European Journal of Operational Research 203 (2010) 494–504

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Analysis of the Hirsch index's operational properties

Fiorenzo Franceschini *, Domenico A. Maisano

Politecnico di Torino, Dipartimento di Sistemi di Produzione ed Economia dell'Azienda (DISPEA), Corso Duca degli Abruzzi 24, 10129 Torino, Italy

ARTICLE INFO

Article history:
Received 4 November 2008
Accepted 4 August 2009
Available online 9 August 2009

Keywords:
Hirsch index
Performance evaluation
Citations
Ranking
Indicator properties
Operational properties
Publication year

ABSTRACT

The *h*-index is a relatively recent bibliometric indicator for assessing the research output of scientists, based on the publications and the corresponding citations. Due to the original characteristics of easy calculation and immediate intuitive meaning, this indicator has become very popular in the scientific community. Also, it received some criticism essentially because of its "low" accuracy. The contribution of this paper is to provide a detailed analysis of the *h*-index, from the point of view of the indicator operational properties. This work can be helpful to better understand the peculiarities and limits of *h* and avoid its misuse. Finally, we suggest an additional indicator (f) that complements *h* with the information related to the publication age, not compromising the original simplicity and immediacy of understanding.

1. Introduction

The evaluation of the scientific work of a scientist is a more and more interesting problem for the scientific community. In 2005, Hirsch proposed a synthetic indicator (h) for evaluating and comparing the research activity of individual scientists according to their output. Specifically, *h* is defined as the number such that, for a general group of papers, *h* papers received at least *h* citations while the other papers received no more than *h* citations (Hirsch, 2005). *h* can be calculated very quickly using some data which are available in specialized databases (such as Scopus or Web of Science – WoS) or public databases (such as Google Scholar – GS). In this way *h* synthetically aggregates two important aspects of the scientific output: respectively diffusion/impact – represented by the number of citations per paper – and productivity – represented by the number of different papers. As *h* grows both with the publications number and the citations number, it encourages the production of good scientific works.

Ever since its introduction, this indicator has received much attention. Two possible reasons are:

- the problem of the evaluation of the scientists' output is very critical, for example for academics or scientists seeking promotion, tenure, faculty positions, research grants etc. (Harzing and van der Wal, 2008; Korhonen et al., 2001);
- the natural simplicity, the immediate understanding and the apparent effectiveness of *h* encourage its diffusion. In fact, *h* has an immediate intuitive meaning and is directly associable to real data (the number of citations and publications), not requiring any conversion, scale transformation or data processing. Probably, this is a fundamental reason why it is so popular and largely diffused.

A tangible sign of the popularity of *h* is the appearance of many publications about reflections, analyses, proposals for new variants and improvements.

h takes account of the history of one author's scientific production and is able to summarize it into a single number. At the same time, this can be a limitation, because it discards many details of citation record (Joint Committee on Quantitative Assessment of Research, 2008). Hirsch's proposal generated a deep debate among scientists, not only about the indicator's validity and applicability, but also about pros and cons of the use of citation analysis in research evaluation (Moed, 2005; Kelly and Jennions, 2006; Rousseau, 2006; Mingers, 2008). Here follows a list of the most interesting discussions on this topic. Many of these points are generally true for any indicator in citation analysis.

* Corresponding author. Tel.: +39 011 5647225; fax: +39 011 5647299.
E-mail address: fiorenzo.franceschini@polito.it (F. Franceschini).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2009.08.001

- *h* does not take into account multiple co-authorship (Burrell, 2007a). A possible correction consists in penalising publications with multiple authors. For example, Schreiber (2008) suggests to fractionalise the number of publications by counting each paper only according to the inverse of the number of authors. Of course, this approach is carried out assuming an equal contribution of all co-authors.
- *h* does not take into account self-citations. Their influence is not so small, especially for younger scientists with low Hirsch index (Schreiber, 2007; Burrell, 2007a).
- *h* is not useful for cross-disciplinary comparisons because citation rates and scholarly productivity vary considerably among disciplines (e.g. physics, medicine, engineering) (Antonakis and Lalive, 2008; Batista et al., 2006; Braun et al., 2006).
- *h* does not take into account the age of publications. A possible correction to *h*-index is to add an age-related weighting to each cited article, giving less weight to older articles, depending on the weighting parameterization (Sidiropoulos et al., 2007).
- *h* does not consider the publication type. For example review articles, open access articles, papers addressing "hot topics" or in fields shared by large communities will often receive far more citations than other papers, all other things being equal (Castillo et al., 2007).
- the *h*-index for a scientist can be easily calculated by using public databases like WoS or GS (Meho and Yang, 2007). Unfortunately, their information can be affected by citation errors – for instance caused by homonymous author names, typographical errors in the source papers, or errors due to some nonstandard reference formats (Bornmann and Daniel, 2007; Harzing and van der Wal, 2008).

As briefly shown, most of these debates suggest new methods aimed at enhancing the *h*-index synthetic information. Nevertheless, they generally undermine its characteristics of easy calculation and immediate intuitive meaning.

The purpose of this paper is providing a detailed analysis of the *h* operational properties, which are a consequence of its definition and construction. This work can be helpful to better understand the peculiarities and limits of *h*, avoiding its misuse. The paper is organised into three sections. Section 1 contains some reflections on the features of *h* and its aggregation criterion. Section 2 is the core of the paper. Firstly, it explains the assumptions and the simplifying hypotheses behind the use of *h*; then identifies and discusses in detail the *h* operational properties. Section 3 suggests a simple procedure to complement *h* with the information related to the publication age, not compromising the simplicity and immediacy of the original indicator. Finally, the conclusions are given, summarising the original contribution of this paper.

2. Reflections on the *h* structure

According to its definition, *h* can be determined:

- knowing the number of citations received by the (most cited) published papers;
- sorting the papers in descending order with respect to the citation number;
- finding the number *h* such that *h* papers received at least *h* citations, while the other papers received no more than *h* citations. The set of the *h* most cited papers is called *h*-core (see the example in Fig. 1).

