h-index: A Review Focused in its Variants, Computation and Standardization for Different Scientific Fields

S. Alonso[a,*], F.J. Cabrerizo[b], E. Herrera-Viedma[c], F. Herrera[c]

[a]*Software Engineering Department, University of Granada, Spain*
[b]*Department of Software Engineering and Computer Systems,*
Distance Learning University of Spain, Spain
[c]*Department of Computer Science and Artificial Intelligence, University of Granada, Spain*

Abstract

The h-index and some related bibliometric indices have received a lot of attention from the scientific community in the last few years due to some of their good properties (easiness of computation, balance between quantity of publications and their impact and so on). Many different indicators have been developed in order to extend and overcome the drawbacks of the original Hirsch proposal. In this contribution we present a comprehensive review on the h-index and related indicators field. From the initial h-index proposal we study their main advantages, drawbacks and the main applications that we can find in the literature. A description of many of the h-related indices that have been developed along with their main characteristics and some of the works that analyze and compare them are presented. We also review the most up to date standardization studies that allow a fair comparison by means of the h-index among scientists from different research areas and finally, some works that analyze the computation of the h-index and related indices by using different citation databases (ISI Citation Indexes, Google Scholar and Scopus) are introduced.

Key words: h-index, bibliometric indicators

1. Introduction

To measure scientific output of researchers is an increasingly important task for the scientific community. In fact, nowadays, almost every research assessment decision (accepting research projects, contracting researchers, awarding scientific prices, concede a grant and so on) depends to a great extent upon the scientific merits of the involved researchers.

To do so, several different indicators have been used. The majority of those indicators were developed to quantify both the production of researchers, e.g.

---

[*]Corresponding author

the total number published papers and the number of papers published in a period of time; and / or the impact of their publications, e.g. the total number of citations, the average number of citations per paper, the number and percentage of significant papers (with more than a certain amount of cites) and some indicators based on the impact of the journals as median impact factor of the journals where the papers are published, relative citation rates and so on. In the specialized literature there have been studies that point out how combining those indicators is a desirable option to obtain a global view of the scientific output of researchers (Leeuwen et al. (2003); Martin (1996)).

Moreover, as new scientific production databases are created and improved to incorporate more information about the production of researchers and citations of papers (some of the most important ones are ISI Web of Science (online resource), Scopus (online resource) and Google Scholar (online resource)) it has been possible to develop new and more robust indicators that take into account different characteristics of the production of scientists.

One of the most recent and successful indicators was proposed by Hirsch (2005) and it is called *h-index*. Even being a quite recent contribution, it has received a lot of attention from the scientific community with more than 150 published papers in the topic since its publication (`http://sci2s.ugr.es/hindex/biblio.php`). Some of the main advantages of the h-index are that it is simple to compute and that it takes into account both the quantity and the impact of the researcher's publications. This good properties made some of the most important journals to take interest on it (Ball (2005)).

The h-index has been extended by many authors that have proposed new variations of the h-index that try to extend it an overcome the main drawbacks of the original Hirsch proposal (Bornmann and Daniel (2009)). Among others, we can cite the g-index (Egghe (2006)), the hg-index (Alonso et al. (2009)), the A-index (Jin (2007)) and the m-index (Bornmann et al. (2008a)).

The aim of this paper is to make a comprehensive review on the h-index, its analysis, computation and related indicators. Specifically, the paper will be focused on the main following topics:

- The description of the original h-index proposal along with a detailed description of its main advantages and drawbacks according to different experts on h-index. Additionally, we show some of the applications developed using the h-index.

- The review of many of the indicators based on similar ideas of the h-index, showing their particular advantages and limitations.

- The presentation of the main comparative studies existing between the h-index and other bibliometric indicators used to evaluate researcher's scientific production. We will also present different h- based indices, their respective correlation studies, and some works about axiomatic and mathematical interpretations of the h- and related indices. Other studies that examine how self-citations affect h- related indices and that analyze the

2

performance of different indices and some of their possible transformations will also be presented.

- The standardization of h- related indices to compare the scientific production among different research areas.

- The computation of the h- index by using different citation databases (ISI Web of Science, Scopus and Google Scholar) and the advantages and disadvantages of using each database.

To do so, in section 2 we present the h-index, its main applications, advantages and disadvantages. In section 3 a description of many of the indicators based on the h-index published in the literature, along with their advantages and limitations is presented. In section 4 we present some published studies that analyze and compare h-related indices. In section 5 we show the studies about the standardization of the h-index in order to effectively compare researchers output from different scientific fields. In section 6 we present how different scientific production databases can be used to compute the h-index and related indices. Finally, in section 7 we summarize the main points of our review, we show some open questions and topics that should be addressed in the future and we point out some conclusions.

2. h-index: Advantages, Disadvantages and Applications

In the following, we pay attention to the original h-index proposal. Along with its definition, we will point out its major advantages and disadvantages and some of its applications that we can found in the specialized literature.

The original definition of the h-index, proposed by Hirsch (2005), was:

> "A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(Np - h)$ papers have $\leq h$ citations each."

In figure 1 the number of cites of the publications of a particular researcher have been represented in decreasing order. The h-index of the researcher corresponds to the point where the number of citations crosses the publication order. In later contribution, Hirsch (2007) conducted an empirical study in order to confirm that the h-index is better than other considered indicators (total citation count, citations per paper and total paper count) in predicting future scientific achievement.

As it has been previously said, the h-index has received a lot of attention since its formulation. As an example of its impact, in table 1 and figure 2 we show the number of citations to date (February 16, 2009) that the original paper has received since its publication according to different citation databases. From the citation counts it is clear that the h-index is an increasingly important hot topic.

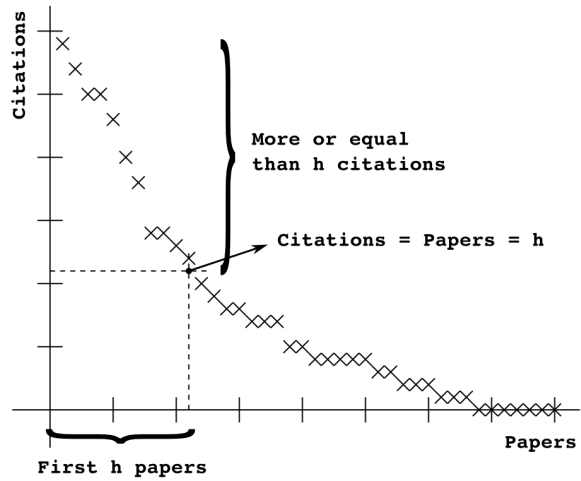[1] Google Scholar reports 414 cites when the search is made without the publication date

3

Figure 1: Graphical interpretation of the h-index

| Citation Database | Number of Cites | | | | | |
	2005	2006	2007	2008	2009	Total
Google Scholar	3	45	108	124	1	*281 (414)*[1]
Scopus	1	29	83	141	13	*267*
ISI Web of Science	1	28	91	115	12	*247*

Table 1: Number of citations received by the original h-index paper Hirsch (2005) according to 3 different citation databases (citations until February 16, 2009)
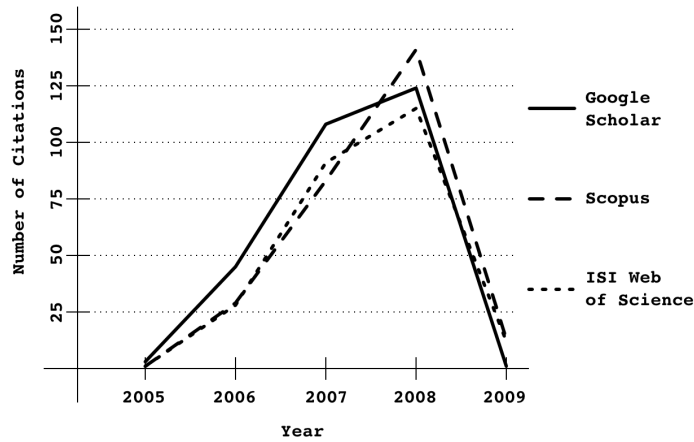

Figure 2: Number of citations received by the original h-index paper (Hirsch (2005)) according to 3 different citation databases (citations until February 16, 2009)

2.1. Advantages and Disadvantages of the h-index

In his original proposal, Hirsch pointed out some of most interesting properties of this indicator. For example, the h-index is easy to compute (and verify) using the ISI Web of Science (online resource) just by ordering one scientist's publications by the *"Times Cited"* field. In addition, Hirsch stated that using a single indicator that mixes both quantity of publications and their impact simplifies to great extent the characterization of researchers' scientific output.

