Journal of Informetrics 7 (2013) 693–698

Contents lists available at ScienceDirect

Journal of Informetrics

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

Are you in *h*?

Marek Kosmulski*

Lublin University of Technology, Lublin, Poland

ARTICLE INFO

Article history:
Received 28 February 2013
Received in revised form 24 April 2013
Accepted 25 April 2013
Available online 22 May 2013

Keywords:
h-index
Scientific output
Ranking of scientists
Ranking of universities

ABSTRACT

A new method of assessment of scientific papers, scientists, and scientific institutions was defined. The significance of a paper was assessed by the definition of the largest (the most prestigious) set, including that paper in its *h*-core. The sets of papers were defined by affiliation (country, city, university, department) or by subject (branches and sub-branches of science, journal). The inclusion of a paper in the *h*-core of certain set(s) was used as an indicator of the significance of that paper, and of the scientific output of its author(s), of their scientific institution(s), etc. An analogous procedure was used to assess the contribution of an individual to the scientific output of his/her scientific institution, branch of science, etc.

1. Introduction

In the original Hirsch' (2005) approach a set of papers of one author is characterized by its *h*-index. The top *h* papers constitute the *h*-core, which can be considered as a set of significant papers of its author. The assignment of a paper to the *h*-core is dynamic, that is, papers enter or exit the *h*-core in the course of time. There is a problem (ambiguity) with assignment of either paper to the *h*-core when the papers ranked *h* and $h+1$ have exactly *h* citations each.

A *h*-type-index can be calculated (and the *h*-core can be defined) at other aggregation levels than author, e.g., journal (Braun, Glanzel, & Schubert, 2006; Harzing & van der Wal, 2009), university (Torres-Salinas, Moreno-Torres, Delgado-Lopez-Cozar, & Herrera, 2011), and even country (Csajbok, Berhidi, Vasas, & Schubert, 2007).

The *h*-type-indices are calculated for pre-defined sets of papers (aggregation levels). In the present approach (which will be referred to as RUinh, which is an abbreviation of aRe yoU in *h*) the aggregation level is not pre-defined, but different aggregation levels are considered, which include given paper in their *h*-cores. Inclusion (of the paper) in the *h*-core is a pre-requisite, and the definition of the set of papers (aggregation level) is a result (score). Every paper, which has been cited at least once is in the *h*-core(s) of certain set(s) of papers. Among different sets, which satisfy the condition (the paper of interest is included in the *h*-core) we are looking for the largest (the most prestigious) one. This study is limited to the simplest definitions based on affiliation (country, city, university, department) or subject (branches and sub-branches of science, journal). The procedure of search for definitions of the largest (most prestigious) sets including the paper of interest in their *h*-cores will be discussed in the next section using an example of specific paper.

2. Assessment of single paper

Paper 1: Thermochim. Acta 412:47(2004), published by authors from Abo Akademi (Turku, Finland), and Lublin University of Technology (Lublin, Poland) has received 134 citations, according to the Web of Science® database provided by Thomson

* Corresponding author. Tel.: +48 81 5384355.
E-mail addresses: mkosmuls@abo.fi, mkosmuls@hektor.umcs.lublin.pl

1751-1577/$ – see front matter
http://dx.doi.org/10.1016/j.joi.2013.04.007

Table 1
The sizes of *h*-cores of sets of papers including Thermochim.Acta 412:47, by affiliation.

Level	Description	No. of papers	*h*	Description	No. of papers	*h*
Country	Poland	440,494 (421,284)	339 (319)	Finland	282,226 (259,846)	439 (400)
City	Lublin	20,682 (21,675)	96 (87)	Abo/Turku	47,735 (39,653)	208 (190)
University				Abo Akademi University	9070 (8168)	112 (107)

Reuters®, accessed on February 26, 2013 (further referred to as WoS). Paper 1 has received exactly the same number of citations according to Scopus (which is a fortunate coincidence).

2.1. By affiliation

Paper 1 is the most-cited paper of one of its authors and the 3rd most-cited paper of each of the other 2 authors. Therefore paper 1 belongs to the *h*-core of each author. There are larger and more prestigious sets of papers (than the sets of papers of single author), in which paper 1 is also in the *h*-core. A few examples of sets of papers including paper 1 are presented in Table 1. We start from very broad (thus large and demanding) sets (aggregation levels), which have very high *h*, and then narrow them until the *h* of the set is lower than the number of citations of paper 1. Table 1 is based on WoS data, and they are supplemented by Scopus data (given in parentheses) when available.

Paper 1 belongs to the *h*-cores of Lublin and of Abo Akademi University, but does not belong to the *h*-cores of larger sets including these sets. Then the most favorable (most impressive) descriptions of paper 1 by affiliation are:

- significant contribution to the scientific output of Lublin
- significant contribution to the scientific output of Abo Akademi University.

