Journal of Informetrics 5 (2011) 64–74

Contents lists available at ScienceDirect

Journal of Informetrics

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

Structured evaluation of the scientific output of academic research groups by recent *h*-based indicators

Fiorenzo Franceschini*, Domenico Maisano

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

ARTICLE INFO

Article history:
Received 29 June 2010
Received in revised form 25 August 2010
Accepted 25 August 2010

Keywords:
Bibliometric positioning
h-Index
h-Spectrum
Successive *h*-indices
Ch-index
Research evaluation
Academic research group
Scientific production

ABSTRACT

Evaluating the scientific output of researchers, research institutions, academic departments and even universities is a challenging issue. To do this, bibliometric indicators are helpful tools, more and more familiar to research and governmental institutions.

This paper proposes a structured method to compare academic research groups within the same discipline, by means of some Hirsch (*h*) based bibliometric indicators. Precisely, five different typologies of indicators are used so as to depict groups' bibliometric positioning within the scientific community. A specific analysis concerning the Italian researchers in the scientific sector of *Production Technology and Manufacturing Systems* is developed. The analysis is supported by empirical data and can be extended to research groups associated to other scientific sectors.

1. Introduction

Evaluating the scientific production of research groups, departments and even universities and research organizations, is not a new issue. For many years, different rankings have been periodically published by many institutions. For example, the doctoral programs' ranking by the National Research Council or several annual universities' rankings, such as those by the US News and World Report, the Shanghai Jiao Tong University, the Center for Science and Technology Studies at the University of Leiden, the German Center for Higher Education Development, and many others [Buela-Casal, Gutiérrez-Martínez, Bermúdez-Sánchez, & Vadillo-Muñoz, 2007; Bach & Llerena, 2007; Kellner & Ponciano, 2008; Lindsey, 1991; Lombardi, Craig, Capaldi, & Gater, 2002; Van den Berghe et al., 1998]. While the inherent limitations of these rankings have been richly discussed [Billaut, Bouyssou, & Vincke, 2010; Bornmann, 2010; Cai Liu, Cheng, & Liu, 2005, 2006; Opthof & Laydesdorff, 2010; Van Raan, 2005], it is interesting to notice that they are becoming more and more demanded.

In general, rankings are based on a number of evaluation criteria, one of which is the scientific production. Traditional approaches for evaluating and comparing the scientific production of researchers are peer-review and bibliometric indicators [Da Luz et al., 2008]. In particular, when large-scale evaluations are performed (e.g. over hundreds or even thousands of scientists), bibliometric indicators seem to be the only practicable instrument [Van Raan, 2000]. This is one of the reasons why they are more and more familiar to the scientific community and governmental organizations of several world countries [Orr, 2004].

The purpose of this paper is proposing a structured comparison of several academic research groups within the same discipline, by some Hirsch (*h*) based bibliometric indicators. We precise that the *h*-index is a relatively recent but very popular

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

1751-1577/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2010.08.003

indicator that synthetically aggregates two important aspects of the output of a scientist: diffusion/impact – represented by the number of citations per paper – and productivity – represented by the number of papers. A scientist has index *h* if *h* of his or her papers have at least *h* citations each and the other papers have $\leq h$ citations each [Hirsch, 2005]. For more on the advantages/disadvantages of *h* and the large number of proposals for new variants and improvements, we refer the reader to the vast literature and extensive reviews [Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Burrell, 2007; Bornmann, Mutz, & Daniel, 2008; Egghe, 2010; Franceschini & Maisano, 2010a; Franceschini & Maisano, 2010c; Glänzel, 2006; Van Raan, 2006a; Rousseau, 2008].

For this comparison to be as much exhaustive as possible, five different typologies of *h*-based indicators are used. They are respectively: the *h*-spectrum [Franceschini & Maisano, 2010b], the *h*-index of a research group (h_{GROUP}) [Moed, 2005], the *successive* *h*-index of a research group (h_2) [Schubert, 2007], the *ch*-index [Ajiferuke & Wolfram, 2010; Franceschini et al., 2010], and the *h*-index of single publications (h_{SINGLE}) [Schubert, 2009].

We remark that analysis is focused on research groups, rather than whole universities. In fact, many universities are quite heterogeneous, containing excellent as well as mediocre research groups, and university assessment often fails to give proper credit to those pockets of excellence. Moreover, ranking of individual research groups is a worthwhile but still quite rare endeavour. For example, a comparative study of two academic departments by bibliometric indicators was published almost twenty years ago by Zachos (1991). A more extensive analysis, aimed at investigating statistical properties of bibliometric indicators in general, was conducted by Van Raan (2006b), considering several chemistry and medical research groups. Another study, directed at comparing homologous departments of different universities, has been recently proposed by Lazaridis (2010). In this study, four typologies of department over six Greek universities are considered and ranked using the *h*-index mean value of academic department members.

