned above!) but we will keep the name in order not to confuse the reader.

Woeginger (2008a) then continues his search for "desired elementary properties" (other than (i) and (ii)) for impact measures. He formulates five "fairly natural axioms" that are desired for impact measures f (a so-called wish-list as was also formulated in a totally different context in Cater and Kraft (1989) in the context of information retrieval). Again the vectors X, Y are decreasing.

    A1.   If the $(N+1)$-dimensional vector Y results from the N-dimensional vector

        X by adding a new article with $f(X)$ citations then $f(Y) \leq f(X)$.

A2. If the $(N+1)$-dimensional vector Y results from the N-dimensional vector X by adding a new article with $f(X)+1$ citations, then $f(Y)>f(X)$.

B. If the N-dimensional vector Y results from the N-dimensional vector X by increasing the number of citations of a single article, then $f(Y)\leq f(X)+1$.

C. If the N-dimensional vector Y results from the N-dimensional vector X by increasing the number of citations of every article by at most 1, then $f(Y)\leq f(X)+1$.

D. If the $(N+1)$-dimensional vector Y results from the N-dimensional vector X by first adding an article with $f(X)$ citations and then increasing the number of citations of every article by at least 1, then $f(Y)>f(X)$.

Although a bit artificial, these are all "desirable" properties of a good impact measure. Only A_1 looks a bit controversial but this axiom shows the needed robustness of an impact measure: it is better understood as: adding an article with $f(X)$ citations will not increase $f(X)$. In fact A_1 and (ii) together imply $f(X)=f(Y)$. Also axioms B and C are not completely natural: why should $f(Y)$ be maximally one unit larger than $f(X)$? These axioms clearly are meant to describe the h-index and certainly not other h-type indices such as the g- or R-index !

Woeginger (2008a) shows $A_2 \Rightarrow D$ but also that A_1 and $A_2 \Rightarrow$ not B, so that there is no impact measure that can satisfy all the above axioms !

He then continues to prove that an impact measure f is the h-index if and only if f satisfies A_1, B and D. This learns us a lot on the true nature of the h-index as impact measure. Other, less important measures are also characterized using some of the above axioms.

A similar characterization of the h-index (but now for the real-valued version of it, i.e. with values in the positive real numbers) has been given in Quesada (2009).

In Woeginger (2008b) another characterization of the h-index is given. First we introduce a new axiom.

S.   For every decreasing vector X, we have $f(X) = f(R(X))$ where $R(X)$ is the mirrored vector $X = (x_1,...,x_N)$ (depicted in the plane with the indices (papers) as abscissa and the x_i (citations) as ordinates) over the first bissectrix (the 45° line).

Then Woeginger proves that an impact measure f is the h-index if and only if f satisfies A_1, S and D.

In all these characterizations it is easy to show that h satisfies these axioms; the difficult part is the reverse (if) part. Axiom S is logic in the connection with the definition of the h-index.

Characterizing the g-index is an even more difficult task that was also performed by Woeginger (2008c). In this paper he adds to (i) and (ii) above the trivial requirement that an impact measure should also satisfy: if $X = (x_1,...,x_N)$ and $Y = (x_1,...,x_N,0)$, then $f(X) = f(Y)$. This is indeed logical: adding a new paper without citations cannot increase (or decrease) the impact. It is a bit strange that Woeginger keeps the same name for a measure f satisfying these three axioms as one that satisfies only (i) and (ii) (standardization problem again !). For characterizing the g-index, Woeginger needs other "desirable" properties than in the case of the h-index, showing the different nature of the g-index. The three needed axioms are (X, Y are decreasing)

E.   If the $(N+1)$-dimensional vector Y results from the N-dimensional vector X by first adding an article with $f(X)$ or $f(X)+1$ citations and then

increasing the number of citations of every article by one, then

$$f(Y) = f(X) + 1.$$

Axiom E is of the same nature as axiom D. The difference is needed for technical reasons. Of a completely different nature are the following two axioms (again for decreasing vectors X and Y).

T_1.   Let $X = (x_1, ..., x_N)$ and let $1 \le i < j \le N$. If the vector Y results from X by setting $x_k = y_k$ for all $k = 1, ..., N$ but $k \ne i$, $k \ne j$ and $y_i = x_i + 1$ and $y_j = x_j - 1$, then $f(Y) \ge f(X)$.

T_2 Let $X = (x_1, ..., x_N)$ and let $1 \le i < j \le f(X)$. If the vector Y results from X by setting $x_k = y_k$ for all $k = 1, ..., N$ but $k \ne i$, $k \ne j$ and $y_i = x_i - 1$ and $y_j = x_j + 1$, then $f(Y) = f(X)$.

Then Woeginger proves that an impact measure f is the g-index if and only if f satisfies E, T_1 and T_2. Axiom T_2 is only a technical one in order to obtain the characterization of the g-index. Axiom T_1 is very interesting: in view of Muirhead's theorem (discussed above) we see that T_1 (called the transfer property) is nothing else than property (8), except for the inequality sign: this is strict in (8) $(>)$ while in T_1 it is not strict $(\ge)$. Due to the requirement that impact measures should be robust with respect to a change of citation quantities we think that the non-strict inequality is a good property in this context. So the g-index (and essentially only the g-index) satisfies this econometric concentration property. What does this mean from the point of view of impact measures ? Essentially it says that, if we have two situations (e.g. two authors) with the same number of papers and the same number of citations in total, then this author, for whom the citations are more concentrated over the papers, has the highest impact. In other words it is better to write 1 paper with 100 citations (and 9 papers without a

citation) than to write 10 papers with 10 citations each. I think that it is a good property of the g-index. That this is a desired property for good impact (performance) measures is also recognized in Lehmann, Jackson and Lautrup (2008) and Leydesdorff (2008). The theorem of Woeginger, characterizing the g-index, is not trivial but it is trivial to see that the g-index satisfies axiom T_1.

Note also that the h-index does not satisfy T_1: for $X = (3,3,3)$ we have $h(X) = 3$ while for $Y = (4,3,2)$ we have $h(Y) = 2 < h(X)$.

In Egghe (2008k), these "econometric" aspects of the g-index are discussed and the Kosmulski variant $h^{(2)}$ of the h-index is also defined for the g-index. It is shown that also this variant satisfies T_1. In Egghe (2008k) we further discuss the case that $L(X) = L(Y)$: then we require that $\sum_{i=1}^{N} x_i > \sum_{i=1}^{N} y_i$ should imply that X has the highest impact, which is proved to be true for most impact measures.

The work of Gagolewski and Grzegorzewski (2009) is similar to the work of Woeginger. Several h-type indices such as the h-index, Woeginger's w-index and Kosmulski's MAXPROD index are characterized by the maximal generalized circles or ellipses that are still dominated by the citation function, i.e. the step function on unit intervals having the decreasing citation scores as ordinates.

In Deineko and Woeginger (2008) one studies a new family of impact measures which generalize the Kosmulski index and axiomatic characterizations of these measures are given. Woeginger (2008d) does the same for a family of generalizations of the g-index. Similar generalizations of the h-index and the g-index are studied in van Eck and Waltman (2008). Rousseau (2008c) checks some Woeginger axioms on the g-index, the $h^{(2)}$-index and the R-index.

Marchant (2009a) gives a characterization of some impact measures (impact in the wide sense): total number of publications, total number of citations, maximal number of citations, number of papers with at least α citations (α = a threshold) and the h-index. All these measures are ranking devices (e.g. of authors) and in Marchant (2009a) they are characterized according to their ranking properties. Most ranking properties are technical and sometimes cumbersome (as indicated by Marchant himself). It takes not less than six ranking properties to characterize the h-index but it is remarked in Rousseau (2008b) that one other important ranking property, defined in Marchant (2009a), is not a property of the h-index: the weak independence property. This ranking property is defined as follows. Consider two scientists A and B (represented by their publications and the citations received by these publications) and suppose we have a ranking device that gives a lower rank to B than to A (i.e. B is preferred to A or said otherwise, according to a certain impact measure, B has more impact than A). Assume that one adds the same publications (each with their citations) to the publication list of scientists A and B. Then B should still be preferred above A, i.e. the impact measure of B should still be larger than the one of A. This apparently logical requirement is, strangely enough, not satisfied by the impact measure h. Indeed, as argued in Rousseau (2008b), let A be represented by the vector $(5,5)$ and B by the vector $(5,3,3)$ (hence A has two publications with each 5 received citations and B has three publications with 5, 3 and 3 received citations. Hence A has h-index 2 and B has h-index 3. Suppose we add to both authors two publications with 4 received citations each. Then A and B are now represented as (still using the notation A and B): $A = (5,5,4,4)$, $B = (5,4,4,3,3)$. Now A's h-index is 4 (increase from 2 to 4) while B's h-index still is 3, now smaller than A's h-index.

Rousseau nor Marchant argue that ranking according to the h-index is a bad ranking but they underline that, whatever method of ranking one uses, one must know the properties of the ranking method. Marchant (2009a) is a good "catalogue" for that.

We add here to Rousseau's comments that also the g-index and the R-index lack this property. For the R-index we can take the example above (which is a slight modification of Rousseau's examples so that it can also be used for the R-index). In the first case we have $R_A = \sqrt{10} < R_B = \sqrt{11}$ while in the second case we have $R_A = \sqrt{18} > R_B = \sqrt{13}$.