This procedure can be easily implemented by using the information from databases that are available online. *h* provides a snapshot of a scholar's most important papers and does not require complex calculations or algorithms. This is not so common among indicators. It is worth reminding that an indicator is successful not only if it is effective, but also if it is easily understood. Indicators that are difficult to understand and interpret, because reference to real data has been "lost", are often rejected by potential users (Montgomery, 2005). For this reason, the *h*-index popularity can be seen as an indirect confirmation of its simplicity and immediate intuitive meaning.

As noticed before, *h* considers two features of a scholar's scientific production:

1. the number of papers (assessment of productivity);
2. the citation rate of the different papers (assessment of diffusion/impact).

These two features are aggregated using a particular and original synthesis criterion, consisting in finding the number *h* such that *h* papers received at least *h* citations. This synthesis makes sense because the two examined parameters have roughly the same order of magnitude, so – despite their different nature – they can be compared (see Figs. 2a and 3). For example, if the number of citations per paper were of the order of magnitude of thousands and the number of papers were generally not larger than few units, then *h* would be very small (limited by the second parameter). Thus, in this condition, *h* would give a rough indication of a scientist's productivity but it would not be useful for representing the corresponding diffusion/impact. A similar synthesis criterion is used by *g*-index. *g* is the most popular variant of

	citations for each paper	rank
h-core	30	1
h-core	20	2
h-core	18	3
h-core	12	4
h-core	9	5
h-core	8	6
h-core	8	**7**
	6	8
	6	9
	5	10
	…	…

Fig. 1. Example of *h*-index calculation. In this case $h = 7:7$ papers received at least 7 citations each. The set of the *h* most cited papers can be identified as *h*-core.

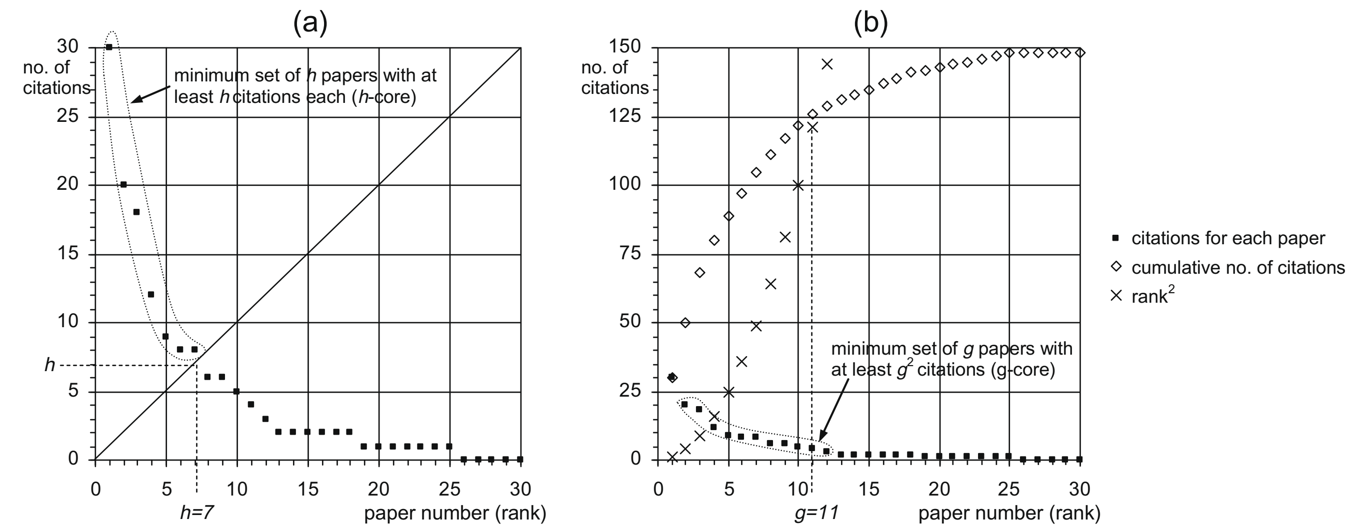

Fig. 2. Graphical representation of (a) the *h*-index and (b) the *g*-index. Graphs are constructed using the input data reported in Fig. 3.

	citations for each paper	rank	cumulative no. of citations	rank^2	
	30	1	30	1	
	20	2	50	4	
	18	3	68	9	
h-core	12	4	80	16	
	9	5	89	25	
	8	6	97	36	*g*-core
	8	**7**	105	49	
	6	8	111	64	
	6	9	117	81	
	5	10	122	100	
	4	11	126	121	
	3	**12**	129	144	
	2	13	131	169	
	2	14	133	196	
	2	15	135	225	
	2	16	137	256	
	2	17	139	289	
	2	18	141	324	
	1	19	142	361	
	1	20	143	400	
	1	21	144	441	
	1	22	145	484	
	1	23	146	529	
	1	24	147	576	
	1	25	148	625	
	0	26	148	676	
	0	27	148	729	
	0	28	148	784	
	0	29	148	841	
	0	30	148	900	

Fig. 3. Example of calculation of the *h*-index and *g*-index using the same (fictitious) input data.

h, defined as the largest number for which the *g* most cited papers have a total of at least g^2 citations (Egghe, 2006). Similarly to *h*, *g* is based on a comparison between number of papers and citation rate of the papers, and corresponds to the size of a particular subset of the most cited papers (see Figs. 2b and 3).

It can be noticed that *g* is more sensitive to one or several highly cited papers than *h*.

From an operating point of view, *h* acts as a "double filter" for the following reasons:

1. once a paper is selected to belong to the *h*-core, this paper is not "used" any more in the determination of *h*, as a variable over time. Indeed, the *h*-index calculated in subsequent years is not at all influenced by this paper's received citations further on (i.e. even if the paper doubles or triples its number of citations, the subsequent *h*-indexes is not influenced);
2. *h* does not take into account the citation rate of the "tail" papers, that is to say those papers with a low number of citations.

Being insensitive to accidental excess of low cited articles, and to one or several highly cited articles, *h* is commonly considered a robust indicator (Orbay et al., 2007). The other side of the coin is that *h* is considered to be rather coarse in evaluating the scholars' scientific output.

As there has yet been no thorough validation of the h-index, that is cross-discipline validation on the basis of broad statistical data, the h-index should be carefully used as a criterion to support decision making in science. The same holds, of course, for the many complementary indices, modifications and alternatives suggested in the literature (Seglen, 1992; Lehmann et al., 2005; Garfield, 2006; Lehmann et al., 2006; Franceschini and Maisano, in press).