Costas and Bordons (2007) mentioned other good properties of the h-index. For example, it is an objective indicator and therefore, it may play an important role when making decisions about promotions, fund allocation and awarding prizes. Moreover, it performs better than other single-number criteria commonly used to evaluate the scientific output of a researcher (impact factor, total number of documents, total number of citations, citation per paper rate and number of highly cited papers).

Vanclay (2007) pays attention to another interesting benefit of the h-index: its robustness, in the sense that it is insensitive to a set of lowly cited papers (the difficulty of increasing the h-index grows exponentially as all the most cited papers of the researcher have to receive new cites to obtain a higher index). Besides, the great majority of errors in citation databases tend to occur in the lower citation portion of a researcher's scientific production which tends not to affect to the computation of the h-index. He also states that being the h-index an integer, it avoids the false impression of precision conveyed by the three decimal points in the ISI impact factor.

However, the h-index presents some drawbacks that have been pointed out in the literature (and often different authors have tried to overcome those drawbacks by defining new indicators, see section 3). Hirsch himself noted that there are inter-field differences in typical h values due to differences among fields in productivity and citation practices, so the h-index should not be used to compare scientists from different disciplines (Hirsch (2005)). Moreover, Hirsch noted that there exist some technical limitations, such as the difficulty to obtain the complete output of scientists with very common names (this problem decreases as citation databases improve their records and search engines). It is important to remark that this drawbacks are shared with almost any indicator that is based in citation counts as there are intrinsic differences in citation practices among fields, and that nowadays, the homonymity problem is an open issue for every bibliography database that is used to compute bibliometric indices.

As it happens with other indicators, another drawback is that the h-index depends on the duration of each scientist's career because the pool of publications and citations increases over time (Kelly and Jennions (2006)), and so it might not be suitable to compare scientists at different stages of their career.

_____

restriction. However, searching in a year by year basis do not show the same total amount of cites. This behaviour might occur due to lack of publication dates in many documents of their database.

Egghe's main concern with the h-index is that highly cited papers are important for the determination of the h-index, but once they are selected to belong to the top h papers, it is unimportant the number of citations they receive (Egghe (2006)). In fact, this feature contradicts to the selection criterion of impact indicators, namely, obtaining more citations, it may be regarded as higher impact (Plomp (1994); Vinkler (1996)). In practical terms, this means that researchers with some extremely cited papers may have a similar or equal h-index as researchers with moderate or high cited papers.

As any other single indicator which is simple to compute, we run the risk of indiscriminate use, such as relying only on it for the assessment of scientists. However, research performance is a complex multifaceted endeavour that usually cannot be assessed adequately by means of a single indicator (Martin (1996)). Leydesdorff (2008) also extends this problem to the case of using the h-index or similar indicators to evaluate journals or more general research activities.

The original formulation of the h-index takes into account self-citations (they are usually included in citation databases). This could provoke changes in the publishing behaviour of scientists, such an artificial increase in the number of self-citations distributed among the documents on the edge of the h-index (Purvis (2006); Raan (2006); Zhivotovsky and Krutovsky (2008)). In fact, several authors propose to exclude self-citations from any citation based index calculation (including the h-index) to obtain a fairer indicator (Schreiber (2007b); Vinkler (2007)). However, it is important to note that excluding self-citations makes the calculation of the h-index a much more complex task as usual citation databases do not automatically allow to avoid self-citations.

In a short letter, Wendl (2007) called attention to the fact that the context of citations is a critical issue for every productivity metric as *"Many citations are used simply to flesh out a paper's introduction, having no real significance to the work"* and *"Citations are also sometimes made in a negative context, or to fraudulent or retracted publications"*. Moreover, Wendl mentions the "Matthew effect", whereby well-established researchers and projects are cited disproportionately more often than those that are less widely known and how review articles receive an important citation bias. Both factors should have a negative impact on the validity of the h-index as an indicator to make research assessment decisions.

2.2. Applications of the h-index

Even being a quite recent proposal, the h-index have been used for different tasks and contexts. In the following we present some of those applications:

- Several authors have used the h-index directly to compare the scientific output of researchers: Hirsch (2005) originally presented a comparison between prominent physicists. On the contrary, Schreiber (2007a) has studied the h-index for 26 non-prominent physicists. Imperial and Rodríguez-Navarro (2007) and Alcaide et al. (2008) use the h-index to evaluate scientific research for several authors in different areas in the Biological Sciences and Biomedicine. Cronin and Meho (2006) and Oppenheim (2007) use

the h-index to rank influential information scientists. In Bornmann and Daniel (2005, 2007a) the h-index is studied and used for post-doctoral research fellowships applicants evaluation. Finally, Salgado and Paez (2007) studies the scientific productivity in the Spanish social psychology field.

- The h-index can also be used to measure the scientific output of complete research groups, institutions and groups of authors (Egghe and Rao (2008); Molinari and Molinari (2008b,a)). Raan (2006) uses it (among other indicators and peer judgement) to compare 147 different chemistry research groups. Mugnaini et al. (2008) presents a comparison based on the h-index between the Brazilian Academy of Sciences and the National Academy of Sciences of the USA.

- Braun et al. (2006) propose a Hirsch-type index for evaluating the scientific impact of journals as a robust alternative indicator that is an advantageous complement to usual journal impact factors. Vanclay (2008) uses a similar approach for the particular case of ranking forestry journals.

- Banks (2006) applies the h-index to study interesting topics and compounds: Bank's h-b-index is found by entering a topic (search string, like "superstring" or "teleportation") or compound (name or chemical formula) into the Web of Science database and then ordering the results in terms of citations, by largest first. The h-b-index is then defined in the same manner as the h-index. With the calculation of the h-b-index, it can be determined how much work has already been done on certain topics or compounds, what the "hot topics" (or "older topics") of interest are, or what topic or compound is mainstream research at the present time.

- Some attempts have also been made to adapt the h-index to evaluate research performance among different countries. For example, Guan and Gao (2008) make a comparison and evaluation of Chinese research performance in the field of bioinformatics using the h-index and compares it to other countries as USA, UK, Germany and so on. Csajbók et al. (2007) have also produced some ranked lists of world's countries attending to their h-index on various scientific fields and have concluded that EU countries have strong positions in each field, but none of them can successfully compete with the USA.

- Schubert (2007) proposed the use of successive h-indices to be able to evaluate scientific output not only of individual scientists, but of institutes, universities and even higher levels of aggregation. His idea is quite simple: the h-index of a level of aggregation n is h_n if h_n of its elements in the immediately lower level of aggregation have an h_{n-1} index of at least h_n each and the rest of elements have $h_{n-1} \leq h_n$. For example, the h_2 level for a research institute is $h_2 = 10$ if at least 10 researchers have an $h_1 \geq 10$ and the rest of researchers have a $h_1 \leq 10$. Note that h_1 refers to the usual h-index.

- Egghe (2008e) presented a model for successive h-indices based on his existing theory on the distribution of the h-index in Lotkaian informetrics and he introduces a global h-index for which tables of individuals (authors, institutes,...) are merged. As an example of application, he calculated the successive and global h-indices for the (still active) D. De Solla Price awardees.

- Following the previous idea by Schubert, Arencibia-Jorge et al. (2008) apply the idea of successive h-indices to perform a case study of the scientific production of different institutions using a *researcher-department-institution* hierarchy as levels of aggregation. They improve their study by using additional h- related indices as complementary indicators.