Also for the g-index we can make an example. Let first $A = (6,2)$ $B = (5,2,2)$. We have $g_A = 2$ (even adding a zero to A does not yield a g_A-value of 3) and $g_B = 3$. Now add two articles with 4 citations each to A and B: $A = (6,4,4,2)$, $B = (5,4,4,2,2)$. Now $g_A = 4 > g_B = 3$.

Impact measures are examples of scoring rules. General scoring rules are characterized in Marchant (2009b).

This section has made it clear that it is not easy to define exactly what we mean by a "good" impact measure. Furthermore there is the intriguing problem on the influence of productivity on impact measures. Productivity can be defined as the total number of publications of an author, but what is its influence on impact measures ? In Egghe (2008i) we study impact measures of h-type (h, g, R, h_w), non-h-type (average number of citations per paper (i.e. general IF), median citation or fraction of the papers with at least a certain number of citations). General relations, such as an increasing relation between an impact measure and productivity, cannot be proved in general although, intuitively, we feel that such an increasing relation is plausible. Indications for this has been given in Egghe, Goovaerts and Kretschmer (2007) in the connection of productivity versus collaboration (in the sense of "fraction of co-authored papers"), at least in fields in the exact sciences.

The study of the general relation between productivity and impact measures has not yet been done and, hence, is formulated here as an open problem.

Lehman, Jackson and Lautrup (2008) (see also Lehmann, Jackson and Lautrup (2005)) also study impact measures of h-type and non-h-type (similar to the ones studied in Egghe (2008i)) and conclude that e.g. the average number of citations per paper is a superior indicator of scientific quality (based on statistical arguments). This is, however, disproved in Hirsch (2007).

Informetric models for h-type indices

Most models for impact measures are made in the simple framework of the law of Lotka. This law is used in two ways. We assume that the number of authors $f(n)$ with $n = 1, 2, 3, ...$ publications equals

$$f(n) = \frac{C}{n^\alpha} \qquad (9)$$

where $C > 0$, $\alpha > 1$. This is the classical law of Lotka as defined in Lotka (1926). This law is not controversial. Theoretically it applies in every IPP where sources have (or produce) items: the number of sources with n items is given by (9). A second application, used in this text is the paper-citation relation (say for an author): the number of papers with n citations (received) is given by (9). Of course, each application requires its dedicated C and α and can be obtained via a fitting procedure – see Rousseau and Rousseau (2000) or Egghe (2005), Appendix.

Applying (9) to the paper-citation relation is a bit more controversial since we can see in van Raan (2001a,b) as well as in Radicchi, Fortunato and Castellano (2008), Brantle and Fallah (2007), Lehmann, Jackson and Lautrup (2008) and Molinari and Molinari (2008a) (the latter one on the rank-frequency version of (9) – see formula (10) below) that (9) does not completely apply: for (9) to apply one needs a linearly decreasing data set in a log-log scale which is not completely true: in practise we see a concavely decreasing graph in a log-log

scale, modelled by van Raan (2001a,b) by Bessel functions - see also earlier work of Glänzel, Burrell and Sichel and the recent article Burrell (2008b) where Lotkaian informetrics is extended using a Pareto type II distribution. All these functions are difficult to calculate with and using such a model would prevent us from the creation of informetric models for impact measures.

Using Lotka's law has learned us in the past (see e.g. Egghe (2005)) that we obtain good approximations of reality and it turns out that also in this application to impact measures, this is the case.

Of course, for calculatory reasons we take $n \geq 1$ as a continuous variable. In this framework it is well-known (see Egghe (2005)) that Lotka's law (9) is equivalent with Zipf's law: in the first application, if we rank the authors in decreasing order of their number of publications, then the number of publications $g(r)$ of the author on rank r equals

$$g(r) = \frac{D}{r^{\beta}} \tag{10}$$

$D > 0$ and where β is given by

$$\beta = \frac{1}{\alpha - 1} \tag{11}$$

The notation $r \circledR g(r)$ is classical and used in Egghe (2005) and references therein. Therefore we use this notation here also but one should not confuse this function $g(.)$ with the g-index !

In the second application we rank the papers (of an author) in decreasing order of the number of received citations. Then the number $g(r)$ of citations of the paper on rank r is given by (10), where (11) is always valid.

Having the rank-frequency function $g(r)$ at our disposition (in the papers-citations context as described above), we can give an elegant variant of the definition of the h-index:

the h-index is – simply – the side of the square determined by the function $g(r)$ and its intersection with the first bissectrix. Otherwise said, if we intersect the curve $r \circledR g(r)$ in the plane with the first bissectrix, then the point of intersection has coordinates (h,h), hereby determining the Hirsch-index h (hence also the equation $g(h) = h$ defines the h-index). See Fig. 2 for a graphical illustration of this.
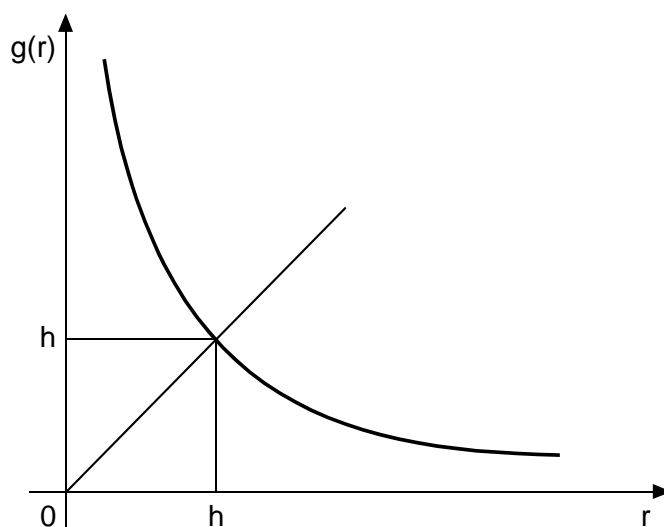

Fig. 2. Geometric illustration of the determination of the h-index

This definition already appears in Hirsch (2005a,b) but also in Valentinuzzi, Laciar and Atrio (2007), Miller (2006), Woeginger (2008b) and indirectly in Woeginger (2008a).

The exponent α is called Lotka's exponent and is very important in informetrics and it will turn out that it will play a similarly important role in the modeling of impact measures. A classical value is $\alpha = 2$ and this value is a turning point of informetric properties.

So far for this short introduction to Lotkaian informetrics; for more see Egghe (2005). The simplest model for impact measures has been obtained in Egghe and Rousseau (2006) for the h-index. Since the result is basic and its proof simple we present it with a proof. Suppose

(9) is valid for the paper-citation relation (of an author or of a journal). Let there be T cited papers in total. Then the h-index of this author or journal is given by

$$h = T^{\frac{1}{\alpha}} \tag{12}$$

Proof: By definition of f is the total number of papers given by

$$T = \int_1^\infty f(n)dn = \frac{C}{\alpha - 1} \tag{13}$$

since $\alpha > 1$. The total number of papers with n or more citations is given by

$$\int_n^\infty f(n')dn' = \frac{C}{\alpha - 1} n^{1-\alpha} \tag{14}$$

So (13) and (14) imply that the total number of papers with n or more citations is given by $Tn^{1-\alpha}$. For $n = h$ (the h-index) we need, by definition of h

$$Th^{1-\alpha} = h$$

whence (12).   □

Result (12) shows that, since $\alpha > 1$, the h-index is a concavely increasing function of T, the total number of papers. We will use (12) in the next section (in a time-dependent context) to model h-index sequences. Here we restrict ourselves to fixed time.

Glänzel (2006a) obtained an approximation of (12) via an argument with discrete values of n (see also Glänzel (2007a) and Glänzel (2008b)). The Zipf variant of (12) was also proved in Egghe and Rousseau (2006) and is reproved in Molinari and Molinari (2008a) (reference to Egghe and Rousseau (2006) is given but, surprisingly, not mentioned at the result itself) – see also Kinney (2007). In the latter publication one is not aware of formula (12) but one seems to indicate – experimentally – that (12) is valid for $\frac{1}{\alpha} \approx 0.4$ (there called the "master-curve"). In view of general experience that citation analysis is highly dependent on the field under study, we doubt that the above experimental findings are universal and in fact are not confirmed in Molinari and Molinari (2008a).

Formula (12) was experimentally verified in Egghe and Rao (2008b) on papers containing N-grams with variable length.

In Egghe and Rousseau (2006) we also found a relation between h and A = the total number of citations: if $\alpha > 2$ we have, by definition of f that

$$A = \int_1^\infty nf(n)dn = \frac{C}{\alpha - 2} \tag{15}$$

Hence (13) and (15) yield

$$A = T\frac{\alpha - 1}{\alpha - 2}$$

and hence, by (12):

$$h = \left(\frac{\alpha - 2}{\alpha - 1}A\right)^{\frac{1}{\alpha}} \tag{16}$$

, also a concavely increasing function of A.

Both formulae (12) and (16) are confirmed in experimental work of van Raan (2006): on a log-log scale, (12) and (16) are linear functions of T and A respectively, meaning: log h is a linear function of log T and log A. In van Raan (2006) one finds a linearly correlated cloud of points. Similar correlations are studied in Costas and Bordons (2007a), Bornmann and Daniel (2007a), Saad (2006), Barendse (2007), Iglesias and Pecharromán (2007), Schubert and Glänzel (2007), Bornmann, Wallon and Ledin (2008) and Ball and Ruch (2008). The concave relation (16) is also found in Minasny, Hartemink and McBratney (2007).