3. Operational properties and assumptions about the *h*-index

Behind the use of *h* there are many implicit assumptions and simplifying hypotheses presented in Table 1. Each of them can be questionable and raised many debates and reflections about the related pros and cons (Castillo et al., 2007; Schreiber, 2007; Wendl, 2007; Hirsch, 2007).

Several variants of the *h*-index have been devised in order to reduce its major limitations (Banks, 2006; Burrell, 2007b; Glänzel, 2006; Katsaros et al., 2007; Saad, 2006). However, they are often naïve attempts to capture a complex citation record with a single number (Joint Committee on Quantitative Assessment of Research, 2008). In addition, most of the indicators derived by *h* do not satisfy the two original requirements of simple calculation and immediate intuitive meaning.

As shown in the previous examples (Figs. 1 and 3), *h* can be interpreted as an indicator able to summarize the information contained in a string of numbers (the citations of each paper, sorted in descending order).

According to the definition presented by Franceschini et al. (2007), an indicator can be classified in two categories:

- *basic* if it is obtained as a direct observation of an empirical system. Examples of basic indicators are the number of manufactured parts or the number of defectives in a production line;
- *derived* if it is obtained combining the information of one or more sub-indicators (basic or derived). An example of derived indicators is the defectiveness of a production line (ratio between defectives and good products for a given time).

Thus, *h* is a derived indicator obtained combining the number of citations and the number of the different papers of a scholar (which are both direct empirical observations). From an operational point of view, derived indicators can satisfy or not some properties (Franceschini et al., 2006; Franceschini et al., 2008). Regarding the *h*-index, the most important properties are summarized in Table 2. They are individually analysed and discussed in the following subsections.

3.1. Papers' passage in the h-core

The articles included in the *h*-core are not fixed during the evolution of a scholar's scientific production. In particular:

- if a paper which is out of the *h*-core earns a number of citations larger than *h*, it will enter the *h*-core not necessarily leading to an increase in the *h* value (see the example of Fig. 4a);
- when a new paper enters the *h*-core it may cause another paper to go out of it. Specifically, the paper or papers that at any given time have exactly *h* citations are at risk of being eliminated from the *h*-core as they are superseded by other papers that are being cited at a higher rate (see Fig. 4b). It is also possible that papers "drop out" and then later come back into the *h*-core (Hirsch, 2005).

3.2. No decreasing

It is always true that *h* cannot decrease with time. This property simply depends on the fact that *h* aggregates the number of papers and the corresponding number of citations, and both these variables can not decrease over time. For example, in case of career interruption or retirement, the *h*-index remains constant or may increase (if already published papers accumulate new citations). The negative consequence of this fact is that *h* is not perfectly suitable to compare scholars with different seniority, being in favour of those with long careers.

3.3. Limited max value of h

For each author, the maximum value of *h* is limited by the following constraints:

- the number of citations of the most cited papers;
- the number of papers.

Table 1
Assumptions and simplifying hypotheses behind the definition and use of *h*.

Assumptions and simplifying hypotheses	
1	The diffusion/impact of an article is evaluated using the number of citations received
2	Self-citations do not increase the *h*-index significantly
3	It is not essential to take the effect of multiple co-authorship into account
4	Citations have the same importance, no matter what their age is or what the paper age is
5	It is not necessary to distinguish between publications of different type (e.g. review, conference, journal articles)
6	As citation rates and scholarly productivity vary considerably among the different research fields, *h* does not allow comparisons between scholars from different fields (Batista et al., 2006; Braun et al., 2006)
7	The information used to determine the *h*-index – provided by public databases like WoS, GS or Scopus – is considered reliable, with no significant errors. In actual fact, using WoS, GS or Scopus can lead to different results since they use different sets of source journals and have different database limitations (Harzing and van der Wal, 2008)

Table 2
Summary of the operational properties of h.

Property	Description
Papers' passage in the h-core	The h-core papers are not fixed Each paper can enter or go out of it depending on the citations earned over time
No decreasing	A scientist's h-index will never decrease
Limited max value of h	The maximum value of h is limited by the number of papers and the corresponding number of citations
Non compensation	Changes in the number of citation of different papers (sub-indicators) do not necessarily compensate each other, not making the value of h (derived indicator) change
Non (strict) monotony	The increase/decrease in number of citations of a paper (sub-indicator) is not necessarily associated to a corresponding increase/decrease of h (derived indicator)
h-Index scale properties	h defines an ordinal scale

(a)

Condition 1 (h=3)			Condition 2 (h=3)		
paper id	citations	rank	paper id	citations	rank
A	15	1	A	15	1
B	9	2	B	9	2
C	3	**3**	D	5	**3**
D	2	4	C	3	4
E	1	5	E	1	5
F	1	6	F	1	6
G	0	7	G	0	7

(b)

Condition 1 (h=3)			Condition 2 (h=4)		
paper id	citations	rank	paper id	citations	rank
A	7	1	A	7	1
B	5	2	B	5	2
C	3	**3**	D	4	3
D	2	4	E	4	**4**
E	1	5	C	3	5
F	1	6	F	1	6
G	0	7	G	0	7

Fig. 4. Examples of a paper entering (a) and going out of the h-core (b). Case (a): paper D increases the citations number (from 2 to 5), entering the h-core (in grey), but not making h increase. Case (b): paper D and E increase their citations number, entering the h-core (in grey) and making h increase from 3 to 4. At the same time, article C goes out of the h-core.

Generally, the first constraint is the most restrictive. However, in some particular conditions where the number of papers is relatively small but they are highly cited, the second constraint can be the most restrictive. This situation can be unfavourable to young brilliant scientists who have few highly diffused articles (Sidiropoulos et al., 2007). For instance, considering a scientist that has only 10 papers with 30 citations each, his h will be limited to 10.