- A related work Glänzel (2008a) addresses the problem of concatenating h-indices without any other additional information, that is, to determine the h-index of the combination of two or more bibliographies were the only information about them is the h-indices of the different bibliographies.

- In addition to the evaluation of scientific output, Glänzel described two new bibliometric applications of h- related indices Glänzel (2008b). The first one is the composite indicator, which expresses a multiplicative connection between derivatives of publication output and citation impact. The second applications is the z statistics, that can be used to analyse the tail of citation distributions in the light of the h-index.

- Another different application of the h-index is the possibility of evaluating the production in different scientific topics. To do so, Bar-Ilan (2008a) proposes to compute the h-index of topics instead of authors. He discusses the methodological issues related to the computation of this index as well as he points out the difficulty of the collection of the data. He also introduces an example of his proposal with some different informetrics topics, being the h-index itself one of them.

3. New Indices Based on the h-index

Despite of the good properties of the h-index, several authors have pointed out several drawbacks of the indicator (see section 2.1). To overcome these drawbacks many new variations of the h-index have been proposed in the literature. In the following, we review these variations and we comment their main properties. For clarity reasons, we have splitted all the proposals into two different categories: i) new indices that try to complement the h-index and ii) new indices that extend the h-index to take into account other variables and dependencies (mainly time and number of authors).

3.1. New Indices That Complement the h-index

Holding that *"a measure which should indicate the overall quality of a scientist ... should deal with the performance of the top articles"*, Egghe (2006)

proposed the **g-index** as a modification of the h-index. For the calculation of the g-index, the same ranking of a publication set -paper in decreasing order of the number of citations received- is used as for the h-index. Egghe defines the g-index as:

> *"The highest number g of papers that together received g^2 or more citations. From this definition it is already clear that $g \geq h$."*

In contrast to the h-index, the g-index gives more weight to highly cited papers. The aim is to avoid a disadvantage of the h-index that *"once a paper belongs to the top h papers, its subsequent citations no longer 'count' "* (Bornmann et al. (2008a)). However, the g-index also suffers from some problems. For example, if a researcher's papers usually recieve few cites, but he achieves a big-hit paper with a extremely big citation count (a successful review paper, for example), the g-index could grow a lot in comparison with other researchers with a much higher average of cites in their papers (Alonso et al. (2009)).

Alonso et al. (2009) presented a new index, called **hg-index**, which is based on both h-index and g-index that tries to keep a balance between the advantages of both measures as well as to minimize their disadvantages. The hg-index of a researcher is computed as the geometric mean of his h- and g- indices, that is: $hg = \sqrt{h \cdot g}$. It is trivial to demonstrate that $h \leq hg \leq g$ and that $hg - h \leq g - hg$, that is, the hg-index corresponds to a value nearer to h than to g. This property can be seen as a penalization of the g-index in the cases of a very low h-index, thus avoiding the problem of the big influence that a very successful paper can introduce in the g-index. The authors noted that this new index has several advantages when compared with h- and g-indices individually: it has greater granularity, which makes easier to compare scientists with similar h- or g- indices; similar to the g-index, it takes into account the number of cites of the most cited papers of an author, but softens the influence that a very successful paper could introduce in the g-index and it is measured in the same scale as h- and g- indices which makes it easy to understand.

In a recent proposal, Eck and Waltman (2008) argued that the definition of the h-index is quite arbitrary: *"Hirsch could equally have defined the h-index as follows: A scientist has h-index of h is h of his n papers have at least $2h$ citations each and the other $n - h$ papers have $\leq 2h$ citations each"*. Thus, they extended the h-index and defined the h_α**-index** as:

> *"A scientist has h_α-index of h_α is h_α of his n papers have at least $\alpha \cdot h_\alpha$ citations each and the other $n - h_\alpha$ papers have fewer than $\leq \alpha \cdot h_\alpha$ citations each."*

where $\alpha \in (0, \infty)$. It is straightforward that $h_1 = h$ and that when $\alpha \neq 1$ the h_α-index need not be an integer. This approach may be useful in situations where the h-index is excessively restrictive. For example, if several authors share a high h-index it may be difficult to discriminate them (because to increase a high h-index is a difficult task as it requires a high number of papers to increment

their citation counts). In that cases an h_α-index with $\alpha < 1$ might show greater granularity among the scientists.

According to Burrell (2007) *"the h-index seeks to identify the most productive core of an author's output in terms of most received citations"*. In fact, Rousseau first introduced the term **Hirsch core** consisting of the first h papers (Liang et al. (2007); Rousseau (2006)): *"The Hirsch core can be considered as a group of high-performance publications, with respect to the scientist's career"*. Based on the definition of the Hirsch core, several authors have proposed new indicators. Jin (2006) was the first to propose the idea of use the average of the cites in the Hirsch core, although the particular name **A-index** for this indicator was first introduced by Rousseau. More formally, the A-index is defined as:

$$A = \frac{1}{h} \sum_{j=1}^{h} cit_j$$

where h is the h-index and cit_j is the number of citations of the j-th most cited paper. Similar as the g-index, the A-index has the particular advantage of taking into account the citations of the Hirsch core, and thus, the A-index may increase even if the h-index remains the same as the scientist's citation counts increase.

Liang et al. (2007) critically observed that with the A-index, *"the better scientist is 'punished' for having a higher h-index, as the A-index involves a division by h"*. Therefore, instead of dividing by h, the authors suggested taking the square root of the sum of citations in the Hirsch core to calculate a new index called the **R-index**. Formally, the r-index is defined as:

$$R = \sqrt{\sum_{j=1}^{h} cit_j}$$

where h is the h-index and cit_j is the number of citations of the j-th most cited paper. However, as the R-index -similarly as with the A-index- measures the citation intensity in the Hirsch core, the index can be very sensitive to just a very few papers receiving extremely high citation counts (Bornmann et al. (2008a)). We must remark that the R-, h- and A- indices are strongly related: $R = \sqrt{h \cdot A}$.

In a recent paper, Bornmann et al. (2008a) noticed that as the distribution of citation counts is usually skewed, the median and not the arithmetic average should be used as the measure of central tendency. Therefore, as a variation of the A-index, the **m-index** is proposed as the median number of citations received by papers in the Hirsch core. Using this approach it is possible to reduce the impact of heavily cited papers in the Hirsch core.

Following the idea presented in the g-index of giving more weight to the most highly cited papers, Kosmulski (2006) proposed the $h^{(2)}$-**index**:

> *"A scientist's $h^{(2)}$-index is defined as the highest natural number such that his $h^{(2)}$ most cited papers received each at least $[h^{(2)}]^2$*

citations."

For example, a $h^{(2)}$ index of 20 means that a scientist has published at least 20 papers of which each one has been cited over 400 times. It is obvious that for any scientist the $h^{(2)}$-index is always lower to his h-index. According to Liang et al. (2007), the main advantage of the $h^{(2)}$-index is that it reduces the precision problem because less work is needed to check the accuracy of the publications data, especially with regard to homographs, that is, to distinguish between scientists that have the same last name and first initial. As only few papers in the set are sufficiently highly cited in order to fulfill the criterion of $[h^{(2)}]^2$ citations, there are also fewer papers to check (Bornmann et al. (2008a)). However, as only a small subset of the authors papers will be used to compute the $h^{(2)}$-index, it could be difficult to discriminate between scientists having different number of publications with quite different citation rates for relatively high $h^{(2)}$-indices.

Sidiropoulos et al. (2007) point out that since scientists do not publish the same number of articles, the original h-index is not a fair enough metric. Thus, they define the **normalized h-index** as follows:

$$h^n = \frac{h}{N_p}$$

where h is the h-index and N_p is the total number of papers published by the researcher. However, this particular metric does in fact reward less productive researchers, which probably is negative property.