In the RUinh approach we are only looking for the most favorable (most impressive) descriptions, thus the above characterization denies (by definition) inclusion of paper 1 to:

- significant Polish papers
- significant contributions to the scientific output of Abo/Turku.

In view of the dynamic character of assignment of paper to certain *h*-core, the above characterization is valid at specific time point. The dynamic character of the RUinh approach is illustrated in Table 2, in which a time dependency of the number of citations of paper 1 and of the size of *h*-core of Abo Akademi University are reported. Paper 1 has not entered the *h*-core of Abo Akademi University until 2010.

Table 2 illustrates the relative character of the RUinh approach. In 2004, sixty-six was a sufficient number of citations, which made a paper “significant contribution to the scientific output of Abo Akademi University”, but in course of time, the threshold increased to 112. Detailed analysis of particular significant Polish or Finnish papers (Table 1) indicates that among the most cited Polish or Finnish papers, there is a substantial fraction of papers reporting multi-hospital clinical trials or other papers with dozens of authors and affiliations. Only a small fraction of the authors and affiliations (in a particular paper) are Polish or Finnish, and those papers can only formally be considered as Polish or Finnish contributions to science. Perhaps Schreiber (2008)-style fractional counting of papers in the calculation of h in sets of multi-author papers would be a fair solution.

2.2. By discipline

A few examples of sets of papers including paper 1 are presented in Table 3.

Paper 1 belongs to the *h*-core of Thermochimica Acta, but does not belong to the *h*-cores of larger sets including this set. Then the most favorable (most impressive) description of paper 1 by discipline is:

Table 2
The *h*-index of Abo Akademi University, and the number of citations of paper 1 as a function of time.

Year	*h* of Abo Akademi University	No. of citations of paper 1
2004	66	3
2005	70	14
2006	73	20
2007	79	38
2008	84	53
2009	88	69
2010	**95**	97
2011	**103**	117
Feb. 2013	**112**	134

The range, where paper 1 is in the *h*-core of Abo Akademi University is in boldface.

Table 3
The sizes of *h*-cores of sets of papers including Thermochim.Acta 412:47, by discipline.

Level	Description	No. of papers	*h*			
Research area	Chemistry	4,469,683 (3,596,872)	998 (797)			
	Description	No. of papers	*h*	Description	No. of papers	*h*
WoS category	Chemistry, physical	1,038,705	627	Chemistry, analytical	581,886	478
	Description	No. of papers	*h*			
Journal	Thermochimica Acta	16,449 (16,826)	88 (87)			

- significant contribution to Thermochimica Acta.

The above characterization denies (by definition) inclusion of paper 1 to:

- significant papers in physical chemistry
- significant papers in analytical chemistry.

The RUinh approach gives fair credit to papers from disciplines with a few scientists, a few publications, and a few citations per paper. Direct comparison of citation numbers of such articles with papers from disciplines with many scientists, many publications, and many citations per paper, puts the former at inherent disadvantage. In the RUinh approach the threshold depends on the branch and sub-branch of science. Table 3 indicates that a paper with 500 citations is a significant contribution to analytical chemistry, but another paper with 500 citations is not a significant contribution to physical chemistry.

2.3. Combined categories

Larger and more demanding sets including paper 1 in their *h*-cores than those reported in Tables 1 and 3 can be defined by combination of different affiliation and subject categories, e.g., Finland, chemistry has a *h* index of 129(124), but for sake of simplicity we will limit ourselves to “pure” affiliation and “pure” subject categories.

3. Assessment of scientists/institutions

The authors of paper 1 “inherit” the properties of paper 1 discussed in the previous section, that is, they can be described as:

By affiliation:

- significant scientist from Lublin
- significant scientists from Abo Akademi University

By discipline:

- significant contributors to Thermochimica Acta

The above characterization of the authors does not deny their inclusion to “higher” (more prestigious) classes, because those authors have published papers, which have more citations than paper 1. Also the scientific institutions (Department of Electrochemistry of Technical University of Lublin, and Department of Physical Chemistry of Abo Akademi University) “inherit” the properties of paper 1 discussed in the previous section, that is, they can be described as:

By affiliation:

- significant department of Abo Akademi University
- significant scientific institution of Lublin.

By discipline:

- significant contributors to Thermochimica Acta.

The above characterization of the departments does not deny their inclusion to “higher” (more prestigious) classes, for the same reasons as discussed above for authors.

The assessment of an author or an institution based upon single paper is controversial. Therefore, a hierarchical approach discussed in the next section was introduced.

Table 4
Top-*h* scientists in an university and its divisions (WoS).