3.4. Non compensation

The property of compensation can be studied when a system is represented by sub-indicators (citations of each paper) aggregated by a derived indicator (h-index). If local changes (i.e. increase/decrease in performance) of sub-indicators may compensate each other – without making the derived indicator value change – then the derived indicator fulfils the property of compensation. In formal terms, a derived indicator (h) fulfils the property of compensation if the following situation is verified (Franceschini et al., 2007):

<u>If</u> $h(C_1) = h(C_2)$
<u>and if</u> $\exists\, I_i \in F : I_i(C_1) > I_i(C_2)$ (increase in the local performance of i-th sub-indicator)
<u>then</u> $\exists$ at least one indicator $I_j(i \neq j) \in F : I_j(C_1) < I_j(C_2)$ (decrease in the local performance of j-th sub-indicator)

Condition 1 (h=3)			Condition 2 (h=4)		
paper id	citations	rank	paper id	citations	rank
A	10	1	A	9 (-1)	1
B	9	2	B	5	2
C	5	**3**	C	5	3
D	3	4	D	4 (+1)	**4**
E	2	5	E	2	5
F	1	6	F	1	6

Fig. 5. h does not fulfil the compensation property. Two conditions are exemplified. In condition 2, paper A has one citation less than in condition 1. Nevertheless, because of the contribution of paper D (one citation more), h is larger in condition 2. h passes from 3 to 4. Compensation is not fulfilled because, in condition 2, the higher citation rate of paper D is not counterbalanced by the lower citation rates of paper A.

being:
F set of sub-indicators (citations of each paper);
I_i and I_j sub-indicators that change in value;
h derived indicator (h-index);
C_1 and C_2 two examined conditions.

h does not follow the property of compensation, as exemplified in Fig. 5. Generally, compensation is a typical property of additive and multiplicative aggregation models.

3.5. Non (strict) monotony

If a system is represented by different sub-indicators (citations of each paper) aggregated into a derived indicator (h-index), and if the increase/decrease of one sub-indicator is not associated to the increase/decrease of the derived indicator, then the derived indicator does not fulfill the condition of (strict) monotony (Franceschini et al., 2007).

In more detailed terms, if the system (scholar's scientific production in this specific case) skips from condition C_1 to condition C_2 increasing/decreasing the performance of one sub-indicator (I_k), then h should increase/decrease too. Otherwise, h is not (strictly) monotonous. In formal terms:

$$\begin{aligned}&\underline{\text{If}}\quad \forall\, I_j \in F \setminus \{I_i\}, I_j(C_1) = I_j(C_2)\\ &\underline{\text{and if}}\quad \exists I_i \in F : I_i(C_1) > I_i(C_2) \Rightarrow h(C_1) > h(C_2)\\ &\underline{\text{then}}\quad \text{the derived indicator } h \text{ is (strictly) monotonous}\end{aligned}$$

being:
F set of sub-indicators (citations of each paper);
I_k sub-indicator that increases;
$F \setminus \{I_k\}$ set of sub-indicators, not including I_k ;
h derived indicator (h-index);
C_1 and C_2 two examined conditions.

Taking into consideration the exemplified data in Fig. 6, the number of citations of paper C goes from 4 to 9, passing from condition 1 to condition 2. This latter condition is surely better compared to the previous one, and the h value would be expected to be higher than the previous one. Nevertheless, h remains unchanged not representing a significant improvement. This proves that h is not strictly monotonic, but only weakly monotonic. In fact:

$$\begin{aligned}&\underline{\text{If}}\quad \forall\, I_j \in F \setminus \{I_i\}, I_j(C_1) = I_j(C_2)\\ &\underline{\text{and if}}\quad \exists I_i \in F :\ I_i(C_1) > I_i(C_2) \Rightarrow h(C_1) \geqslant h(C_2)\\ &\underline{\text{then}}\quad \text{the derived indicator } h \text{ is weakly monotonous}\end{aligned}$$

3.6. h-Index scale properties

Analyzing the h-index empirical system – that is the string of numbers of each paper's citations, sorted in descending order – it can be noticed that the h scale has only equivalence and ordinal properties.

- *Equivalence property*. Each string can be associated to one and only one h class. Strings with the same h class are considered equivalent (see Fig. 7).The h classes of equivalence are mutually exclusive.
- *Ordinal property*. h defines an ordinal relation between strings. The larger h is, the more important the string. Each comparison between two h values results in establishing an ordinal relation between the corresponding strings.

Condition 1 (h=3)

paper id	citations	rank
A	10	1
B	9	2
C	4	**3**
D	3	4

Condition 2 (h=3)

paper id	citations	rank
A	10	1
B	9	2
C	9	**3**
D	3	4

Fig. 6. h does not fulfil the property of (strict) monotony. The increase in the number of citations of paper C (from 4 to 9) is not associated to a corresponding increase of h, which remains unchanged.

String 1 (h=3)

paper id	citations	rank
A	9	1
B	8	2
C	3	**3**
D	2	4

String 2 (h=3)

paper id	citations	rank
A	8	1
B	7	2
C	3	**3**

$h \in$ □

Fig. 7. Example of equivalent classes with $h = 3$.

Scholar *a*

citations	rank
12	1
7	2
6	3
4	**4**
3	5
2	6
1	7
1	8
1	9

h-index	4
P	9
C	36

Scholar *b*

citations	rank
20	1
15	2
10	**3**
3	4
3	5
3	6
3	7
3	8
3	9
3	10
3	11

h-index	3
P	11
C	69

$h(a) > h(b)$
$P(a) < P(b)$
$C(a) < C(b)$

Fig. 8. The ordinal property of the *h*-scale can be in contrast to other empirical measures, such as the total number of publications (P) or the total number of citations (C).

h-Index is used to compare the scientists' research output. This is possible because the *h* measurement scale has the ordinal property. In other terms, if we compare two scientists – *a* and *b* – and if $h(a) > h(b)$, then *a* is considered better than *b*. However, this scale property can be in contrast to other measures of diffusion and productivity. Let consider the example in Fig. 8, in which two scholars (*a* and *b*) are compared. In this case, the indication provided by the *h*-index is in contrast with the indication provided by other measures (total number of papers produced – *P* – and total number of citations – *C*). According to the *h*-index, $h(a) = 4 > \ h(b) = 3$. On the other hand, if we use *P* and *C* as empirical measures of diffusion and productivity, scholar *b* is considered better than *a*. This example shows that the aggregation made by *h* may introduce some anomalies. However, this kind of "rank reversal" is rare because of the general positive correlation of the *h*-index with the total number of citations and the total number of papers (Joint Committee on Quantitative Assessment of Research, 2008). More precisely, *h* is approximately proportional to the square root of *C* and linearly proportional to *P* (Hirsch, 2005; Van Raan, 2006; Burrell, 2006; Burrell, 2007b; Antonakis and Lalive, 2008).