Anderson et al. (2008) presented a new proposal that they called **tapered h-index** that presents the interesting property of taking into account all cites (not only from the Hirsch core). The main idea for computing this index comes from a representation of the cites of the papers in a Ferrers graph (see figure 3) where each column represents a partition of the cites amongst papers. The largest completed (filled in) square of points in the upper left hand corner of a Ferrers graph is called the Durfee square. Following this interpretation, the h-index is equal to the length of the side of the Durfee square (in the case of Figure 3, $h = 3$), effectively assigning no credit (zero score) to all points that fall outside.

The main idea of the tapered h-index is to evaluate all the citations giving to each of them a value equal to the inverse of the increment that would supposse to increase the h-index one unit. For example, in Figure 3, we can see that to increment the h-index from 3 to 4, 7 citations are needed. Thus, all the citations that are in the size 4 Durfee square will add $1/7$ to the final tapered h-index. The tapered h-index of the cites represented in Figure 3 is $H_T = 3+4\cdot1/7+3\cdot1/9+2\cdot1/11+1\cdot1/13 \simeq 4.16$. A more formal definition can be found in Anderson et al. (2008). This approach has the clear advantages of taking into account the complete production of a researcher. However, the computations needed to obtain this measure and the difficulty of obtaining accurate data from bibliographic databases even for the less cited works of the author may make it difficult to implement.
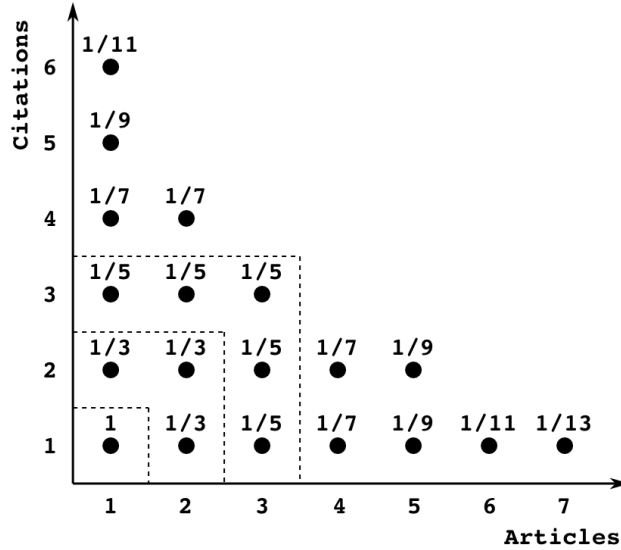
Figure 3: Example of a Ferrers diagram of an author's citations

Following a similar but simpler approach than to the tapered h-index, Ruane and Tol (2008) presented the rational h-index h_{rat}-**index**, which has the advantage of increasing in smaller steps than the standard h-index (it could provide more granularity to an evaluation process). This index is defined as:

$$h_{rat} = (h+1) - \frac{n_c}{2 \cdot h + 1}$$

where h is the h-index, n_c is the number of citations that are already needed to obtain a h-index of $h+1$ (it corresponds to the blank spaces in the Durfee square of size $h+1$). Note that $2 \cdot h + 1$ corresponds to the maximum amount of cites that could be needed to increment the h-index in one unit. It is obvious that $h \leq h_{rat} < h+1$. It a very recent contribution, Guns and Rousseau (2009a) review different real and rational variants of the h- and g- indices.

3.2. New Indices That Take Into Account Other Variables and Dependencies

Some of the critics to the h-index that can be found in the literature is that it does not take into account several different variables that are often useful to evaluate the production of researchers. The h- based indices that we present in the following were introduced to take into account some of those variables, as time and number of authors:

According to a stochastic model for an author's production/citation process, Burrell (2007) conjectures that the h-index is approximately proportional to career length. Thus, an improvement to the h-index to compare scientists with

different lengths of scientific careers is to divide it by the number of years of research activity. For this reason, Hirsch (2005) already proposed dividing the h-index by number of years since a scientist's first publication and called it the **m-quotient**. Formally, the m-quotient is defined as:

$$m = \frac{h}{y}$$

where h is the h-index, and y is the number of years since the first paper published by the author.

Jin (2007) introduced a new proposal to complement the R-index and called it the AR-index. This index takes into account not only the citation intensity of the Hirsch core, but also the age of the publications in the core:

> *"The AR-index is defined as the square root of the sum of the average number of citations per year of articles included in the h-core."*

In fact, Liang et al. (2007) state that taking into account the age of the publications is a necessary condition to be able to evaluate performance changes. Thus, the AR-index can increase or decrease over time, a property that is not shared by the previously mentioned indices. Formally, the AR-index is defined as:

$$AR = \sqrt{\sum_{j=1}^{h} \frac{cit_j}{a_j}}$$

where h is the h-index, cit_j is the number of citations of the j-th most cited paper, a_j is the number of years since the publication of the j-th most cited paper.

Egghe and Rousseau (2008) presented a new h-index variation that they called citation-weighted h-index (h_w-**index**) which is, as the AR-index, sensitive to performance changes. It is defined as:

$$h_w = \sqrt{\sum_{j=1}^{r_0} cit_j}$$

where cit_j is the number of citations of the j-th most cited paper, r_0 is the largest row index i such that $r_w(i) \leq cit_i$ and $r_w(i) = \frac{\sum_{j=1}^{i} cit_j}{h}$. It is clear that indicators that are sensitive to performance changes can be useful in certain environments. For example, they could allow to identify journals whose impact is changing (in both a positive or a negative way).

Other time dependent indices were proposed by Sidiropoulos et al. (2007). In that contribution, the authors investigate how the h-index could be adapted to take into account the age of the papers, thus differentiating between senior scientists that have received many of their cites due to papers that where published a long time ago and brilliant young scientists who are expected to contribute a

large number of significant works in the near future but that nowadays have a small number of important papers due to the time constraint, and the age of the cites that would allow to identify scientists whose contributions are still influential even if they were published a long time ago and would allow to disclose trendsetters, i.e., scientists whose work is considered pioneering and sets out a new line of research that currently is hot ("trendy"), thus this scientist's works are cited very frequently. For the first case, they defined the **contemporary h-index** (Sidiropoulos et al. (2007)):

> "A researcher has contemporary h-index h^c if h^c of its N_p articles get a score of $S^c(j) \geq h^c$ each, and the rest $(Np - h^c)$ articles get a score of $S^c(j) \leq h^c$."

where $S^c(j)$ is a novel score for an article j based on citation counting defined as:

$$S^c(j) = \gamma \cdot (Y(now) - Y(j) + 1)^{-\delta} \cdot cit_j$$

where $Y(j)$ is the publication year of an article j and cit_j is the number of citations of article j. If we set $\delta = 1$, then $S^c(j)$ is the number of citations that article j has received, divided by the "age" of the article. Since the number of citations is divided with the time interval, the quantities $S^c(j)$ will be too small to create a meaningful h-index; thus, the coefficient γ is used. This way, an old article gradually loses its "value", even if it still gets citations. In other words, in the calculations, we mainly take into account the newer articles. For the second case, the authors defined the **trend h-index**:

> "A researcher has trend h-index h^t if h^t of its N_p articles get a score of $S^t(j) \geq h^t$ each, and the rest $(N_p - h^t)$ articles get a score of $S^t(j) \leq h^t$."

where $S^t(j)$ is a score for article j defined as:

$$S^t(j) = \gamma \cdot \sum_{\forall x \in cit_j} (Y(now) - Y(x) + 1)^{-\delta}$$

where γ, δ and $Y(j)$ are defined as in the contemporary h-index.

A similar aproach was taken by Rousseau and Ye (2008). The authors presented a **dynamic h-type index** which depends on the Hirsch core, the actual number of citations received by articles belonging to the Hirsch core, and the recent increase in h. It tries to detect situations where two scientists have the same h-index and the same number of citations in the h-core, but that one has no change in his h-index for a long time while the other scientist's h-index is on the rise. To do so, the definition contains three time-dependent elements: the size and contents of the Hirsch core, the number of citations received, and the h-velocity:

$$h_d = R(T) \cdot v_h(T)$$

where $R(T)$ denotes the previously defined R-index computed at time T and v_h is the h-velocity. In practice it is imperative to stablish a starting point, $T = 0$ and a way to determine v_h. The authors suggested that the starting point should not be the beginning of a scientist's career, but when T is "now", then $T = 0$ can be taken 10 or 5 years ago (or any other appropriate time). In addition, if one has a good-fitting continuous model for $h(t)$ over this period, then this function should be used to determine $v_h(T)$.