In Schubert and Glänzel (2007) and Glänzel (2008b) – see also Glänzel (2007a), statistical relations between the h-index and – so called – composite indicators (combining publication output and mean citation rate) are studied. Strong linear statistical relations are found. It must, however, be noted that these composite indicators have one factor equalling (in our notation) $T^{\frac{1}{\alpha}}$ which is h itself according to (12), the other factor being a power of the

mean citation rate. In Glänzel (2008b), z-statistics are proposed to be used to analyse the tail of citation distributions in the light of the h-index.

Other h-type indices can, likewise, be modelled (if $\alpha > 2$): for the g-index (Egghe (2006b))

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} \tag{17}$$

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} h \tag{18}$$

by (12); for the R-index

$$R = \sqrt{\frac{\alpha-1}{\alpha-2}} \ T^{\frac{1}{\alpha}} \tag{19}$$

$$R = \sqrt{\frac{\alpha-1}{\alpha-2}} \ h \tag{20}$$

by (12) (see Jin, Liang, Rousseau and Egghe (2007)) and similar for the weighted h-index h_w (Egghe and Rousseau (2008)). One could note that, based on (18) and (20), that g and R are linear functions of h. This is not really true: if h varies, then, by (12) and (13), it is likely that α varies and hence not only h varies in (18) and (20). Only when α is constant and when C varies in (13) we have that h varies (by (12)) and that g and R are linear in h (by (18) and (20)).

In Beirlant, Glänzel, Carbonez and Leemans (2007) an alternative for T-dependent indices is proposed: $\frac{1}{\alpha-1}$ (incidentally equal to Zipf's exponent β (see formula (11)) and is called the extreme value index. In Rousseau (2008d) one calculates (and compares) the h- and g-index in some simple distribution models.

Let us further study the simple formula (12) for the h-index: $h = T^{\frac{1}{\alpha}}$. This formula allows us to see how h will change if the publication-citation relation changes. This was done

in Egghe (2008c), using results obtained in Egghe (2007c) and Egghe (2008a). If we apply a

power law transformation to the paper rank r: $\psi(r)$ and to the citation density n: $\varphi(n)$ of the

form

$$\psi(r) = Ar^b \tag{21}$$

$$\varphi(j) = Bj^c \tag{22}$$

$(A, B, b, c > 0)$ then we prove that the h-index of the transformed system, denoted h^*, is

given by

$$h^* = B^{\frac{\delta-1}{\delta}} A^{\frac{1}{\delta}} T^{\frac{b}{\delta}} \tag{23}$$

where T is the total number of papers before the transformation and where

$$\delta = 1 + \frac{b(\alpha-1)}{c}, \tag{24}$$

the Lotka exponent of the transformed system. In Egghe (2008c) we considered simple cases:

$b = c = 1$ which gives, by (23) and (24) that

$$h^* = B^{\frac{\alpha-1}{\alpha}} A^{\frac{1}{\alpha}} h \tag{25}$$

This has the following simple application:

(i) $\psi(r) = r$, $\varphi(j) = 2j$: papers remain the same but they double their number

of citations. Then

$$h^* = 2^{\frac{\alpha-1}{\alpha}} h \tag{26}$$

which yields

$$h < h^* < 2h \tag{27}$$

: doubling of citations does not multiply h by 2, a logical fact.

(ii) $\psi(r) = 2r$, $\varphi(j) = j$: double the number of papers with the same number of

citations. Now we have

$$h^* = 2^{\frac{1}{\alpha}} h \tag{28}$$

and again we have

$$h < h^* < 2h \tag{29}$$

, again a logical fact.

(iii) $\psi(r) = 2r$, $\varphi(j) = 2j$: if we double papers and citations then we have

$$h^* = 2h \tag{30}$$

which is logical.

(iv) $\psi(r) = 2r$, $\varphi(j) = \dfrac{j}{2}$: double the number of papers but divide their number

of citations by 2. This principle occurs (approximately) in uncontrolled lists where the same paper occurs twice and where the citations to this paper are divided over the two occurrences of this paper. This also occurs (approximately) when an author has "publicitis" i.e. where an author publishes "the least publishable unit" and where, therefore, less citations per paper are obtained. Now we have

$$h^* = 2^{\frac{2}{\alpha} - 1} h \tag{31}$$

Does this practise raise the h-index ? We have from (31) that $h^* > h$ if and only if $\alpha < 2$, so for low values of α. By (12) this means high values of h, i.e. for prolific authors this practise pays off !

The next problem that we address is the one of the distribution of the h-index (given a group of authors). First we deal with the size-frequency function $\varphi(h)$, i.e. the number of authors with h-index h. As discussed in the beginning of this section we will assume that both the paper-citation relation and the author-paper relation conforms with a law of Lotka. Let $f(n) = \dfrac{C}{n^{\alpha}}$ be the number of papers with n citations and $\varphi(T) = \dfrac{C^*}{T^{\alpha^*}}$ the number of authors

with T papers. If there are T papers and since $f(n) = \dfrac{C}{n^{\alpha}}$ we have that this author has h-index

$h = T^{\frac{1}{\alpha}}$ (see formula (12)) hence $T = h^{\alpha}$. If we put this in the formula for φ we have that

$$\varphi(h) = \frac{C^*}{h^{\alpha\alpha^*}} \tag{32}$$

Hence the h-index is distributed according to Lotka's law with exponent $\alpha\alpha^*$. This was proved in Egghe (2007d). A similar result was proved for the g-index. Empirical evidence for the high degree of skewness of (32) (since the Lotka exponent is $\alpha\alpha^*$ with $\alpha > 1$ and $\alpha^* > 1$) is given in Rao (2007).

We have two applications of this result. First we can model successive h-indices. If there are S authors it follows from (32) and (12) that

$$h_2 = S^{\frac{1}{\alpha\alpha^*}} \tag{33}$$

the h-index of the group of authors (e.g. an institute). This result was obtained in Egghe (2008f). Further successive h-indices can likewise be modelled.

The second application is given in Egghe (2008l). Since we have result (32) we have that the rank-frequency distribution $h(r)$ is the law of Zipf, as remarked in the beginning of this section, hence is convexly decreasing. Since Zipf's law is a power law on a log-log scale, the function $h(r)$ is a decreasing straight line. This result was also experimentally confirmed based on the h-indices of the mathematics journals in WoS. Confirmation is also provided in Pyykkö (2008) where most $h(r)$-curves in a log-log scale are linear. This also shows that $h(r)$ has a totally different shape than $IF(r)$, the impact factor rank distribution. Based on the Central Limit Theorem, it was proved in Egghe (2008j) that $IF(r)$ has an S-shape: first convexly decreasing and then concavely decreasing, hence completely different from the

shape of $h(r)$, although Beirlant and Einmahl (2008) show the asymptotic normality of the h-index (as $T \rightarrow \infty$).

Merging aspects of the h-index are studied in diverse papers. Glänzel (2008a) discusses the merging of disjoint sets. The stochastic proof essentially boils down to the following observation. Let us have two sets of papers consisting of T_1 respectively T_2 papers. If both papers-citations systems satisfy the same law of Lotka (9) with the same α then their h-indices are (see (12)): $h_1 = T_1^{\frac{1}{\alpha}}$ and $h_2 = T_2^{\frac{1}{\alpha}}$ respectively. Hence $T_1 = h_1^{\alpha}$ and $T_2 = h_2^{\alpha}$. So $T_1 + T_2 = h_1^{\alpha} + h_2^{\alpha}$, hence $h = (T_1 + T_2)^{\frac{1}{\alpha}} = (h_1^{\alpha} + h_2^{\alpha})^{\frac{1}{\alpha}}$ is the h-index of the merged system. A similar result can be found in Molinari and Molinari (2008a). More general merging models are discussed in Egghe (2008d), where the merged sets are not necessarily disjoint. Only inequalities between the h-indices of the sets and the h-index of the merged set can be given. Two merging devices are given: in the merging one can add the citation scores or one can take the maximum of the citation scores. Examples of both scoring systems are given. In Vanclay (2007) appears an example of merging, using the max-device.

Robustness of the h-index is discussed in Rousseau (2007a) and Vanclay (2007). Rousseau (2007a) shows the robustness of the h-index with respect to missing publications and Vanclay (2007) shows robustness with respect to citation errors.

We end with some references on different topics. Basu (2007) compares the Random Hierarchical Model (see Basu (1992)) with a formula derived in Hirsch (2005b) under unexplained assumptions on the rank-order citation distribution of papers. We refer to Egghe (2008b) for the mathematical derivation of some inequalities between the h-index h (where authorship is counted in the total way) and the fractional h-indices h_f and h_F (described in the second section) and to Egghe (2008i) for some mathematical results on the influence of

productivity on some impact measures. Finally Torra and Narukawa (2007) interpret the h-index and the total number of citations as fuzzy integrals.

Time series of h-type indices

All h-type indices are single numbers − as is the case for any indicator. We can add − literally − an extra dimension by looking at h-type indices over time t. In case of e.g. an author, we can let t vary over the career of an author (or any other time span). This yields a much stronger evaluation tool than just one number and can also indicate (by extrapolation) what will be the value of this index in the (near) future.

Most of this section will be devoted to h-index sequences but most theoretical results are also valid for other impact measures (because they are the h-index multiplied with a number only dependent on Lotka's exponent α - see e.g. formulae (17)-(20)).