We show that the *h*-scale does not have interval properties. This means that it is not possible to say how much *a* is better than *b*, according to the difference $h(a) - h(b)$, regardless of the values of $h(a)$ and $h(b)$. An indirect demonstration is that, for high values of the *h*-index, it becomes more and more difficult to increase it (Egghe, 2007). In other words, the "gap" between two scientists with *h*-indexes 16 and 18 is much larger than the gap between two scientists with *h*-indexes 3 and 5.

Considering a scholar's scientific production, Hirsch empirically showed that the total number of citations (C) is approximately proportional to h^2 (Hirsch, 2005):

$$C \approx ah^2. \tag{1}$$

Thus, *h* value is roughly proportional to $C^{1/2}$. The same result is confirmed by the study of Egghe and Rousseau (2006): they prove that *h* is proportional to $C^{1/\alpha}$, where α equals the exponent in the law of Lotka, which most classical value is 2. According to Eq. (1), Fig. 9 illustrates the values of *h* depending on the value of *C*. For the purpose of simplicity *a*is assumed to be unitary.

Considering a particular class (h), the distance from the higher consecutive class ($h + 1$) – in terms of number of citations – can be calculated using the relationship of Eq. (1):

$$C(h + 1) - C(h) \approx a \cdot (h + 1)^2 - ah^2 = a \cdot (2h + 1). \tag{2}$$

As shown in Eq. (2), the distance (in terms of citations) between two consecutive *h* classes increases proportionally with the *h* value. For example, if $h = 5$ this distance is about $a \cdot 11$(a times 11), and if $h = 10$ it is about $a \cdot 21$(a times 21).

Another key to the reading of the same issue is provided by Anderson et al. (2008): considering the minimum possible citation records associated to a specific *h*-index score, we assume that an author with $h = 1$ has a single paper that has one citation. Subsequently, $h = 2$ is achieved with two papers each with two citations. To move from $h = 1$ to $h = 2$, an additional 3 citations are required, one for the first paper and two for the second paper. In turn, moving from $h = 2$ to $h = 3$ requires a further 5 citations (one for the first two papers and three for the third), and so on.

Additionally, it can be noticed that *h* does not necessarily reflect compositions of the input citation strings. This fact is exemplified in Fig. 10. In case (a), the union of two strings (*a* and *b*), with the corresponding *h*-indexes of 3 and 5, respectively, originates a third citation string ($a + b$) with a corresponding *h* of 5. In case (b), the union of two citation strings (*a* and *c*), with both *h*-indexes of 3, originates a third citation string ($a + c$) with a corresponding *h* of 5.

From this example it emerges that there is no direct relationship between the *h* values of the single input citation strings, and the *h* values of the corresponding joined string. A similar but more detailed study of the problem of merging different citation strings is presented by Egghe (2008).

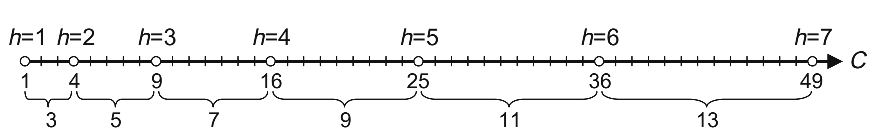

Fig. 9. *h* values represented on the *C* axis. For the purpose of simplicity we assumed $a = 1$, so $h \approx C^{1/2}$. The distance (in terms of citations) between two consecutive *h*-classes is not constant.

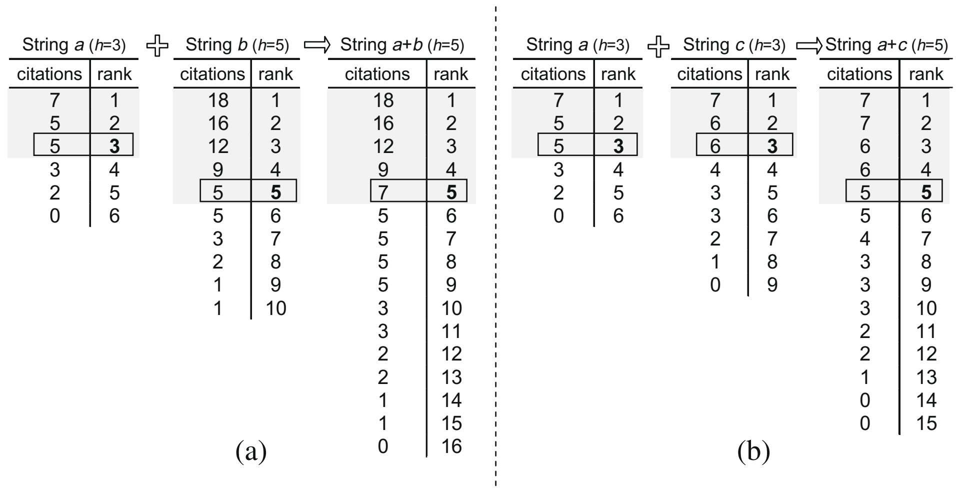

Fig. 10. h does not reflect compositions of the input citation strings. In case (a), string a ($h = 3$) is joined to string b ($h = 5$), obtaining a string with corresponding $h = 5$. In case (b), string a ($h = 3$) is joined to string b ($h = 3$), obtaining a string with corresponding $h = 5$. There is no direct relationship between the h values of the single input citation string, and the h values of the corresponding joined strings.

3.7. Further considerations on the h properties

After discussing some peculiar properties of h, it is natural to ask oneself: "what are the ideal properties that we would like h to satisfy?" As said before, measuring the research contribution of scholars is not an easy task. If – for a moment – we forget about the complexity of the problem, we would like h able to show the research output of a scholar by a *representative* number: ideally, an indicator able to describe in detail all the aspects of one scholar's career. It is probably impossible to conceive a single synthetic indicator with the characteristics. This because the synthesis of every problem–not necessarily in the bibliometric field – produces a simplification, which (at least partially) undermines the problem completeness (Carreras, 2005; Mingers, in press). The same issue is analyzed in an interesting way by Woeginger, who provides a characterization of h and other h-based bibliometric indicators from the viewpoint of some symmetry axioms that capture certain desired elementary properties (Woeginger, 2008a; Woeginger, 2008b; Woeginger, 2008c; Rousseau, 2008; Gagolewski and Grzegorzewski, in press; Quesada, 2009).