In addition to the new indices that take into account the time variable, some other authors have been interested in the adaptation of the h-index to take into account the number of co-authors of the publications. For example, in a recent contribution, Schreiber proposes a new index called the h_m-**index** (Schreiber (2008d,c)) that pretends to soften the influence of the number of co-authors for a researcher's publications. To do so, it counts the papers fractionally according to the number of authors, for example, only as one third for three authors. This yields an effective number which is utilized to define the h_m-index as that effective number of papers that have been cited h_m or more times. To compute the h_m-index we part from a ranking of the researcher's papers, where j is the rank attributed to a paper when the publication list of an author is sorted by the number cit_j of citations. Thus, h-index is determined as $h = max(j) \mid j \leq cit_j$ where each paper is fully counted for the (trivial) determination of its rank $j = \sum_{j'=1}^{j} 1$. Counting a paper with $auth_j$ authors only fractionally, i.e. by $1/auth_j$ yields an effective rank $j_{eff}(j) = \sum_{j'=1}^{j} \frac{1}{auth_{j'}}$ which is used to define the h_m-index as

$$h_m = max(j) \mid j_{eff}(j) \leq cit_j.$$

In the same paper Schreiber mentions that an alternative but similar approach called h_I-**index** was used by Batista et al. (2006) and further discussed in Bornmann and Daniel (2007b); Imperial and Rodríguez-Navarro (2007). In this approach, the papers are counted fractionally according to the mean number of authors of the papers contributing to the index. To compute it, first the average number of authors of the first j papers is computed as $\overline{auth}(j) = \frac{1}{r} \sum_{j'=1}^{j} auth'_j$. This average number of authors is used as a normalization yielding

$$h_I = \frac{h}{auth(h)}.$$

Egghe (2008d) has also considered the problem of multiple authors. In his paper, he studies two different possibilities to compute h- and g- indices for authors using a fractional crediting system in which papers with several authors do have less weight than single authored papers. He studies two different approaches: the first one is called "fractional counting on citations", and implies giving an author of an m-authored paper only a credit of $\frac{c}{m}$ if the paper received c citations. The second one is called "fractional counting on papers" where for each author in a m-authored paper, the paper occupies only a fractional rank of $\frac{1}{m}$ (note that this is a equivalent approach as the previously presented h_m-index). This kind of approaches to take into account the number of authors in

15

papers might be useful to reward scientists whose papers are entirely produced by themselves from the authors that work in big groups (and which naturally would publish a larger amount of papers).

4. Studies Analyzing the Indices

In this section we present some studies dedicated to the analysis and comparison of the different h- related indices:

- Firstly, we discuss some of the studies that compare the h-index with other bibliometric indicators to evaluate researcher's scientific production (subsection 4.1),

- we also present some studies that analyze different h- based indices and their correlations and some papers that establish some axioms and mathematical interpretations of the h- and related indices (subsection 4.2),

- we introduce some works that examine how self-citations affect h- related indices (subsection 4.3),

- we present some studies that stablish some axioms and mathematical interpretations of h- based indices (subsection 4.4) and, finally,

- we present other studies that analyze the performance of different indices and some of their possible transformations (subsection 4.5).

4.1. Studies Comparing h-index and Other Bibliometric Indicators

Bornmann et al. (2008b) used some comprehensive data sets of applicants to the long-term fellowship and young investigator programmes of the European Molecular Biology Organization. They determined the relationship between the h-index and three standard bibliometric indicators (total number of publications, total citation counts, and average journal impact factor) as well as peer assessments to test the convergent validity of the h-index. Their results suggest that the h-index is a promising rough measurement of the quality of a young scientist's work as it is judged by internationally renowned scientists.

Costas and Bordons (2007) analyze the relationship of the h-index with other bibliometric indicators at the micro level for Spanish CSIC scientists in Natural Resources, using publications downloaded from the Web of Science (1994–2004). Different activity and impact indicators are obtained to describe the research performance of scientists in different dimensions, being the h-index located through factor analysis in a quantitative dimension highly correlated with the absolute number of publications and citations. The need to include the remaining dimensions in the analysis of research performance of scientists and the risks of relying only on the h-index are stressed. The hypothesis that the achievement of some highly visible but intermediate-productive authors might be underestimated when compared with other scientists by means of the h-index is tested. The authors suggest that the h-index tends to underestimate the

achievement of scientists with a "selective publication strategy", that is, those who do not publish a high number of documents but who achieve a very important international impact. In addition, a good correlation is found between the h-index and other bibliometric indicators, especially the number of documents and citations received by scientists, that is, the best correlation is found with absolute indicators of quantity. Finally, they notice that the widespread use of the h-index in the assessment of scientists' careers might influence their publication behaviour. It could foster productivity instead of promoting quality, and it may be increasing the presence of least publishable units or "salami publications", since the maximum h-index an author can obtain is that of his/her total number of publications.

Lehmann et al. (2008) employ Bayesian statistics in order to analyze the h-index and several other different indicators of scientific performance to try determine each indicator's ability to discriminate between scientific authors. They demonstrate that the best of the indicators that they studied requires approximately 50 papers to draw conclusions regarding long term scientific performance. In addition, they stated how their approach allows a statistical comparison among scientists from different fields.

Van Leeuwen (2008) conducted a bibliometric study in the Netherlands focusing on the level of the individual researcher in relation to an academic reward system. He compared the h-index with various bibliometric indicators and other characteristics of researchers and tested its usefulness in research assessment procedures. He found that there is a strong bias towards the research field(s) in which the researcher is active, and thus, he concludes that this limits the validity of the h-index for the specific interest of evaluation practices.

4.2. Studies that Analyze h- Based Indices and Their Correlations

Schreiber (2008a) works out 26 practical cases of physicists from the Institute of Physics at Chemnitz University of Technology, and compare the h and g values. It is demonstrated that the g-index discriminates better between different citation patterns. As expected, the g-index allows for a better discrimination between the datasets and yields some rearrangement of the order. The rearrangements can be traced to different individual citation patterns, in particular distinguishing between one-hit wonders and enduring performers: One-hit wonders advance in the g-sorted list. In his opinion, this makes the g-index more suitable than the h-index to characterize the overall impact of the publications of a scientist. Specially for not-so-prominent scientists, the small values of h do not allow for a reasonable distinction between the datasets. This also can be achieved by evaluating the A-index, which reflects the average number of citations in the h-core, and interpreting it in conjunction with the h-index. h and a can be combined into the r-index to measure the Hirsch core's citation intensity. He also determines the a and R values for the 26 datasets. For a better comparison, he uses interpolated indices. The correlations between the various indices as well as with the total number of papers and the highest citation counts are discussed. The largest Pearson correlation coefficient is found between g and R. Although the correlation between g and h is relatively strong,

17

the arrangement of the datasets is significantly different depending on whether they are put into order according to the values of either h or g.

Bornmann et al. (2008a) examined empirical results on the h-index and its most important variants in order to determine whether the developed variants are associated with an incremental contribution for evaluation purposes. The compared indices were: the m-quotient, the g-index, the $h^{(2)}$-index, the A-index, the R-index, the AR-index, the m-index and the h_w-index. The results of the analysis indicate that with the h-index and its variants, we can assume that there are two types of indices: (i) The first type of indices (h-index, m-quotient, g-index and $h^{(2)}$-index) describe the most productive core of the output of a scientist and tell us the number of papers in the core, and (ii) the second type of indices (A-index, m-index, R-index, AR-index and h_w-index) depict the impact of the papers in the core.