The h-type measures can only be calculated when we know the publication period and the citation period. So it is clear that different h-index sequences are possible. In Liu and Rousseau (2008a) not less than 10 different h-index sequences are defined, using dedicated tables indicating what publication and citation periods are used. Four different h-index sequences are studied in Egghe (2008h). One of them is the "real career h-index sequence" which, we think, is the most important h-index sequence. It could be defined as follows: for every year $t = 1, 2, ..., t_m$ (t_m is, e.g., the present (or last) year of the career) we calculate the h-index $h(t)$ based on the papers published in the years $1, 2, ..., t$ and on the citations that these papers received in the same period. Such an h-index sequence is, obviously, increasing. Further in this section we will indicate when such a sequence is concavely, linearly or convexly increasing. The real career h-index sequence was originally considered by Hirsch (2005a,b), see also Burrell (2007a).

It is best that $h(t)$ in this real career h-index sequence is calculated at every present year t (i.e. calculate $h(t)$ from year to year, say each time at the end of the year). In this way one does not have to truncate for citation periods that end in the past. Such truncations are very tedious to execute e.g. in WoS where citation data are presented up to now. Of course, since the h-index was defined in 2005, for most authors this is not possible since their careers started (long) before 2005. So e.g. in the case of this author (who's publication career started in 1978) we had to truncate citation from that year onwards. This was done in Egghe (2008g) where a linearly increasing h-index sequence was found (more on this further in this section).

This problem lead Liang (2006) to study another h-index sequence, i.e. one where one goes backwards in time: so the h-index at time t, denoted $h^*(t)$ uses papers and citations to these papers in the years $t_m - t, ..., t_m$. Note that for this h-index sequence one always has a citation period up to now, so that truncation in WoS is not necessary. Note, by the way, that Liang was the first to study h-index sequences. Burrell (2007c) noticed that Liang's h-index sequence used time going backwards. He remarks that the real career h-index sequence, as introduced above, is the more natural one, which is true. Burrell further remarks in Burrell (2007c) that, for a theoretical model, the difference between the two h-index sequences will not be large. This is, however, not true as the mathematical theory, developed in Egghe (2008g) shows. It is proved there that, for every t in the interval $[0, t_m]$ (continuous time model).

$$h^*(t) = \left(T(t_m) - h(t_m - t)^\alpha \right)^{\frac{1}{\alpha}} \qquad (34)$$

in the Lotkaian model (12) and where $T(t_m)$ is the total number of (cited) publications at time t_m (e.g. now). Based on (34) one can show that the sequences $h(t)$ and $h^*(t)$ are very different as e.g. expressed by the following result: if $h(t)$ is convex (including the linear

case) then $h^*(t)$ is strictly concave. Within the Lotkaian model it is shown that both sequences are equal if and only if the researcher has a constant production of publications per time unit.

Another h-index sequence, where time goes forward (the normal case) but where one does not have to execute tedious citation truncation actions is the so called "total career h-index sequence". Here, as in the real career h-index sequence, the h-index at time t is calculated based on publications of years $1,...,t$ but one considers, for every t, the total citation period $1,...,t_m$. Several fitting exercices for this total career h-index sequence (for authors, journals, universities and companies) were executed in Ye and Rousseau (2008).

Let us go back to the real career h-index sequence $h(t)$, which I still consider as the most important one. In Egghe (2008g) a simple mathematical model for $h(t)$, based on (12) has been given. Denote by $T(t)$ the total number of cited publications at time t. Then (12), applied to $T(t)$ gives

$$h(t) = T(t)^{\frac{1}{\alpha}} \tag{35}$$

Assume that the author has a constant number b of publications per year. Hence at time t there are $T(t) = bt$ publications yielding, by (35)

$$h(t) = b^{\frac{1}{\alpha}} t^{\frac{1}{\alpha}} \tag{36}$$

which is a concavely increasing sequence (since $\alpha > 1$). As mentioned above, this author's real career h-index sequence is linearly increasing. This can be modelled by assuming that the author produces more papers as t increases: assume that the author produces at year t, bt^{β} papers ($\beta > 0$; the case $\beta = 0$ corresponds to the above case of constant production). In Egghe (2008g) one then finds that

$$h(t) = ct^{\frac{\beta+1}{\alpha}} \hspace{4cm} (37)$$

where c is a constant. We now have that $h(t)$ is strictly concave if $\beta+1 < \alpha$, is linear if $\beta+1 = \alpha$ and is convex if $\beta+1 > \alpha$. Since this author's $h(t)$ is linear we can conclude that, for this author, $\beta+1 \approx \alpha$ (e.g. for $\alpha \approx 2$ we have $\beta \approx 1$ yielding also a linear increase of the number of publications of this author, which is confirmed in practise).

Burrell (2007a,b) studies time dependence of the h-index using a different model for the publication and citation process: in each case a Poisson process is assumed (of course, with different rates). In this model one proves in Burrell (2007a) that the h-index is approximately linear in career length and in the logarithm of the publication rate and citation rate. In Burrell (2007b) it is shown that Jin's A-index (formula (2)) is approximately a linear function of time and of h while the total number of citations to the papers in the h-core has, approximately, a square-law relationship with time and hence also with Jin's A and the h-index. The linear dependence of the h-index on career length was also confirmed in Burrell (2007c,d) (see also Burrell (2008a) for a case study) and in a simple model in Hirsch (2005b). In Burrell (2007d) it is noted that also the g-index is proportional to career length and hence to the h-index. This also follows from (18) proved in Egghe (2006b).

In da Fontoura Costa (2006), linear as well as concave h-index sequences are found for authors.

In Anderson, Hankin and Killworth (2008), a comparison between h-index sequences and h_T-index sequences (h_T = tapered h-index, described in the second section) for authors is given and they show that – essentially – there is a linear relationship between h and h_T. This needs an explanation and is left here as an open problem. The result, however, shows that the "trouble" of calculating the more difficult-to-obtain h_T-index is not really worth the effort

(but, as mentioned in the second section, there is some theoretical interest in the tapered h-index h_T).

In Egghe (2009) a general Lotkaian growth model for sources and for items is proposed. Based on this model, a time-dependent growth model for the h-index $h(t)$ and the g-index $g(t)$ is derived. Egghe (2007a,b) are special cases of the above result.

In Rousseau and Ye (2008) one defines a dynamic h-type index based on the R-index at time t multiplied by $h'(t)$, the derivative of the h-index sequence. This is a difficult to execute calculation. Using computer simulations, Guns and Rousseau (2009b) generate h-index sequences. For these simulations they use peak models (fast linear increase followed by a slow linear decrease) for the distribution of the number of received citations. They find that many simulations yield a linear h-index sequence, although convex as well as concave cases occurred as well.

Finally we mention the case studies of journal h-index sequences in Liu, Rao and Rousseau (2008) and Rousseau (2006a) and of h-index sequences of countries in Mahbuba and Rousseau (2008).

Conclusions and open problems

As is clear from this review, the h-index and its variants are extensively studied in informetrics (experimentally and theoretically), although its introduction was in Hirsch (2005b), a physicist. Outside informetrics, the h-type indices are mostly studied as case studies.

We described advantages and disadvantages of the h-index and variants of the h-index (such as the g-index and R-index) which try to avoid some disadvantages of the h-index.

The h-index, originally defined to measure the citation impact of a researcher and to predict future performance was soon applied to other sources. An important application of the h-index (or its variants) is to journals, as proposed in Braun, Glänzel and Schubert (2005, 2006). One is then immediately interested in the performance comparison between the h-index for a journal and its (classical) impact factor IF. Although some papers on this contradict each other, we are inclined to say that the h-index will become a severe concurrent for the IF of a journal. One reason is that the IF has time-thresholds both for the publication period (2 years) and for the citation period (1 year) which is not optimal in several fields; the h-index does not suppose such time limitations. Also in Scopus and WoS, one can calculate the h-index for any set of articles, hence not only for the publication set of an author but also of a journal (in a certain publication year or for all publication years).

As mentioned in the text, the application of h-type indices to topics and compounds (Banks (2006)) is very interesting (hereby determining "hot" topics).

We have formulated as a challenge to study h-type indices related to downloads (of articles in a e-journal or a repository) or even outside our field, e.g. in econometrics.

Further study is needed on "the real nature" of impact measures. Not all given "desirable properties" for impact measures, presented in the fourth section (mainly by Marchant and Woeginger) are so natural, since they are – simply – needed to characterise some h-type indices (such as the h- and g-index). More standardized properties are needed as e.g. is the case in concentration theory, using Lorenz curves.

One needs more studies on informetric models for h-type indices. We formulate here the open problem to extend the informetric models to some non-Lotkaian cases of the paper-citation relation as described in the fifth section.

Finally, as described in the last section, time series of h-type indices are very important. They add, literally, an extra dimension to this assessment tool, hence hereby describing the

evolution of a researcher's career (as one example). We noted that many different time series can be studied (different publication and citation periods but also dependent on the used source (Scopus, WoS, Google Scholar,…).

The shapes of these time series still have to be determined in many research areas and not only for authors but also for institutes, topics,… .

More and more research-funding bodies will use h-type indices (and their sequences) for assessment purposes and consequent funding (cf. Ball (2007)). More and more e-literature sources will add such indicators to their information products (Scopus, WoS, but also institutional repositories), which are, indeed, readily available for these research funding bodies such as governments, universities and other research centers e.g. in industry.