Table 3 reports some of the most popular bibliometric indicators (including h) and compares them with respect to the most important properties discussed in the previous subsections. These results are obtained by extending the analysis that has been carried out for h. Being not extendable to other indicators, the property *Papers' passage in the h-core* is omitted. Also, it can be noticed that the *Compensation* and *Monotony* properties can not be applied to P and C because they are basic indicators.

As expected, Table 3 shows that there is not an indicator satisfying all the quoted properties. However, we concur with Glänzel (2006) who describes h as a interesting and simple index representing "a useful supplement to the bibliometric toolset". In our opinion, h – compared to other competing indicators – is effective in providing a quick and easily understandable approximation to a researcher's profile. Some studies suggest that the h is indeed a promising (rough) measurement of the quality; in fact, the statistical correlation of h with standard bibliometric indicators and peer judgment was shown to be quite high (Schreiber, 2007). However, as recommended by Hirsch himself, many other factors should be considered in combination in evaluating the multifaceted profile of one individual (Hirsch, 2005).

We leave open for future work a more detailed analysis of additional properties, which are not strictly peculiar to h and other h-based bibliometric indicators.

Table 3
Comparison table of the most popular bibliometric indicators, according to the properties discussed in Section 2.

Indicator name	Author(s) and publ. date	Short description	Comparison of scholars with different seniority	Compensation	Monotony	Scale properties
P	–	Total number of publications	Satisfied	Not applicable	Not applicable	Ratio
C	–	Total number of citations	Satisfied	Not applicable	Not applicable	Ratio
CPP	–	Mean citations per paper (C/P)	Not satisfied	Satisfied	Satisfied	Ratio
h-Index	Hirsch (2005)	The number (h) such that, for a general group of papers, h papers received at least h citations while the other papers received no more than h citations	Not satisfied	Not satisfied	Not satisfied	Order
(raw) h_{ratio}	Burrell (2007c)	$h(T)/T$, h is the Hirsch index and T is the career time length of a scholar	Satisfied	Not satisfied	Not satisfied	
AR-index	BiHui et al. (2007)	$AR = \sqrt{\sum_{j=1}^{h} \frac{C_j}{a_j}}$, where a_j denotes the age of an article j in the h-core and C_j the corresponding citations	Not satisfied	Satisfied	Satisfied	
g-Index	Egghe (2006)	The largest number (g) for which the g most cited papers have a total of at least g^2 citations	Not satisfied	Not satisfied	Not satisfied	Order

4. A proposal for complementing h

As mentioned before, several attempts to improve *h* and reduce its intrinsic limitations have been done. Nevertheless, most of the proposed corrections and new suggested methods tend to undermine its original qualities of easy calculation and immediate intuitive meaning.

Citations apart, the *h*-index does not use some information about publications, which is readily available using on-line databases. They consist of:

- the number of authors of each publication;
- the date of each publication;
- the destination of each publication (journal, conference, etc.).

Particularly, the publication date can be very important for:

1. providing a temporal collocation of the most important publications;
2. determining the duration of the scientific production. This can be useful to have an idea of the continuity of a scholar's research output over time.

As remarked before, *h* does not take into account the time-width of the scientific research. Some variants of the *h*-index, proposed by Sidiropoulos et al. (2007), BiHui (2007), BiHui et al. (2007), give less importance to older articles introducing an age-related weighting system. However, these attempts complicate the calculation of a final synthetic indicator, introduce some weights and parameters which can be questionable.

A simpler way to take into account the age of the publications consists in complementing *h* with an additional indicator (f), defined as the time-range of the papers with at least one citation (added to 1 to consider the time spent to prepare the first paper included in the set):

$$f = \underset{i \in \Omega}{\text{range}}(y_1, y_2, , \ldots, y_i, \ldots) + 1, \qquad (3)$$

being:

y_i publication year related to the *i*th paper;
Ω set of publications which have been cited at least once.

In the Example in Fig. 11, *h* (equivalent to 7) is complemented with the information of *f* (equivalent to 10). Using a short form, it can be said that the examined scholar is *h*7-*f*10.

f provides a rough indication on the temporal extension of one author's scientific production expressed in years. Non cited articles are not taken into account. *f* is a natural number which cannot decrease with time and cannot exceed the total duration of an author's scientific production.

It is interesting to build a map associating *h* values with *f* values (see Fig. 12).

The map can be divided, for example, into four operating regions, with reference to the evolution of the career of a scholar. They are:

- region I (low *h* and low *f*): debut of a scholar;
- region II (large *h* and low *f*): scholar with a large number of citations, accumulated in a small time (brilliant young scholar);
- region III (large *h* and large *f*): scholar with a large number of citations distributed over a large time interval (good and continuous productivity);
- region IV (low *h* and large *f*): scholar with few citations, accumulated over a large time interval (poor productivity).

The size of this map (in terms of maximum values of *h* and *f*), as well as the division in quadrants is not univocal. This because the average scientific production of scholars is strongly dependent on the scientific field and the operating context (Imperial and Rodriguez-Navarro, 2007).

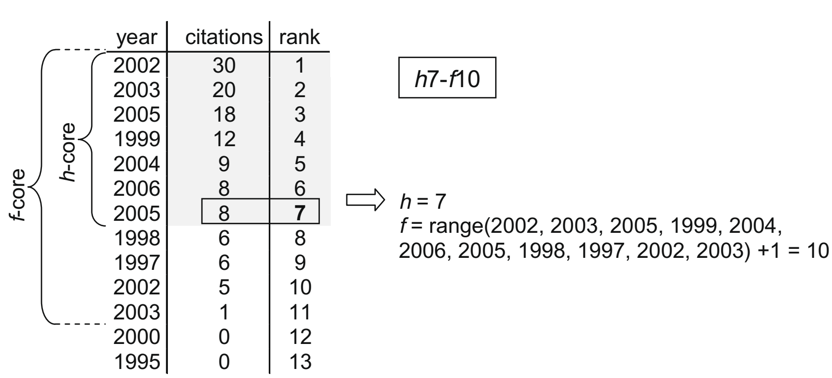

Fig. 11. Example of *h*-index complemented with *f*-index.