In a very recent paper Bornmann et al. (2009) make a further empirical analysis on the h-index and several of its variants (g-index, $h^{(2)}$-index, a-index and r-index) to measure the performance of journals. In their case, they compare 20 organic chemistry journals with those indices and with the Journal Impact Factor and they found very high intercorrelations among all indices. Thus, the authors conclude that all the examined measures could be called redundant for empirical applications.

Costas and Bordons (2008) analyze the ability of both the g-index and h-index to discriminate between different types of scientists (low producers, big producers, selective scientists and top scientists) in the area of Natural Resources at the Spanish CSIC (Web of Science, 1994–2004). Their results show that these indicators clearly differentiate low producers and top scientists, but do not discriminate between selective scientists and big producers. However, they show that the g-index is more sensitive than the h-index in the assessment of selective scientists, since this type of scientist shows in average a higher g-index/h-index ratio and a better position in g-index rankings than in the h-index ones. Therefore, this research suggests that these indexes do not substitute each other but that they are complementary.

Schubert and Glänzel (2007) presented a theoretical model of the dependence of h- related indices on the number of publications and the average citation rate. They successfully tested it against some empirical samples of journal h-indices. Their results demostrated that it is possible to stablish a kind of "similarity transformation" of h-indices between different fields of science.

4.3. Studies About How Self Citation Affect the h-index

As aforementioned, one of the most criticized aspects of h- related indices is the possible influence of self citations. In fact, Schreiber (2007b) studies several anonymous datasets concludes that self-citations do have a great impact on the h-index, specially in the case of young scientists with a low h-index. In his paper he proposes three different ways to sharpen the h-index to avoid the self citation problem. Each proposal has an increasing level of difficulty as usual citation databases do not allow to easily differentiate among self citations and external citation.

18

In a later work, Schreiber (2008b) studies how both the h- and g- indices are affected by self citations by means of an analysis of nine practical cases in the physics field. He concludes that the g-index is more influenced by self citations than the h-index and thus, he proposes to exclude those citations in the computation of the g-index.

However, it is important to note that some authors do not agree with the argument of self citation being an important drawback of h- related indices. For example, L. Engqvist (2008) argue that to increase one's own h-index would be necessary to cite many self papers and that it is difficult to predict which papers should be cited in order to improve the author's h-index. In fact, they performed a literature study, selecting 40 authors from the fields of evolutionary biology and ecology and identified the citation causing their most recent increases in h. Next, they distinguished the first citation appearing thereafter, which would have caused the same increase in the author's h. The difference between the publication dates of these two citations give the time that the h-index is dependent on one single citation. This time measure is an estimation of how long selective self-citation of target papers would be effective.

4.4. Studies That Stablish Some Axioms and Mathematical Interpretations of h- Based Indices

Woeginger (2008b) provides a new axiomatic characterization of the h-index in terms of three natural axioms (concerning the addition of single publications, the addition of new citations to old publications and the joint case of adding new publications and citations). The author extended his work (Woeginger (2008c)) stablishing a new axiomatic characterization based on a simple symmetry axiom which imposes that the number of citations and the number of publications should be treated in the same way and measured in the same scale. Woeginger (2008a) has recently applied a similar axiomatic characterization to the g-index. Rousseau (2008) has proposed a slight generalization of Woeginger's characterization of the h-index which has allowed him to check that other h- related indices also satisfy the previous axioms.

Torra and Narukawa (2008) have stablished the connection of the h-index (and the number of citations) with the Choquet and Sugeno integrals. In particular they show that the h-index is a particular case of the Sugeno integral and that the number of citations corresponds to the Choquet integral (in both cases using the same fuzzy measure). This conclusion allow the authors to envision new indexes defined in terms of fuzzy integrals using different types of fuzzy measures.

Liang (1996) studied how the h-index changes over time using time series. After his initial work there have been several studies about time series, the h-index and its mathematical properties (Guns and Rousseau (2009b); Liu and Rousseau (2008); Egghe (2009)). In a related paper, Egghe and Rousseau (2006) presented an informetric model for the h-index in which the index is derivated using different power laws and hence, they could derivate some of its mathematical properties. Ye and Rousseau (2008) complemented the previous work to find out if power law models for a specific type of h-index time series fit real

data sets. Rousseau (2007) has also used a continuous power law model in order to show that the influence of missing articles is largest when the total number of publications is small and non-existing when the number of publications is very large (the same conclusion is drawn for missing citations).

4.5. Other Studies that Analyze the Performance of Different Indices and Their Transformations

To illustrate the necessity of a decreasing index in a concrete application, Liang et al. (2007) calculated the h-index, R-index, and the AR-index for the articles written by B.C. Brookes (Brookes, who was the Derek de Solla Price Medallist in 1989, died in 1991). They concluded that *"Brookes' h-index over the whole period (2002-2007) stays fixed at h = 12 (hence here h > AR). Between 2002 and 2007 his R-index increased by 5% while the AR-index decreased by about 5%"*.

Egghe (2008a,c) makes a comparative study about how the h-index, the g-index, the R-index and the h_w-index are affected by simple transformation as doubling the production per source, doubling the number of sources, doubling the number of sources but halving their production, halving the number of sources but doubling their production (fusion of sources) and some special cases of general power law transformations. The author demonstrated that this kind of transformations affect in a similar way to all the h- related indices that he studied.

Egghe (2008b) has also studied the importance of merging h- type indices for different information production processes. In fact, he studies two types of information production processes mergings for the h-, g-, R- and h_w- indices: one where common sources add their number of items and one where common sources get the maximum of their number of items in the two information production processes.

5. Standardization of the h-index to Compare Scientists that Work in Different Fields

Even in Hirsch's initial proposal, the fact that the h-index cannot directly be used to compare research workers of different areas, mainly due to lack of normalization for reference practices and traditions in the different fields of science (Glänzel and Moed (2002); Pinski and Narin (1976)) was pointed out. It is worth to note that the differences in the bibliometric features of the fields influences not only the h-index and related indices, but any indicator that uses citations. To try to overcome this problem, different standardizations of the h-index to compare scientists that work in different scientific fields have been developed. In the following we introduce three of those studies.

Iglesias and Pecharromán (2007) suggest a rational method to account for different citation practices, introducing a simple multiplicative correction to the h-index which depends basically on the ISI field the worker researcher works in, and to some extent, on the number of papers the researcher has published.

To do so, the authors, using the average number of citations per paper and fixing the physics field as reference computed several normalizing factors for the rest of scientific fields. In fact, the authors provide a very practical table of these normalizing factors (see Iglesias and Pecharromán (2007), page 313), so the corrected h remains relatively simple to obtain, and thus, they allow an easy comparison between scientists from different fields.

Imperial and Rodríguez-Navarro (2007) suggest that, in general, publications in applied areas are less cited that publications in dynamic, basic areas, and therefore, scientists in the former areas show lower values of h. These differences are mainly caused by:

(i) the different sizes of the populations that can potentially cite the publication, and

(ii) the lower emphasis placed on research by scientists in applied areas.

Although the complex dependence of h on the citing population size precludes an overall h normalization across scientific areas, they empirically observed that the highest h values attained for a given area correlate well with the impact factor of journals in that area. They calculated h-indexes for the most highly cited scientists in different areas and subareas (reference h-index or h_R) and observed that h_R indexes are more dependent on journal impact factors than on specific publication patterns. In general, and for most areas, they observe that $h_R \sim 16 + 11f$ where f is the impact factor of the top journals that characterize that specific scientific area or subarea. Since h_R exhibits a linear dependence on f, it is possible to compute it as an average for scientists who publish in more than one area.

Podlubny and Kassayova (2006) proposed a new approach to produce multidisciplinary lists of highly cited researchers. Their approach is based on the recently discovered law of the constant ratio (Podlubny (2005)) that says:

> "The ratio of the total number of citations in any two fields of science remains close to constant."

In fact, several previous works (Rousseau (1992); Schubert and Glänzel (1983); Vinkler (2004)) had previously related the chance for citations with regards to the field of study. The law of the constant ratio allows normalization of all fields with respect to mathematics, where the total number of citations is always the smallest, although also growing with time. In addition, the authors produced the first multidisciplinary list of the top 200 highly cited researchers by selecting the first 100 highly cited researchers in each of the 22 ISI fields (see Podlubny and Kassayova (2006), pages 158–161), normalizing their citation counts using the obtained constant ratios among fields and finally merging them into a single list.