Rank	Chair R	Faculty	University
1	16 P	31 P	31 P
2	15 P	29 P	29 P
3	13 H	28 P	28 P
4	7 H	27 P	27 P
5	7 H	21 P	27 P
6	**7 H**	21 H	23 P
7	6 H	19 P	23 P
8	...	17 P	22 P
9		17 P	21 P
10		16 P	21 H
11		16 P	19 P
12		16 P	18 H
13		16 P	17 P
14		16 D	17 P
15		**15 P**	17 P
16		15 P	**17 P**
		15 P	16 P
		15 P	16 P

The *h*-index of the last scientist in the *i*-core is in boldface.

4. Hierarchical approach

Hierarchical *h*-type indices (Kosmulski, 2006; Prathap, 2006; Schubert, 2007) can be used in the framework of RUinh. A scientific institution (city, country, discipline) has an *i*-index of *i* when it employs *i* scientists, with *h* indices of at least *i* each, but does not employ *i* + 1 scientists with *h* indices of at least *i* + 1 each. The top *i* scientists constitute the *i*-core, which can be considered as a set of significant scientists of that institution (city, country, discipline). Different approaches (methodologies) can produce various results. Inclusion of recently deceased or retired scientists into the ranking list, inclusion (or exclusion) of papers, which belong to the *h*-core of the author, but not to the aggregation level of interest (e.g., different affiliation or discipline) are among the most controversial issues (Cronin & Meho, 2006).

4.1. By affiliation

Lists of top-*h* scientists at a level of faculty or university can be readily compiled. Table 4 lists the top-*h* scientists of an university in Poland, a chemical faculty of that university and a chair R, which is a sub-division of the chemical faculty. All lists are truncated, that is, low-*h* scientists were ignored. Only living scientists who are current employees of that university were taken into account, that is, deceased and retired scientists, and those who have left the university for other reasons were ignored. Only the papers, which have the name of the university in the "Organization-enhanced" field were taken into account. Several scientists have a higher personal h than the numbers reported in Table 3, because they have published highly cited papers under other affiliations.

The present approach may lead to paradoxical situations: e.g., an outstanding scientist who has moved many times may be a significant contributor to the science worldwide without being a significant contributor to the scientific output of any country, city and even university. Yet such a situation is rather hypothetical, and there are very few such examples in the real world.

Homonyms were manually filtered out in the current analysis (Section 4.1), and the *h*-indices were calculated for actual human beings. The present author knows most scientists subjected to the analysis personally. This should be emphasized that the problem of homonymous names in Poland is less significant than in many other countries due to a specific distribution of the last names. Only 0.6% of the population carry the most popular last name, and 0.3% of the population carry the second most popular last name. Moreover, the spellings of the male and female forms of popular last names are often different.

A referee of this paper called this "very unfair to compare scientists without taking their complete publication and citation record into account", and discussed that "this would strongly disadvantage scientists who have been willing and able to change institutions". This is a semantic nuance to distinguish between top-*h* scientists of certain scientific institution (city, country, etc.) on the one hand, and significant contributors to the scientific output of that institution, etc., on the other. For example my colleague, Prof. Ryszard Kornijów (personal *h* of 15) recently moved from Lublin (city) to National Marine Fishers Research Institute in Gdynia (another city). Under his new affiliation he published two papers which have together 1 self-citation and no independent citations. He belongs to the top-personal-*h* scientists currently working in Gdynia, but his actual contribution to the scientific output of his institute or of Gdynia is not significant.

My other colleague, Prof. Keshra Sangwal (personal *h* of 21) is a physicist, who recently published a few papers in bibliometrics. These papers have together 10 self-citations and no independent citations. This would be controversial to call him a significant information-scientist, although his personal *h* is comparable with that of Prof. Leo Egghe, the Editor of Journal

Table 5
Top-*i* chairs in a chemical faculty (WoS).

Rank	*i*	No. of full professors
1	7	4
2	7	3
3	7	3
4	6	3
5	6	2
6	**6**	1
7	6	1
8	6	0

The *i*-index of the last chair in the second-level *i*-core is in boldface.

of Informetrics. The problem of people publishing in different fields, trans-field citations, etc., is intriguing and important, and its detailed discussion is beyond the scope of the present paper.

The scientific degrees of the scientists are also reported in Table 4. Professor (P) is the highest scientific degree in Poland followed by H (habilitation), and D (Ph.D.). Table 4 shows a high correlation between h and the scientific degree, that is, scientists who have the highest h (at certain aggregation level) have also the highest scientific degrees (with very few exceptions).