We also note that the controversy around citation analysis (which was certainly there e.g. 30 years ago) is diminishing thanks to the creation of objective simple indicators such as the h-index. This is certainly the case in the exact, applied and medical sciences ("every researcher wants to know his/her h-index"). As it is a challenge in citation analysis as a whole, it is also a challenge to apply h-type indices to the human and social sciences in an appropriate way.[2]

BIBLIOGRAPHY

Aguayo-Albasini, J. L., & Campillo-Soto, A. (2008). Evaluation of research activity by means of the Hirsch h index. *Medica Clinica*, 131(6), 239.

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

Andrews, G. E. (1998). *The theory of partitions.* Cambridge: Cambridge University Press.

---

[2] Although some references refer to the year 2009 (known to us in 2008), this review ends effectively at the end of 2008.

Antonakis, J., & Lalive, R. (2008). Quantifying scholarly impact: IQp versus the Hirsch h. *Journal of the American Society for Information Science and Technology*, 59(6), 956-969.

Arencibia-Jorge, R., Barrios-Almaguer, I., Fernández-Hernández, S., & Carvajal-Espino, R. (2008). Applying successive H indices in the institutional evaluation: a case study. *Journal of the American Society for Information Science and Technology*, 59(1), 155-157.

Arencibia-Jorge, R., & Rousseau, R. (2008). Influence of individual researchers' visibility on institutional impact: an example of Prathap's approach to successive h-indices. *Scientometrics*, to appear.

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

Ball, P. (2007). Achievement index climbs the ranks. *Nature*, 448, 737.

Ball, R., & Ruch, S. (2008). New correlations between the h-index, citation rate (CPP) and number of papers (NP) in neuroscience and quantum physics: the new S-index. In J. Gorraiz & E. Schiebel (Eds.), *Book of Abstracts. 10th International Conference on Science and Technology Indicators*, 262-263.

Banks, M. G. (2006). An extension of the Hirsch index: Indexing scientific topics and compounds. *Scientometrics*, 69(1), 161-168.

Bar-Ilan, J. (2006). H-index for Price medalists revisited. *ISSI Newsletter*, 2(1), 3-5.

Bar-Ilan, J. (2007). The h-index of h-index and of other informetric topics. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 826-827.

Bar-Ilan, J. (2008a). Which h-index? - A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257-271.

Bar-Ilan, J. (2008b). The h-index of h-index and of other informetric topics. *Scientometrics*, 75(3), 591-605.

Bar-Ilan, J. (2008c). Informetrics at the beginning of the 21st century - A review. *Journal of Informetrics*, 2(1), 1-52.

Barendse, W. (2007). The strike rate index: a new index for journal quality based on journal size and the h-index of citations. *Biomedical Digital Libraries*, 4(3) (no page numbers).

Basu, A. (1992). Hierarchical distributions and Bradford's law. *Journal of the American Society for Information Science*, 43(7), 494-500.

Basu, A. (2007). A note on the connection between the Hirsch index and the random hierarchical model. *ISSI Newsletter*, 3(2), 24-27.

Batista, P. D., Campiteli, M. G., Kinouchi, O., & Martinez, A. S. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics*, 68(1), 179-189.

Beirlant, J., & Einmahl, J. H. J. (2008). Asymptotics for the Hirsch index. *CentER Discussion Paper Series*, 2007-86.

Beirlant, J., Glänzel, W., Carbonez, A., & Leemans, H. (2007). Scoring research output using statistical quantile plotting. *Journal of Informetrics*, 1(3), 185-192.

Boell, S. K., & Wilson, C. S. (2008). Use of h-like indicators based on journal impact factors for evaluation scientific performance: Australian cancer research case study. In J. Gorraiz & E. Schiebel (Eds.), *Book of Abstracts. 10th International Conference on Science and Technology Indicators*, 253-255.

Bornmann, L., & Daniel, H.-D. (2005). Does the h-index for ranking of scientists really work? *Scientometrics*, 65(3), 391-392.

Bornmann, L., & Daniel, H.-D. (2007a). Convergent validation of peer review decisions using the h-index. Extent of and reasons for type I and type II errors. *Journal of Informetrics*, 1(3), 204-213.

Bornmann, L., & Daniel, H.-D. (2007b). What do we know about the h index? *Journal of the American Society for Information Science and Technology*, 58(9), 1381-1385.

Bornmann, L., & Daniel, H.-D. (2009). The state of h index research. Is the h index the ideal way to measure research performance ? *EMBO reports*, 10(1), 2-6.

Bornmann, L., Marx, W., & Schier, H. (2009). Hirsch-type index values for organic chemistry journals: A comparison on new metrics with the Journal Impact Factor. *European Journal of Organic Chemistry*, to appear.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2007). The b index as a measure of scientific excellence. A promising supplement to the h index. *Cybermetrics*, 11(1), Paper 6.

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. (2009). Do we need the h index and its variants besides standard bibliometric measures? *Journal of the American Society for Information Science and Technology*, to appear.

Bornmann, L., Mutz, R., Daniel, H.-D., Wallon, G., & Ledin, A. (2008). Are there really two types of h index variants? A validation study by using molecular life sciences data. In J. Gorraiz & E. Schiebel (Eds.), *Book of Abstracts. 10th International Conference on Science and Technology Indicators*, 256-258.

Bornmann, L., Wallon, G., & Ledin, A. (2008). Is the h index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the h index by using molecular life sciences data. *Research Evaluation*, 17(2),149-156.

Brahler, E., & Decker, O. (2005). The H-index. *Psycotherapie Psychosomatik Medizinische Psychologie*, 55(11), 451.

Brantle, T. F., & Fallah, M. H. (2007). Complex innovation networks, patent citations and power laws. *PICMET 2007 Proceedings*, 540-549.

Braun, T., Glänzel, W., & Schubert, A. (2005). A Hirsch-type index for journals. *The Scientist*, 19(22), 8-10.

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

Burrell, Q. L. (2007a). Hirsch's h-index: A stochastic model. *Journal of Informetrics*, 1(1), 16-25.

Burrell, Q. L. (2007b). On the h-index, the size of the Hirsch core and Jin's A-index. *Journal of Informetrics*, 1(2), 170-177.

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

Burrell, Q. L. (2007d). Hirsch's h-index and Egghe's g-index. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11ᵗʰ International Conference of the International Society for Scientometrics and Informetrics*, 162-169.

Burrell, Q. L. (2007e). Should the h-index be discounted? In W. Glänzel, A. Schubert & B. Schlemmer (Eds.), *The multidimensional world of Tibor Braun*: *a multidisciplinary encomium for his 75ᵗʰ birthday.* ISSI, 65-68;

Burrell, Q. L. (2008a). The publication/citation process at the micro level: a case study. In H. Kretschmer & F. Havemann (Eds.), *Proceedings of WIS 2008. Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting.*

Burrell, Q. L. (2008b). Extending Lotkaian informetrics. *Information Processing and Management*, 44(5), 1794-1807.

Campiteli, M. G., Batista, P. D., & Martinez, A. S. (2007). A research productivity index to account for different scientific disciplines. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11ᵗʰ International Conference of the International Society for Scientometrics and Informetrics*, 184-188.

Cater, S. C., & Kraft, D. H. (1989). A generalization and clarification of the Waller-Kraft wish list. *Information Processing and Management*, 25(1), 15-25.

Chai, J., Hua, P., Rousseau, R., & Wan, J. (2008). The adapted pure h-index. In H. Kretschmer & F. Havemann (Eds.), *Proceedings of WIS 2008. Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting.*

Costas, R., & Bordons, M. (2007a). 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. (2007b). A critical view of the h-index: observations based on a practical application. *Profesional de la Informacion*, 16, 427-432.

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

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, 57(9), 1275-1278.

Csajbók, E., Berhidi, A., Vasas, L., & Schubert, A. (2007). Hirsch-index for countries based on Essential Science Indicators data. *Scientometrics*, 73(1), 91-117.

da Fontoura Costa, L. (2006). On the dynamics of the h-index in complex networks with coexisting communities. arXiv:physics/0609116v1.

da Silva, R., Palazzo de Oliveira, J., Valdeni de Lima, J., Warpechowski, M., Hannel, K., Medianeira Denardi, R., & Moreira Orengo,V. (2009). Statistics for ranking program committees and editorial boards. Preprint.

de Araujo, A. F. P. (2008). Increasing discrepancy between absolute and effective indexes of research output in a Brazilian academic department. *Scientometrics*, 74(3), 425-437.

Deineko, V. G., & Woeginger, G. J. (2008). A new family of scientific impact measures: The generalized Kosmulski-indices. *Scientometrics*, to appear.

Dodson, M. V. (2008). Research paper citation record keeping: It is not for wimps. *Journal of Animal Science*, 86(10), 2795-2796.

Dorta-Contreras, A. J., Arencibia-Jorge, R., Marti-Lahera, Y., & Araujo-Ruiz, J. A. (2008). Productivity and visibility of Cuban neuroscientists: bibliometric study of the period 2001-2005. *Revista de Neurologia*, 47, 355-360.

Edwards, A. W. F. (2005). System to rank scientists was pedalled by Jeffreys. *Nature*, 437, 951.

Egghe, L. (2005). *Power laws in the information production process: Lotkaian informetrics.* Oxford: Elsevier.

Egghe, L. (2006a). How to improve the h-index. *The Scientist*, 20(3), 14.