Finally, Radicchi et al. (2008) study the distributions of citations received by a single publication within several disciplines, spanning broad areas of science. The authors show that the probability that an article is cited a number of times has large variations between different disciplines, but all distributions

are rescaled on a universal curve when they consider a relative indicator. In fact, they show that the same behavior occurs when citation distributions of articles published in the same field, but in different years, are compared. Thus, they propose an unbiased generalization of the h-index suitable for comparing scientists working in different fields.

6. The Computation of the h-index Using Different Databases: ISI Web of Science, Scopus and Google Scholar

Until just a few years ago, when citation information was needed, the single most comprehensive source was the ISI Citation Indexes. Although the Citation Indexes were often criticized for various reasons, there was no other source to rely on. Data from the ISI Citation Indexes and the Journal Citation Reports are routinely used by promotion committees at universities all over the world. In this paper we refer to the Web version of Citation Indexes, i.e., to the ISI Web of Science (online resource). In fact, the original h-index proposal relied on the ISI Citation Indexes as its main source of information.

However, in the last few years, different alternatives to the ISI Citation Indexes have become available. Some of those new bibliography databases are discipline-oriented, like *Chemical Abstracts* produced by the American Chemical Society, *MathSciNet* by the American Mathematical Society and *PsycINFO* by the American Psychological Association (Neuhaus and Daniel (2008)). However, there are still not many studies that use this thematic sources to compute the h- and related indices. On the other hand, there are two new multidisciplinary databases that are increasingly being used for scientific evaluation purposes:

- The first one is Scopus (online resource) developed by Elsevier, and

- the second one is the freely available Google Scholar (online resource).

Each of these has a different collection policy which affects both the publications covered and the number of citations to the publications (Bar-Ilan (2008b)). In the following we present some studies that analize the results of computing the h-index and its variations by using these three different citation databases that can be found in the literature.

- In a recent paper, Jacso (2008a) aims to provide a general overview, focusing on the analysis of pros and cons of the three largest, cited-reference-enhanced, multidisciplinary databases (Google Scholar, Scopus, and Web of Science) for determining the h-index. In addition, the practical aspects of determining the h-index also need scrutiny, because some content and software characteristics of reference-enhanced databases can strongly influence the h-index values.

- In a following paper, Jacso (2008b) focus on Google Scholar, from the perspective of calculating the h-index for individuals and journals. The paper shows that effective corroboration of the h-index and its two component

indicators can be done only on persons and journals with which a researcher is intimately familiar. Corroborative tests must be done in every database for important research. Furthermore, the paper highlights the very time-consuming process of corroborating data, tracing and counting valid citations and points out Google Scholar's unscholarly and irresponsible handling of data.

- In a third paper, Jacso (2008c) focuses on the practical limitations in the content and software of the databases that are used to calculate the h-index for assessing the publishing productivity and impact of researchers. To celebrate F.W. Lancaster's biological age of seventy-five, and "scientific age" of forty-five, this paper discusses the related features of Google Scholar, Scopus, and Web of Science, and demonstrates in the latter how a much more realistic and fair h-index can be computed for F.W. Lancaster than the one produced automatically. Browsing and searching the cited reference index of the 1945–2007 edition of Web of Science, which in his estimate has over a hundred million "orphan references" that have no counterpart master records to be attached to, and "stray references" that cite papers which do have master records but cannot be identified by the matching algorithm because of errors of omission and commission in the references of the citing works, can bring up hundreds of additional cited references given to works of an accomplished author but are ignored in the automatic process of calculating the h-index. The partially manual process doubled the h-index value for F.W. Lancaster from 13 to 26, which is a much more realistic value for an information scientist and professor of his stature.

- Meho and Rogers (2008) examine the differences between Scopus and Web of Science in the citation counting, citation ranking, and h-index of 22 top human-computer interaction (HCI) researchers from EQUATOR large British Interdisciplinary Research Collaboration project. Results indicate that Scopus provides significantly more coverage of HCI literature than Web of Science, primarily due to coverage of relevant ACM and IEEE peer-reviewed conference proceedings. No significant differences exist between the two databases if only citations in journals are compared. Although broader coverage of the literature does not significantly alter the relative citation ranking of individual researchers, Scopus helps distinguish between the researchers in a more nuanced fashion than Web of Science in both citation counting and h-index. Scopus also generates significantly different maps of citation networks of individual scholars than those generated by Web of Science. The study also presents a comparison of h-index scores based on Google Scholar with those based on the union of Scopus and Web of Science. The authors conclude that Scopus can be used as a sole data source for citation-based research and evaluation in HCI, especially when citations in conference proceedings are sought, and that researchers should manually calculate h scores instead of relying on system calculations.

- Using citations to the work of 25 library and information science (LIS) faculty members as a case study, Meho and Yang (2007) examine the effects of using Scopus and Google Scholar on the citation counts and rankings of scholars as measured by Web of Science. Overall, more than 10.000 citing and purportedly citing documents were examined. Results show that Scopus significantly alters the relative ranking of those scholars that appear in the middle of the rankings and that Google Scholar stands out in its coverage of conference proceedings as well as international, non-English language journals. The use of Scopus and Google Scholar, in addition to Web of Science, helps to reveal a more accurate and comprehensive picture of the scholarly impact of authors. The Web of Science data took about 100 hours of collecting and processing time, Scopus consumed 200 hours, and Google Scholar 3.000 hours.

Remark: At the time of the writing, ISI Web of Science does include conference proceedings citations, and thus, the differences in the computation of the h-index with Scopus and Google Scholar that have been previously pointed out have been minimized.

7. Discussion and Concluding Remarks

In this contribution we have made a state of the art review on the h-index and related indicators. The h-index, even being a quite recent proposal, has received plenty of attention from the scientific community due to its good properties when measuring the scientific production of researchers: for example, it takes into account both the production of the researcher and its impact and it is easy to compute using existing bibliography databases. However, after the initial h-index proposal, many different bibliometry researchers have pointed out some drawbacks of the h-index. Among other drawbacks, it has been noted that the h-index does not allow to directly compare scientists from different disciplines (although many other indicators do suffer from the same problem), that it is very sensitive to some variables as career length, being then not suitable to compare senior scientists and new ones, that it is quite insensitive to changes in the citation counts of the most cited papers of a researcher, that as self citations increase the h-index, it may induce changes in the publishing behaviour of scientists (just to increase their h-index) and finally that the h-index does not take into account the context of the citations.

To overcome these drawbacks different authors have proposed new variations of the h-index. Some of those variants that have received a great attention from the scientific community are the g- and A-indices, which do explicitly take into account the citation counts in the h-core, thus avoiding the problem of the h-index insensitivity to citation count changes; the m-index, which reduces the effect that a very successful paper could introduce in some of the h-index variations and the $h^{(2)}$-index whose main advantage is to reduce the amount of data that has to be retrieved from the bibliography databases to be computed.

Other researchers have introduced several variations of the h-index to take into account variables that were not originally included in the h-index proposal. For example, we can cite the AR-index that takes into account the age of the publications in the Hirsch core, thus being able to detect and evaluate performance changes and the h_m-index to reduce the advantage and influence that multi-authored papers could introduce in the original h-index.

We have also presented many studies that analyze the h-index and its variants in both the theoretical and empirical sides. One of the most interesting facts that we observe in those studies is that many h-index variations, although being designed to overcome some of its supposed limitations, do indeed correlate quite heavily. This fact has made some researchers think that there is probably no need to introduce more h-index variations if it is not possible to prove that they are not redundant in real examples.

We have also paid special attention to the works devoted to use the h-index and its variations to compare scientists from different research areas. As it has been said, this is a quite difficult problem due to the inherent differences in citation practices among different research fields. Although it is not an exclusive problem of the h-index (many other bibliometric indicators based on citation counts do present similar issues), there have been several different efforts in the literature to overcome it.