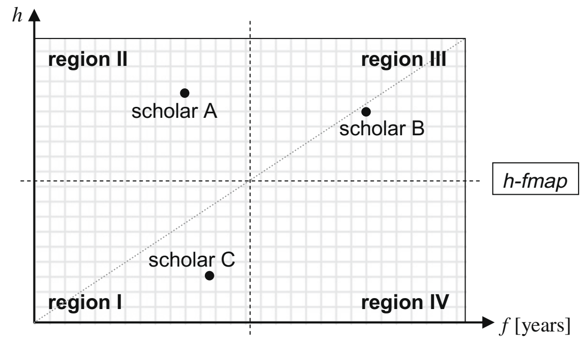

Fig. 12. Map associating h values (vertical axis) with f values (horizontal axis).

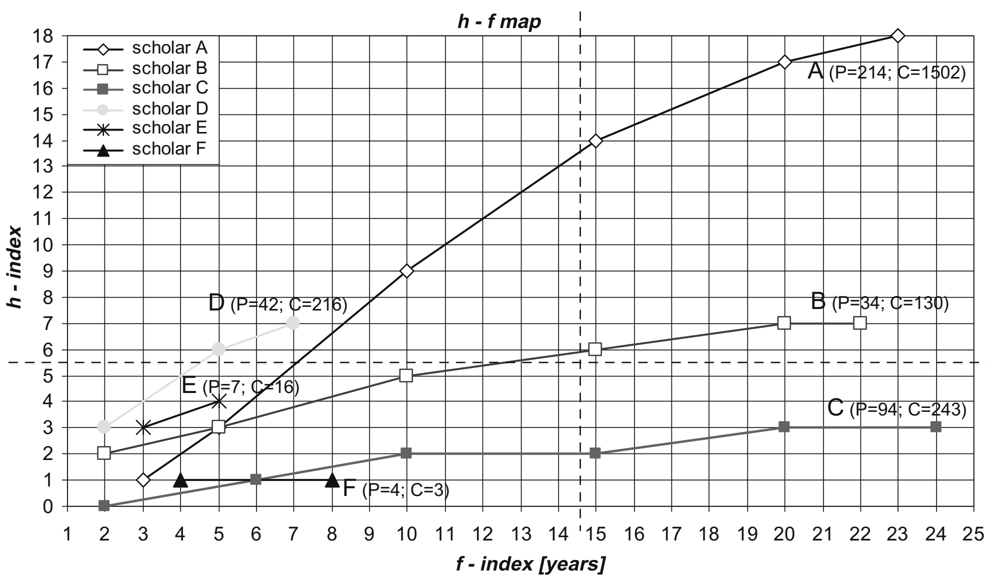

Fig. 13. Comparison of the scientific production of six anonymous scholars from Politecnico di Torino (Italy). The map makes it possible to monitor the evolution of the scholars' output over time. Points in the *hf*-plane are obtained calculating the h and f indexes for different periods of activity (e.g. for scholar D: 1, 4 and 6 years after the publication of the first work), considering the papers and citations accumulated till the precise period of interest. In order to give an idea of the typical relationship between P (total number of publications), C (total number of citations) and the map regions, each scholar's P and C values are reported in brackets.

The map can be used for comparing different scholars and also for monitoring the evolution of their scientific production over time. Also, it is interesting to notice that the *hf*-map can be somehow related to the Hirsch's m coefficient and the Burrell's h-rate (Hirsch, 2005; Burrell, 2007c).

For the purpose of example, the map in Fig. 13 illustrates the evolution of the scientific production of six anonymous scholars from Politecnico di Torino (Italy), in the industrial engineering scientific field. The map is obtained calculating the h and f indexes at different time periods of activity, considering the papers and citations accumulated by the scholars in precise periods (e.g. 3, 5 or 10 years after the publication of the first work).

Scholars A, B and C are senior scientists with more than 20 years of activity. It can be noticed that scholars A and B have a good productivity, while the scientific production of scholar C is inferior and has a smaller growth rate. In addition, the map illustrates the scientific outputs of three young scholars (D, E and F), with less than 10 years of activity. Scholar D stands out against the other two for its good productivity concentrated in a short period (7 years). On the other hand, the productivity of scholar F can be considered rather poor.

5. Conclusions

The h-index is a relatively recent bibliometric indicator for assessing the research output of scientists. Owing to the immediate intuitive meaning and simplicity of calculation, h received much attention in the scientific community. On the other hand, the h-index greatest disadvantage is the not very high accuracy, consequent upon many implicit assumptions and simplified hypotheses behind its definition and construction.

The original contribution of this work is to analyze h defining its operational properties and discussing them in detail, with the help of examples. Particular attention is given to the ordinal property of the h-scale. Finally, an additional indicator (f) that complements h taking into account the age of the publications is suggested. This new indicator enriches the information provided by h, with a minimum additional effort. h and f can be represented in a map, which is useful to compare different scholars and to monitor the evolution of their scientific production over time.

Future activities will focus on the use of h-based performance indicators in other fields, like manufacturing or Quality engineering.