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

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

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

Egghe, L. (2007b). Item-time-dependent Lotkaian informetrics and applications to the time-dependent h- and g-index. *Mathematical and Computer Modelling*, 45(7-8), 864-872.

Egghe, L. (2007c). General evolutionary theory of IPPs and applications to the evolution of networks. *Journal of Informetrics*, 1(2), 115-122.

Egghe, L. (2007d). Distributions of the h-index and the g-index. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 245-253.

Egghe, L. (2008a). The influence of transformations on the h-index and the g-index. *Journal of the American Society for Information Science and Technology*, 59(8), 1304-1312.

Egghe, L. (2008b). Mathematical theory of the h-index and g-index in case of fractional counting of authorship. *Journal of the American Society for Information Science and Technology*, 59(10), 1608-1616.

Egghe, L. (2008c). Examples of simple transformations of the h-index: Qualitative and quantitative conclusions and consequences for other indices. *Journal of Informetrics*, 2(2), 136-148.

Egghe, L. (2008d). The influence of merging on h-type indices. *Journal of Informetrics*, 2(3), 252-262.

Egghe, L. (2008e). Conjugate partitions in informetrics: Lorenz-curves, h-type indices, Ferrers graphs and Durfee squares in a discrete and continuous setting. Preprint.

Egghe, L. (2008f). Modelling successive h-indices. *Scientometrics*, 77(3), 377-387.

Egghe, L. (2008g). Mathematical study of h-index sequences. Preprint.

Egghe, L. (2008h). Comparative study of h-index sequences. *Scientometrics*, to appear.

Egghe, L. (2008i). Performance and its relation with productivity in Lotkaian systems. *Scientometrics*, to appear.

Egghe, L. (2008j). Mathematical derivation of the impact factor distribution. Preprint.

Egghe, L. (2008k). An econometric property of the g-index. Preprint.

Egghe, L. (2008l). A rationale for the Hirsch-index rank-order distribution and a comparison with the impact factor rank-order distribution. Preprint.

Egghe, L. (2009). Time-dependent Lotkaian informetrics incorporating growth of sources and items. *Mathematical and Computer Modelling*, 49(1-2), 31-37.

Egghe, L., Goovaerts, M., & Kretschmer, H. (2007). Collaboration and productivity: An investigation into "Scientometrics" journal and "Uhasselt" repository. *COLLNET Journal of Scientometrics and Information Management*, 1(2), 33-40.

Egghe, L., & Rao, I. K. R. (2008a). Study of different h-indices for groups of authors. *Journal of the American Society for Information Science and Technology*, 59(8), 1276-1281.

Egghe, L., & Rao, I. K. R. (2008b). The influence of the broadness of a query of a topic on its h-index: Models and examples of the h-index of N-grams. *Journal of the American Society for Information Science and Technology*, 59(10), 1688-1693.

Egghe, L., & Rousseau, R. (1990). *Introduction to Informetrics.* Amsterdam: Elsevier.

Egghe, L., & Rousseau, R. (1991). Transfer principles and a classification of concentration measures. *Journal of the American Society for Information Science*, 42(7), 479-489.

Egghe, L., & Rousseau, R. (1996a). Average and global impact of a set of journals. *Scientometrics*, 36(1), 97-107.

Egghe, L., & Rousseau, R. (1996b). Averaging and globalising quotients of informetric and scientometric data. *Journal of Information Science*, 22(3), 165-170.

Egghe, L., & Rousseau, R. (2006). An informetric model for the Hirsch-index. *Scientometrics*, 69(1), 121-129.

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

Engqvist, L., & Frommen, J. G. (2008). The h-index and self-citations. *Trends in Ecology & Evolution*, 23(5), 250-252.

Ferrand, L. (2007). Hirsch's h index: A new measure to quantify the research output of individual scientists. *Année Psychologique*, 107, 531-536.

Frangopol, P. T. (2005). The Hirsch index - a new scientometric indicator for the evaluation of the results of a scientific researcher. *Revista de Chimie*, 56(12), 1279-1281.

Gagolewski, M., & Grzegorzewski, P. (2009). A geometric approach to scientific impact indices construction. Preprint.

Gini, C. (1909). Il diverso accrescimento delle classi sociali e la concentrazione della richezza. *Giornale degli Economisti*, serie 11, 37.

Glänzel, W. (1996). The need for standards in bibliometric research and technology. *Scientometrics*, 35, 167-176.

Glänzel, W. (2006a). On the h-index - A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, 67(2), 315-321.

Glänzel, W. (2006b). On the opportunities and limitations of the H-index. *Science Focus*, 1(1),10-11.

Glänzel, W. (2007a). Some new applications of the h-index. *ISSI Newsletter*, 3(2), 28-31.

Glänzel, W. (2007b). The R- and AR-indices: Complementing the h-index - Comments. *Chinese Science Bulletin*, 52(6), 863.

Glänzel, W. (2008a). H-index concatenation. *Scientometrics*, 77(2), 369-372.

Glänzel, W. (2008b). On some new bibliometric applications of statistics related to the h-index. *Scientometrics*, 77(1), 187-196.

Glänzel, W. (2008c). What are your best papers? *ISSI Newsletter*, 4(4), 64-67.

Glänzel, W., & Persson, O. (2005). H-index for Price medalists. *ISSI Newsletter*, 1(4), 15-18.

Gracza, T., & Somoskovi, E. (2008). Research papers in the crosshairs: Newer viewpoints on the development of library resources. *Library Collections Acquisitions & Technical Services*, 32(1), 42-45.

Grothkopf, U., & Stevens-Rayburn, S. (2007). Introducing the h-index in telescope statistics. In S. Ricketts, C. Birdie & E. Isaksson (Eds.), *Proceedings of the 5th Library and Information Services in Astronomy Conference*, 86-92.

Guan, J. C., & Gao, X. (2009). Exploring the h-index at patent level. *Journal of the American Society for Information Science and Technology*, 60(1), 35-40.

Guns, R., & Rousseau, R. (2009a). Real and rational variants of the h-index and the g-index. *Journal of Informetrics*, 3(1), 64-71.

Guns, R., & Rousseau, R. (2009b). Simulating growth of the h-index. *Journal of the American Society for Information Science and Technology*, 60(2), 410-417.

Hardy, G., Littlewood, J. E., & Pólya, G. (1952). *Inequalities*. Cambridge: Cambridge University Press.

Harnad, S. (2007). Open access scientometrics and the UK research assessment exercise. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 27-33.

Harzing, A.-W., & van der Wal, R. (2009). A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology*, 60(1), 41-46.

Heaney, P. J. (2007). What's your h-index? *Elements*, 3(4), 229-230.

Hermes-Lima, M., Alencastro, A. C. R., Santos, N. C. F., Navas, C. A., & Beleboni, R. O. (2007). The relevance and recognition of Latin American science. Introduction to the fourth issue of CBP-Latin America. *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, 146, 1-9.

Hirsch, J. E. (2005a). An index to quantify an individual's scientific research output. arXiv:physics/0508025v3.

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

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

Honekopp, J., & Kleber, J. (2008). Sometimes the impact factor outshines the H index. *Retrovirology*, 5, 88.

Hu, X., & Chen, J. (2009). An exploration of the Major Contribution Index for leading scientists in the field of health sciences in China. Preprint.

Hua, P.-H., Rousseau, R., Sun, X.-K., & Wan, J.-K. (2009). A download $h^{(2)}$-index as a meaningful indicator of Web impact of academic journals. Preprint.

Humberto Romero, A., Garcia, A., & Kiwi, M. (2009). Evaluation of the scientific impact and biological age based upon the H-index in three Latin American countries: the material science case. Preprint.

Iglesias, J. E., & Pecharromán, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics*, 73(3), 303-320.

Imperial, J., & Rodriguez-Navarro, A. (2007). Usefulness of Hirsch's h-index to evaluate scientific research in Spain. *Scientometrics*, 71(2), 271-282.

Jacsó, P. (2007). Software issues related to cited references. *Online Information Review*, 31(6), 892-905.

Jacsó, P. (2008a). The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced databases. *Online Information Review*, 32(2), 266-283.

Jacsó, P. (2008b). The pros and cons of computing the h-index using Google Scholar. *Online Information Review,* 32(3), 437-452.

Jacsó, P. (2008c). Testing the calculation of a realistic h-index in Google Scholar, Scopus and Web of Science for F. W. Lancaster. *Library Trends*, 56(4), 784-815.

Jacsó, P. (2008d). The pros and cons of computing the h-index using Scopus. *Online Information Review*, 32(4), 524-535.

Jacsó, P. (2008e). The pros and cons of computing the h-index using Web of Science. *Online Information Review*, 32(5), 673-688.

Järvelin, K., & Persson, O. (2008). The DCI-index: discounted impact-based research evaluation. *Journal of the American Society for Information Science and Technology*, 59(9), 1433-1440. Erratum. *Journal of the American Society for Information Science and Technology*, 59(14), 2350-2351.

Jeang, K. T. (2007). Impact factor, H index, peer comparisons, and Retrovirology: is it time to individualize citation metrics? *Retrovirology*, 4(42).

Jin, B. (2006). H-index: An evaluation indicator proposed by scientist. *Science Focus*, 1(1), 8-9.

Jin, B. (2007). The AR-index: complementing the h-index. *ISSI Newsletter*, 3(1), 6.

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.