Finally, we have reviewed some works that compare the h- and related indices when they are computed using different bibliography databases. To date, those works have been centered in the use of ISI Web of Science, Scopus and Google Scholar. Some of the most interesting results in this aspect is that the h-index is quite dependant on the database that it is used and that, in general, it is much more difficult to compute those indices using Google Scholar than ISI Web of Science or Scopus.

The h-index has been a very productive bibliometry field in the last years, and it will probably continue being a hot topic in the future. In fact, there are still several open questions and topics that should be addressed in forthcoming contributions. For example, we can cite:

- Produce deeper analyses and comparisons among all h- related indices. Those new studies should include real world examples to test the correlation among the indices and then analyze the adequacy of using the different variations or even the inclusion of new proposals.

- As new bibliography databases are being produced (specially discipline-oriented ones) it is necessary to conduct different studies to test the different h- variants among the new data. Moreover, to test the easiness of computation of those measures using new databases is an interesting question as it may be an important factor to use certain indicators (indicators that are difficult to compute are not usually used in real evaluation problems).

- As many different theoretical models and studies in the h-index topic have been developed, it is necessary to apply them to real evaluation problems

and situations: for example, additional studies comparing individual researchers in particular fields, comparing journals, institutions and even regions or countries are still to be developed.

- As research assessment decisions can be partially based on measures and indices like the h-index, there is still room to develop evaluation models and decision support systems that incorporate these indicators to reach good decisions.

It is important to note that this review does not pretend to be a complete list of all the publications in the topic. For additional material and a complete list of publications about the h-index and related indices, please visit

http://sci2s.ugr.es/hindex/biblio.php

Acknowledgements

References

Alcaide, G., Gomez, M., Zurian, J., Benavent, R., 2008. Scientific literature by spanish authors on the analysis of citations and impact factor in biomedicine (1981-2005). Revista Española de Documentación Científica 31 (3), 344–365.

Alonso, S., Cabrerizo, F., Herrera-Viedma, E., Herrera, F., 2009. hg-index: A new index to characterize the scientific output of researchers based on the h- and g- indices. Scientometrics In press.

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

Arencibia-Jorge, R., Barrios-Almaguer, I., Fernandez-Hernandez, 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.

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

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

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

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

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

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

Bornmann, L., Daniel, H., 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., 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., 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 of new metrics with the journal impact factor. European Journal of Organic Chemistry 10, 1471–1476.

Bornmann, L., Mutz, R., Daniel, H., 2008a. 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., Wallon, G., Ledin, A., 2008b. Is the h-index related to (standard) 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.

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

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

Costas, R., Bordons, M., 2007. Advantages, limitations and its relation with other bibliometric indicators at the micro level. Journal of Informetrics 1 (3), 193–203.

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

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.

Eck, N. V., Waltman, L., 2008. Generalizing the h- and g-indices. Journal of Informetrics 2 (4), 263–271.

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

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

Egghe, L., 2008c. 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., 2008d. Mathematical theory of the h- 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., 2008e. Modelling successive h-indices. Scientometrics 77 (3), 377–387.

Egghe, L., 2009. Mathematical study of h-index sequences. Information Processing and Management In press.

Egghe, L., Rao, I., 2008. 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., 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 & Management 44 (2), 770–780.

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., Moed, H., 2002. Journal impact measures in bibliometric research. Scientometrics 53 (2), 171–193.

Google Scholar, online resource. `http://scholar.google.com/`.

Guan, J., Gao, X., 2008. Comparison and evaluation of Chinese research performance in the field of bioinformatics. Scientometrics 75 (2), 357–379.

Guns, R., Rousseau, R., 2009a. Simulating growth of the h-index. Journal of Informetrics 3, 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.

Hirsch, J., 2005. An index to quantify an individual's scientific research out-put. Proceedings of the National Academy of Sciences 102, 16569–16572.

Hirsch, J., 2007. Does the h-index have predictive power? Proceedings of the National Academy of Sciences 104, 19193–19198.

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

Imperial, J., Rodríguez-Navarro, A., 2007. Usefulness of hirsch's h-index to evaluate scientific research in Spain. Scientometrics 71 (2), 271–282.

ISI Web of Science, online resource. `http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/`.

Jacso, 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.

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

Jacso, 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.

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.

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

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.

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

Leeuwen, T. V., Visser, M., Moed, H., Nederhof, T., Raan, A. V., 2003. The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. Scientometrics 57 (2), 257–280.

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

Leydesdorff, L., 2008. Caveats for the use of citation indicators in research and journal evaluations. Journal of the American Society for Information Science and Technology 59 (2), 278–287.

Liang, B. J. L., Rousseau, R., Egghe, L., 2007. The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin 52 (6), 855–863.

Liang, L., 1996. h-index sequence and h-index matrix: Constructions and applications. Scientometrics 69 (1), 153–159.

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

Martin, B., 1996. The use of multiple indicators in the assessment of basic research. Scientometrics 36 (3), 343–362.

Meho, L., 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., 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.

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

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

Mugnaini, R., Packer, A., 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.

Neuhaus, C., Daniel, H., 2008. Data sources for performing citation analysis - an overview. Journal of Documentation 64 (2), 193–210.

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.

Pinski, G., Narin, F., 1976. Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. Information Processing and Management 12 (5), 297–312.

Plomp, R., 1994. The highly cited papers of professors as an indicator of a research group's scientific performance. Scientometrics 29 (3), 377–393.

Podlubny, I., 2005. Comparison of scientific impact expressed by the number of citations in different fields of science. Scientometrics 64 (1), 95–99.

Podlubny, I., Kassayova, K., 2006. Law of the constant ratio. towards a better list of citation superstars: compiling a multidisciplinary list of highly cited researchers. Research Evaluation 15 (3), 154–162.

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

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

Radicchi, F., Fortunato, S., Castellano, C., 2008. Universality of citation distributions: toward an objective measure of scientific impact. Proceedings of the National Academy of Sciences 105 (45), 17268–17272.

Rousseau, R., 1992. Why am I not cited or why are multi-authored papers more cited then others? Journal of Documentation 48, 79–80.

Rousseau, R., 2006. New developments related to the Hirsch index. Industrial Sciences and Technology, Belgium, available at `http://eprints.rclis.org/6376/`.

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

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), 263–372.

Rousseau, R., Ye, F., 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., 2008. Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics 75 (2), 395–405.

Salgado, J., 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.

Schreiber, M., 2007a. A case study of the Hirsch index for 26 non-prominent physicists. Annalen der Physik 16 (9), 640–652.

Schreiber, M., 2007b. Self-citation corrections for the Hirsch index. EPL (Europhysics Letters) 78 (3), 30002.

Schreiber, M., 2008a. 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.

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

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

Schreiber, M., 2008d. To share the fame in a fair way, h_m for multi-authored manuscripts. New Journal of Physics 10 (040201), 1–9.

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

Schubert, A., Glänzel, W., 1983. Statistical reliability of comparisons based on the citation impact of scientific publications. Scientometrics 5, 59–74.

Schubert, A., Glänzel, W., 2007. A systematic analysis of Hirsch-type indices for journals. Journal of Informetrics 1 (3), 179–256.

Scopus, online resource. `http://www.scopus.com/`.

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

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.

Van Leeuwen, T., 2008. Testing the validity of the hirsch-index for research assessment purposes. Research Evaluation 17 (2), 157–160.

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

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

Vinkler, P., 1996. Some practical aspects of the standarization of scientometric indicators. Scientometrics 35 (2), 237–245.

Vinkler, P., 2004. Characterization of the impact of sets of scientific papers: The garfield (impact) factor. Journal of the American Society for Information Science and Technology 55 (5), 431–435.

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.

Wendl, M., 2007. h-index: however ranked, citations need context. Nature 449 (7161), 403–403.

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

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

Woeginger, G., 2008c. A symmetry axiom for scientific impact indices. Journal of Informetrics 2 (4), 298–303.

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

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