References

Anderson, T.R., Hankin, R.K.S., Killworth, P.D., 2008. Beyond the Durfee square: Enhancing the *h*-index to score total publication output. Scientometrics 76 (3), 577–588.
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.
Banks, M.G., 2006. An extension of the Hirsch index: Indexing scientific topics and compounds. Scientometrics 69 (1), 161–168.
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.
BiHui, J., 2007. The AR-index: Complementing the *h*-index. ISSI Newsletter 3 (1), 6.
BiHui, J., LiMing, L., Rousseau, R., Egghe, L., 2007. The R- and AR-indices: Complementing the *h*-index. Chinese Science Bulletin 52 (6), 855–963.
Bornmann, L., Daniel, H.D., 2007. What do we know about the *h* index? Journal of the American Society for Information Science and Technology 58 (9), 1381–1385.
Braun, T., Glänzel, W., Schubert, A., 2006. A Hirsch-type index for journals. The Scientist 69 (1), 169–173.
Burrell, Q.L., 2006. Hirsch's *h*-index: A stochastic model. Journal of Informetrics 1 (1), 16–25.
Burrell, Q.L., 2007a. Should the *h*-index be discounted? In: Glänzel, W., Schubert, A., Schlemmer, B. (Eds.), In the Multidimensional World of Tibor Braun: A Multidisciplinary Encomium for his 75th Birthday, Leuven, ISSI, pp. 65–68. <http://www.issi-society.info/tiborbraun75/tiborbraun75.pdf>.
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.
Carreras, F., 2005. A decisiveness index for simple games. European Journal of Operational Research 163 (2), 370–387.
Castillo, C., Donato, D., Gionis, A., 2007. Estimating number of citations using author reputation. In: String Processing and Information Retrieval. Springer, Berlin, Heilberg, pp. 107–117.
Egghe, L., 2006. Theory and practise of the *g*-index. Scientometrics 69 (1), 131–152.
Egghe, L., Rousseau, R., 2006. An informetric model for the Hirsch-index. Scientometrics 69 (1), 121–129.
Egghe, L., 2007. Dynamic *h*-index: The Hirsch index in function of time. Journal of the American Society for Information Science and Technology 58 (3), 452–454.
Egghe, L., 2008. The influence of merging on *h*-type indices. Journal of Informetrics 2 (3), 252–262.
Franceschini, F., Galetto, M., Maisano, D., 2006. Classification of performance and quality indicators in manufacturing. International Journal of Services and Operations Management (2), 294–311.
Franceschini, F., Galetto, M., Maisano, D., 2007. Management by Measurement: Designing Key Indicators and Performance Measurement Systems. Springer-Verlag, Berlin.
Franceschini, F., Galetto, M., Maisano, D., Mastrogiacomo, L., 2008. Properties of performance indicators in operations management: A reference framework. The International Journal of Productivity and Performance Management (57), 137–155.
Franceschini, F., Maisano, D., in press. The hirsch index in manufacturing and quality engineering. Quality and Reliability Engineering International. doi:10.1002/qre.1016.
Gagolewski, M., Grzegorzewski, P., in press. A geometric approach to the construction of scientific impact indices. Scientometrics, doi:10.1007/s11192-008-2253-y.
Garfield, E., 2006. The history and meaning of the journal impact factor. JAMA 295 (1), 90–93.
Glänzel, W., 2006. On the opportunities and limitations of the *h*-index. Science Focus 1 (1), 10–11.
Harzing, A.W., van der Wal, R., 2008. Google Scholar as a new source for citation analysis. Ethics in Science and Environmental Politics 8 (11), 61–73.
Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America 102, 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.
Korhonen, P., Tainio, R., Wallenius, J., 2001. Value efficiency analysis of academic research. European Journal of Operational Research 130 (1), 121–132.
Imperial, J., Rodriguez-Navarro, A., 2007. Usefulness of Hirsch's *h*-index to evaluate scientific research in Spain. Scientometrics 71 (2), 271–282.
Joint Committee on Quantitative Assessment of Research, 2008. Citation Statistics. <http://www.mathunion.org>.
Katsaros D., Sidiropoulos A., Manolopous Y., 2007. Age decaying *h*-index for social network of citations. In: Proceedings of Workshop on Social Aspects of the Web Poznan, Poland, April 27, 2007.
Kelly, C.D., Jennions, M.D., 2006. The *h* index and career assessment by numbers. Trends in Ecology & Evolution 21 (4), 167–170.
Lehmann, S., Jackson, A.D., Lautrup, B.E., 2005. Measures and Mismeasures of Scientific Quality. <http://arxiv.org/abs/physics/0512238>.
Lehmann, S., Jackson, A.D., Lautrup, B.E., 2006. Measures for measures. Nature 444, 1003–1004.
Meho, L.I., Yang, K., 2007. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar. Journal of the American Society for Information Science and Technology 58 (13), 2105–2125.
Mingers, J., in press. Measuring the research contribution of management academics using the Hirsch-index. Journal of the Operational Research Society. doi:10.1057/jors.2008.94.
Mingers, J., 2008. Exploring the dynamics of journal citation: Modelling with *s*-curves. Journal of the Operational Research Society 59, 1013–1025.
Moed, H.F., 2005. Citation Analysis in Research Evaluation, Springer, Dordrecht. ISBN: 1402037139.
Montgomery, D.C., 2005. Introduction to Statistical Quality Control, fifth ed. Wiley, New York.
Orbay, M., Karamustafaoglu, O., Oner, F., 2007. What does Hirsch index evolution explain us? A case study: Turkish Journal of Chemistry. Biblios 27 (8), 1–5.
Quesada, A., 2009. Monotonicity and the Hirsch index. Journal of Informetrics 3 (2), 158–160.
Rousseau, R., 2006. New developments related to the Hirsch index. E-prints in Library and Information Science (ELIS). <http://eprints.rclis.org>.
Rousseau, R., 2008. Woeginger's axiomatisation 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.
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.
Schreiber, M., 2008. A modification of the *h*-index: The h_m-index accounts for multi-authored manuscripts. <http://arxiv.org/abs/0805.2000v1>.
Schreiber, M., 2007. Self-citation corrections for the Hirsch index. EuroPhysics Letters 78. doi:10.1209/0295-5075/78/30002.
Seglen, P.O., 1992. The skewness of science. Journal of the American Society for Information Science 43, 628–638.
Sidiropoulos, A., Katsaros, D., Manolopoulos, Y., 2007. Hirsch *h*-index for disclosing latent facts in citation networks. Scientometrics 72 (2), 253–280.
Van Raan, A.F.J., 2006. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. Scientometrics 67 (3), 491–502.
Wendl, M., 2007. H-index: However ranked, citations need context. Nature 449, 403.
Woeginger, G.J., 2008a. An axiomatic characterization of the Hirsch-index. Mathematical Social Sciences 56 (2), 224–232.
Woeginger, G.J., 2008b. An axiomatic analysis of Egghe's *g*-index. Journal of Informetrics 2 (4), 364–368.
Woeginger, G.J., 2008c. A symmetry axiom for scientific impact indices. Journal of Informetrics 2 (3), 298–303.