Jin, B., & Rousseau, R. (2007). R-index and AR-index: complementing indicators to h-index. *Science Focus*, 2(3), 1-8.

Jin, B., & Rousseau, R. (2008). Dust storms. *E-LIS*: ID 14592, http://eprints.rclis.org/14592/

Jones, A. W. (2008). Hirsch-index for winners of TIAFT's mid-career achievement award. *Journal of Analytical Toxicology*, 32(4), 327-328.

Kellner, A. W. A. (2008). Critical review of radiopharmaceutical-drug interactions, unexpected new age for a Tupiguarani settlement and the h-index in the Brazilian Academy of Sciences. *Anais da Academia Brasileira de Ciencias*, 80(4), 595.

Kellner, A. W. A., & Ponciano, L. C. M. O. (2008). H-index in the Brazilian Academy of Sciences - comments and concerns. *Anais da Academia Brasileira de Ciencias*, 80(4), 771-781.

Kelly, C. D., & Jennions, M. D. (2006). The h index and career assessment by numbers. *TRENDS in Ecology and Evolution*, 21(4), 167-170.

Kelly, C. D., & Jennions, M. D. (2007). H-index: age and sex make it unreliable. *Nature*, 449, 403.

Kim, W. J., & Seo, J. (2007). Evaluation of an individual's scientific productivity using author rank. *Information - An International Interdisciplinary Journal*, 10(3), 365-372.

Kinney, A. L. (2007). National scientific facilities and their impact on nonbiomedical research. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 17943-17947.

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.

Kosmulski, M. (2007). MAXPROD - A new index for assessment of the scientific output of an individual, and a comparison with the h-index. *Cybermetrics*, 11(1), Paper 5.

Krauskopf, M., & Krauskopf, E. (2008). A scientometric view of Revista Medica de Chile. *Revista Medica de Chile*, 136(8), 1065-1072.

Lehmann, S., Jackson, A. D., & Lautrup, B. E. (2005). Measures and mismeasures of scientific quality. arXiv:physics/0512238.

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

Levitt, J. M., & Thelwall, M. (2007). Two new indicators derived from the h-index for comparing citation impact: Hirsch frequencies and the normalized Hirsch index. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 876-877.

Levitt, J. M., & Thelwall, M. (2008). Is multidisciplinary research more highly cited? A macrolevel study. *Journal of the American Society for Information Science and Technology*, 59(12), 1973-1984.

Levitt, J. M., & Thelwall, M. (2009). The most highly cited Library and Information Science articles: Interdisciplinarity, first authors and citation patterns. *Scientometrics*, 78(1), 45-67.

Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometrics toolbox? *Journal of the American Society for Information Science and Technology*, to appear.

Liang, L. (2006). h-index sequence and h-index matrix: Constructions and applications. *Scientometrics*, 69(1), 153-159.

Liu, Y., Rao, I. K. R., & Rousseau, R. (2008). Empirical series of journal h-indices: the JCR category Horticulture as a case study. *Scientometrics*, to appear.

Liu, Y., & Rousseau, R. (2007). Hirsch-type indices and library management: the case of Tongji University Library. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 514-522.

Liu, Y., & Rousseau, R. (2008a). Definitions of time series in citation analysis with special attention to the h-index. *Journal of Informetrics*, 2(3), 202-210.

Liu, Y., & Rousseau, R. (2008b). Properties of Hirsch-type indices: the case of library classification categories. *Scientometrics*, to appear.

Lorenz, M. O. (1905). Methods of measuring concentration of wealth. *Journal of the American Statistical Association*, 9, 209-219.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317-324.

Lovegrove, B. G., & Johnson, S. D. (2008). Assessment of research performance in biology: how well do peer review and bibliometry correlate? *Bioscience*, 58, 160-164.

Lufrano, F., & Staiti, P. (2008). A bibliometric analysis of the international literature in supercapacitors, preprint.

Machacek, M., & Kolcunova, E. (2008). Hirsch index and rankings of Czech economists. *Politicka Ekonomie*, 56(2), 229-241.

Mahbuba, D., & Rousseau, R. (2008). Scientific research in Bangladesh and a comparison with India and Pakistan. In H. Kretschmer & F. Havemann (Eds.), *Proceedings of WIS 2008. Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting*.

Marchant, T. (2009a). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics*, to appear.

Marchant, T. (2009b). Score-based bibliometric rankings of authors. *Journal of the American Society for Information Science and Technology*, to appear.

Marx, W., & Barth, A. (2008). Carbon nanotubes - A scientometric study. *Physica Status Solidi B - Solid State Physics*, 245(10), 2347-2351.

McKercher, B. (2008). A citation analysis of tourism scholars. *Tourism Management*, 29(6), 1226-1232.

Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology*, 59(11), 1711-1726.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105-2125.

Miller, C. W. (2006). Superiority of the h-index over the impact factor for physics. arXiv:physics/0608183 v1.

Minasny, B., Hartemink, A. E., & McBratney, A. (2007). Soil science and the h index. *Scientometrics*, 73(3), 257-264.

Molinari, A., & Molinari, J.-F. (2008a). Mathematical aspects of a new criterion for ranking scientific institutions based on the h-index. *Scientometrics*, 75(2), 339-356.

Molinari, J.-F., & Molinari, A. (2008b). A new methodology for ranking scientific institutions. *Scientometrics*, 75(1), 163-174.

Mugnaini, R., Packer, A. L., & Meneghini, R. (2008). Comparison of scientists of the Brazilian Academy of Sciences and of the National Academy of Sciences of the USA on the basis of the h-index. *Brazilian Journal of Medical and Biological Research*, 41(4), 258-262.

Muirhead, R. F. (1903). Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. *Proceedings of the Edinburgh Mathematical Society*, 21, 144-157.

O'Leary, D. E. (2008). The relationship between citations and number of downloads in Decision Support Systems. *Decision Support Systems*, 45(4), 972-980.

Olden, J. D. (2007). How do ecological journals stack-up? Ranking of scientific quality according to the h index. *Ecoscience*, 14(3), 370-376.

Oppenheim, C. (2007). Using the h-index to rank influential british researchers in information science and librarianship. *Journal of the American Society for Information Science and Technology*, 58(2), 297-301.

Orbay, M., & Karamustafaoğlu, O. (2008). Chinese Journal of Chemistry's Hirsch index: A case study of 1995-2005. *Chinese Librarianship: an International Electronic Journal*, 24 (1/12/2007).

Pilc, A. (2008). The use of citation indicators to identify and support high-quality research in Poland. *Archivum Immunologiae et Therapiae Experimentalis*, 56(6), 381-384.

Pires Da Luz, M., Marques-Portella, C., Mendlowicz, M., Gleiser, S., Silva Freire Coutinho, E., & Figueira, I. (2008). Institutional h-index: The performance of a new metric in the evaluation of Brazilian psychiatric post-graduation programs. *Scientometrics*, 77(2), 361-368.

Popov, S. B. (2005). A parameter to quantify dynamics of a researcher's scientific activity. arXiv:physics/0508113v1.

Prathap, G. (2006). Hirsch-type indices for ranking institutions' scientific research output. *Current Science*, 91(11), 1439.

Purvis, A. (2006). The h index: playing the numbers game. *TRENDS in Ecology and Evolution*, 21(8), 422.

Pyykkö, P. (2008). Power-law distribution of individual Hirsch indices, the comparison of merits in different fields, and the relation to a Pareto distribution. arXiv:physics/0608282 v1.

Qiu, J., Ma, R., & Cheng, N. (2008). New exploratory work of evaluating a researcher's output. *Scientometrics*, 77(2), 335-344.

Quesada, A. (2009). Monotonicity and the Hirsch index. *Journal of Informetrics*, to appear.

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Towards an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45), 17268-17272.

Rao, I. K. R. (2007). Distributions of Hirsch-index and g-index: An empirical study. In D. Torres-Salinas and H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 655-658.

Rau, J. R. (2007). h-index (2000-2004) of the most cited environmental researchers based at Chilean institutions. *Revista Chilena de Historia Natural*, 80, 381-383.

Reedijk, J., & Moed, H. F. (2008). Is the impact of journal impact factors decreasing? *Journal of Documentation*, 64(2), 183-192.

Riikonen, P., & Vihinen, M. (2008). National research contributions: A case study on Finnish biomedical research. *Scientometrics*, 77(2), 207-222.

Rodriguez, V., Janssens, F., Debackere, K., & De Moor, B. (2007). On material transfer agreements and visibility of researchers in biotechnology. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 659-671.

Rodriguez, V., Janssens, F., Debackere, K., & De Moor, B. (2008). On material transfer agreements and visibility of researchers in biotechnology. *Journal of Informetrics*, 2(1), 89-100.

Romero, A. G., Cortes, J. N., Escudero, C., Lopez, J. A. F., & Moreno, J. A. C. (2007). Measuring the contribution of clinical trials to bibliometric indicators: Citations and journal impact factor. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics*, 300-304.

Rosenstreich, D., & Wooliscroft, B. (2009). An improved method for ranking management journals using the g-index and Google Scholar. Preprint.

Rousseau, B., & Rousseau, R. (2000). LOTKA: a program to fit a power law distribution to observed frequency data. *Cybermetrics*, 4(1), paper 4.

Rousseau, R. (2002). Lack of standardisation in informetric research. Comments on "Power laws of research output. Evidence for journals of economics" by Matthias Sutter and Martin G. Kocher. *Scientometrics*, 55(2), 317-327.