The *i*-indices (hierarchical version of the *h*-index) of the chair R, faculty, and university are 6, 15, and 16, respectively. Table 4 shows a few examples of the ambiguity in the assignment to the *h*-core (*i*-core in this case). The scientist ranked 7 in R has *h* equal to *i* of R, but does not belong to the *i*-core or R, because the top *i* scientists in R have $h > 6$. The scientists ranked 15–18 in the faculty have h equal to *i* of the faculty, but only one of them can be assigned to the *i*-core. This can be accomplished either by using additional criteria (e.g. total number of citations) to select one of them or by fractional assignment.

Let us consider *S*, the top-*h* scientist of the chair R ($h = 16$). In the hierarchical RUinh approach *S* is a significant contributor to the scientific output of his faculty ($i = 15$). This description implies that *S* is also a significant contributor to the scientific output of the chair R (which is a sub-unit of the faculty), and that *S* is not a significant contributor to the scientific output of higher aggregation levels (university, city, etc.). Due to the dynamic character of *h* and *i*, the above description of *S* is valid at certain time point only.

The hierarchical RUinh approach can also be applied at a level of a chair. The top-*i* chairs in the discussed above chemical department are listed in Table 5.

The chemical faculty has a (second-level) *i* index of 6 because its 6 chairs have *i*-indices of at least 6, but the chair ranked 7 has an *i* index lower than 7. The (first level) *i*-index of the chair R equals the (second-level) *i*-index of the chemical department, but because of the tied ranks (there are 5 chairs with an *i*-index of 6, but the number of available positions in the 2-nd level *i*-core is limited to 3), its assignment to the 2-nd level *i*-core of the faculty is ambiguous.

Further development opens multiple options. For example the position of the chemical faculty in the university can be defined in terms of its (first level) *i*-index (equal to 15, cf. Table 4) or of its (second level) *i*-index (equal to 6, cf. Table 5). It can be also defined in terms of the (zeroth-level) *i*-indices, that is, by the *h*-indices of individual scientists, as discussed in Section 3.

4.2. *By discipline*

Lists of top-*h* scientists in a few disciplines are available. Chemistry World (2007) published 10 lists of top-*h* living chemists updated between May 2007 and December 2011. The *i*-index of chemistry increased from 76 in 2007 to 86 in 2011, that is, the living chemist ranked 86 had a *h*-index of 86. Thus chemists with a *h*-index of 86 or more are in the *i*-core of chemistry, and they can be considered as significant living chemists in the “hierarchical” RUinh framework. The substantial increase in the *i*-index of chemistry over barely four years illustrates the dynamic character of the RUinh approach (Table 2 and discussion thereof).

The above threshold is very high, and the number of the above-defined significant chemists is much lower than the number of contributors to h of chemistry (Table 1 and discussion in Section 3). Many contributors to *h* of chemistry (defined in Section 3, the ranking list obtained by a mechanical search of WoS database) are deceased, so they have not been taken account in the *i*-index of chemistry based on the Chemistry World (2007) data.

The ranking list of Chemistry World (2007) includes information about sub-fields of chemistry, represented by the top-*h* chemists, but the assignment of different researchers to certain fields is controversial. For example the Nobel laureate Heeger (No. 4 in the list) is a representative of organic chemistry according to Chemistry World (2007), while in WoS only 2 of over 1000 papers by Heeger belong to the WoS category chemistry, organic, and most his papers belong to physics and materials science rather than to chemistry. Formally the *i*-index of different sub-branches of chemistry can be determined from the Chemistry World (2007) compilation, and it equals 66 for physical chemistry.

Palsberg (2013) compiled a ranking list of top-*h*-computer scientists (living and deceased) by means of Google Scholar. The results reported by Palsberg (2013) are not comparable with the results obtained in this study (obtained by means of WoS) or those reported by Chemistry World (2007).

5. Further research

The above procedures can be applied to independent citations. Yet exclusion of self citations requires extra work.

References

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

Chemistry World. (2007). *Hirsch Index Ranks Top Chemists*. http://www.rsc.org/chemistryworld/news/2007/april/23040701.asp

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science Technology*, *57*(9), 41–46.

Csajbok, E., Berhidi, A., Vasas, L., & Schubert, A. (2007). Hirsch-index for countries based on essential science indicators data. *Scientometrics*, *73*(1), 91–117.

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

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.

Kosmulski, M. (2006). I – wskaźnik bibliometryczny (I – a bibliometric indicator, in Polish). *Forum Akademickie*, *11*, 31.

Palsberg, J. (2013). *The h Index for Computer Science*. http://www.cs.ucla.edu/~palsberg/h-number.html

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

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

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

Torres-Salinas, D., Moreno-Torres, J. G., Delgado-Lopez-Cozar, E., & Herrera, F. (2011). A methodology for institution-field ranking based on a bidimensional analysis: the IFQ^{2}A index. *Scientometrics*, *88*, 771–786.