Rousseau, R. (2006a). A case study: evolution of JASIS' h-index. *Science Focus*, 1(1), 16-17.

Rousseau, R. (2006b). New developments related to the Hirsch index. *Science Focus*, 1(4), 23-25.

Rousseau, R. (2006c). Simple models and the corresponding h- and g-index. *E-LIS*: ID 6153, http://eprints.rclis.org/archive/00006153/

Rousseau, R. (2007a). The influence of missing publications on the Hirsch index. *Journal of Informetrics*, 1(1), 2-7.

Rousseau, R. (2007b). Hungary - and Tibor Braun - on top. In W. Glänzel, A. Schubert & B. Schlemmer (Eds.), *The multidimensional world of Tibor Braun: a multidisciplinary encomium for his 75th birthday*. ISSI, 25-28.

Rousseau, R. (2008a). Reflections on recent developments of the h-index and h-type indices. In H. Kretschmer & F. Havemann (Eds.), *Proceedings of WIS 2008. COLLNET Journal of Scientometrics and Information Management*, 2(1), 1-8.

Rousseau, R. (2008b). Luckily, science focuses on achievements. Some thoughts related to the h-index. *ISSI Newsletter*, 4(3), 49-50.

Rousseau, R. (2008c). Woeginger's axiomatization of the h-index and its relation to the g-index, the h$^{(2)}$-index and the R^2-index. *Journal of Informetrics*, 2(4), 335-340.

Rousseau, R. (2008d). Simple models and the corresponding h- and g-index. Preprint.

Rousseau, R., Guns, R., & Liu, Y. (2008). The h-index of a conglomerate. *Cybermetrics*, 12(1), Paper 2.

Rousseau, R., & Jin, B. (2008). The age-dependent h-type AR^2-index: Basic properties and a case study. *Journal of the American Society for Information Science and Technology*, 59(14), 2305-2311.

Rousseau, R., & Rons, N. (2008). Another h-type index for institutional evaluation. *Current Science*, 95(9), 1103.

Rousseau, R., & Small, H. (2005). Escher staircases dwarfed. *ISSI Newsletter*, 1(4), 8-10.

Rousseau, R., & Ye, F. Y. (2008). A proposal for a dynamic h-type index. *Journal of the American Society for Information Science and Technology*, 59(11), 1853-1855.

Ruane, F., & Tol, R. S. J. (2008). Rational (successive) h-indices: An application to economics in the Republic of Ireland. *Scientometrics*, 75(2), 395-405.

Saad, G. (2006). Exploring the h-index at the author and journal levels using bibliometric data of productive consumer scholars and business-related journals respectively. *Scientometrics*, 69(1), 117-120.

Salgado, J. F., & Paez, D. (2007). Scientific productivity and Hirsch's h index of Spanish social psychology: Convergence between productivity indexes and comparison with other areas. *Psicothema*, 19(2), 179-189.

Sanderson, M. (2008). Revisiting h measured on UK LIS and IR academics. *Journal of the American Society for Information Science and Technology*, 59(7), 1184-1190.

Sangam, S. L., & Girji, R. M. (2008). Hirsch index: a new measure for assessing scientific productivity of an individual researcher. *Current Science*, 94, 291.

Satyanarayana, K., & Sharma, A. (2008). Impact factor: Time to move on. *Indian Journal of Medical Research*, 127(1), 4-6.

Schreiber, M. (2007a). Self-citation corrections for the Hirsch index. *Europhysics Letters*, 78(3).

Schreiber, M. (2007b). A case study of the Hirsch index for 26 non-prominent physicists. *Annalen der Physik*, 16, 640-652.

Schreiber, M.. (2008a). A modification of the h-index: The h_m-index accounts for multi-authored manuscripts. *Journal of Informetrics*, 2(3), 211-216.

Schreiber, M. (2008b). To share the fame in a fair way, h(m) modifies h for multi-authored manuscripts. *New Journal of Physics*, 10.

Schreiber, M. (2008c). The influence of self-citation corrections on Egghe's g index. *Scientometrics*, 76(1), 187-200.

Schreiber, M. (2008d). An empirical investigation of the g-index for 26 physicists in comparison with the h-index, the A-index, and the R-index. *Journal of the American Society for Information Science and Technology*, 59(9), 1513-1522.

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

Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, 1(3), 179-184.

Sebire, N. J. (2008). H-index and impact factors: assessing the clinical impact of researchers and specialist journals. *Ultrasound in Obstetrics & Gynecology*, 32(7), 843-845.

Sezer, E., & Gokceoglu, C. (2009). A fuzzy approach for the assessment of academic performance. Preprint.

Sidiropoulos, A., & Katsaros, D. (2008). Unfolding the full potentials of H-index for bibliographic ranking. Preprint.

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*, 72(2), 253-280.

Small, H. (2007). Personal communication.

Smith, L. (1981). Citation analysis. *Library Trends*, 30, 83-106.

The STIMULATE6 Group. (2007). The Hirsch index applied to topics of interest to developing countries. *First Monday*, 12(2).

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*, 34(4), 605-621.

Timofieiev, A., Snasel, V., & Dvorsky, J. (2008). h-index calculation in Enron corpus. In J. Cordeiro, B. Shishkov, A. Ranchordas & M. Helfert (Eds.), *ICSOFT 2008: Proceedings of the third International Conference on Software and Data Technologies*, Vol. ISDM/ABF, 206-211.

Tol, R. S. J. (2008). A rational, successive g-index applied to economics departments in Ireland. *Journal of Informetrics*, 2(2), 149-155.

Torra, V., & Narukawa, Y. (2008). The h-index and the number of citations: two fuzzy integrals. *IEEE Transactions on Fuzzy Systems*, 16(3), 795-797.

Torro-Alves, N., Herculano, R. D., Tercariol, C. A. S., Filho, O. K., & Graeff, C. F. O. (2007). Hirsch's index: a case study conducted at the Faculdade de Filosofia, Ciencias e letras de Ribeirao Preto, Universidade de Sao Paulo. *Brazilian Journal of Medical and Biological Research*, 40, 1529-1536.

Ursprung, H. W., & Zimmer, M. (2007). Who is the "Platz-Hirsch" of the German economics profession? A citation analysis. *Jahrbücher für Nationalökonomie und Statistik*, 227(2), 187-208.

Valentinuzzi, M. E., Laciar, E., & Atrio, J. L. (2007). Two new discipline-independent indices to quantify individual's scientific research output. *Journal of Physics: Conference Series*, 90, 012018.

van Eck, N. J., & Waltman, L. (2008). Generalizing the h- and g-indices. *Journal of Informetrics*, 2(4), 263-271.

van Haselen, R. (2007). The h-index: A new way of assessing the scientific impact of individual CAM authors. *Complementary Therapies in Medecine*, 15, 225-227.

van Leeuwen, T. (2008). Testing the velidity of the Hirsch-index for research assessment purposes. *Research Evaluation*, 17(2), 157-160.

van Raan, A. F. J. (2001a). Two-step competition process leads to quasi power-law income distributions. Application to scientific publication and citation distributions. *Physica A*, 298, 530-536.

van Raan, A. F. J. (2001b). Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. *Scientometrics*, 51(1), 347-357.

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

Vanclay, J. K. (2006). Refining the h-index. *The Scientist*, 20(7), 14-15.

Vanclay, J. K. (2007). On the robustness of the h-index. *Journal of the American Society for Information Science and Technology*, 58(10), 1547-1550.

Vanclay, J. K. (2008a). Gauging the impact of journals. *Forest Ecology and Management*, 256(4), 507-509.

Vanclay, J. K. (2008b). Ranking forestry journals using the h-index. *Journal of Informetrics*, 2(4), 326-334.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics*, 10, 157-177.

Vinkler, P. (2007). Eminence of scientists in the light of the h-index and other scientometric indicators. *Journal of Information Science*, 33(4), 481-491.

Wan, J.-K., Hua, P.-H., Du, J., & Song, H.-M. (2007). Focus on the front of scientific evaluation, putting new indicators into practise: Three new bibliometric indicators adopted by China Academic Journals Comprehensive Citation Report. *Digital Library Forum*, 3, 36-41.

Wan, J., Hua, P., & Rousseau, R. (2007). 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), 1-5.

Wan, J.-K., Hua, P.-H., Rousseau, R., & Sun, X.-K. (2009). The downloaded immediacy index (DII): Experiences using the CNKI full-text database. Preprint.

Wendl, M. C. (2007). H-index: however ranked, citations need context. *Nature* 449(7161), 403.

Woeginger, G. J. (2008a). An axiomatic characterization of the Hirsch-index. *Mathematical Social Sciences*, 56(2), 224-232.

Woeginger, G. J. (2008b). A symmetry axiom for scientific impact indices. *Journal of Informetrics*, 2(4), 298-303.

Woeginger, G. J. (2008c). An axiomatic analysis of Egghe's g-index. *Journal of Informetrics*, 2(4), 364-368.

Woeginger, G. J. (2008d). Generalizations of Egghe's g-index. *Journal of the American Society for Information Science and Technology*, to appear.

Ye, F. Y., & Rousseau, R. (2008). The power law model and total career h-index sequences. *Journal of Informetrics*, 2(4), 288-297.

Zhivotovsky, L. A., & Krutovsky, K. V. (2008). Self-citation can inflate h-index. *Scientometrics*, 77(2), 373-375.