Our analysis concerns some Italian academic research groups. In the Italian university system, academic staff is divided in fourteen scientific areas, depending on the discipline (i.e. Mathematical Sciences, Physical Sciences, Biological Sciences, Industrial and Information Engineering, etc.). Scientific areas are in turn divided in a number of sectors concerning more specific subjects of interest [MIUR, 2010]. The attention of this work is focused on the sector of *Production Technology and Manufacturing Systems*. Given a specific university, we define as a (local) research group the whole of the researchers involved in one scientific sector. The five *h*-based indicators are used for delineating the "bibliometric positioning" of the local research groups investigated. Of course, the number of members of each research group will change depending on the size of each University. The choice of limiting the analysis to a specific scientific sector is due to the fact that it is familiar to the authors and, therefore, it is easier to manage input data (citations statistics) and results. The same procedure could be easily extended to other scientific sectors at national or even international level. We remark that the analysis of two or more scientific sectors would require a normalization to allow comparisons among the results in the different fields, owing to the differences in *h*-index levels.

The remaining of this paper is organized into three sections. Section 2 illustrates the methodology, with particular attention to the description of bibliometric indicators; Section 3 presents and discusses the analysis results; Section 4 examines pros and cons of the proposed indicators and introduces the concept of bibliometric profile, a graphical tool to support comparison among research groups. Finally, the conclusions are given, summarising the original contribution of the paper.

2. Methodology

Analysis is carried out following a *bottom-up* approach, according to the definition of Noyons and Van Raan (2002). Firstly, researchers associated to the examined research groups are identified by an institutional database [MIUR, 2010]. Then, for each researcher are determined: the number of (1) scientific publications, (2) citations and (3) citers – i.e. different (co-) authors of the citing paper(s) associated to each citation. The information about citers is important, not to establish self-citation patterns, but to construct an indicator (i.e. the *ch*-index), which evaluates the diffusion/impact of a publication on the basis of the portion of members of the scientific community that are interested in it. Citation statistics are collected using the Google Scholar (GS) search engine. It was decided to use this database because of the greater coverage and since it can be automatically queried through dedicated software applications, such as Publish or Perish or other *ad hoc* applications [Bar-Ilan, 2010; Harzing & van der Wal, 2008]. While determining publications and citations is quite easy and fast, on the other hand, the determination of citers is much more complicated. This procedure was completely automated by an *ad hoc* software application able to query GS automatically. For more information we refer the reader to [Franceschini et al., 2010].

It is worth mentioning that in the Web of Science, the number of citers to a given scientist can be obtained by few clicks, by using the Analyze Results feature. Unfortunately, search is limited to citing papers published on journals/proceedings that are listed by Thomson Scientific. Also, according to many bibliometrists, coverage of Web of Science and Scopus databases – in many fields such as Social Sciences, Computer Science or Engineering Science – is not sufficient [Harzing & van der Wal, 2008]. GS, on the other hand, includes citations to books, book chapters, dissertations, working papers, conference papers, and journal articles published in non-ISI and Open-Access journals. Hence, it probably provides a more comprehensive picture of recent impact. This is one of the reasons why GS has been preferred to other databases for our analysis.

On the other hand, it should not be forgotten that GS is generally less accurate than Web of Science or Scopus. Database mistakes – such as false references, duplications, author ambiguities, etc. – are not so infrequent in GS [Franceschini & Maisano, in pressb; Jacso, 2006]. This is probably due to the automatic generation of the GS data set by scanning and parsing

Table 1
Analysis dimensions and corresponding indicators used to represent them.

Analysis dimension	Indicator(s)
1. Average and variability of the productivity/impact of one research group's members	*h*-Spectrum
2. Overall impact of the publications of each group on the scientific community	*h*-Index of a research group (h_{GROUP})
3. Size of the most productive core of researchers within each group	Group's successive *h*-index (h_2)
4. Portion of the scientific community "infected" by the publications of one research group	*ch*-Index and other derived indicators
5. Citation reputation of the "masterpiece" paper of each research group	*h*-Index of a single publication (h_{SINGLE})

PDF files, to extract reference lists. To reduce database inaccuracies and to achieve useful and correct results, data preparation and data cleaning are therefore very important.

One of the specific problems encountered during data collection is represented by homonym scientists/authors. In general, authors with common names or identified by full surname and first name(s)' initial(s) – rather than full first name(s) – are subject to this kind of problem. The practical effect is that contributions of different homonym authors are erroneously added up, with the result of distorting and "inflating" bibliometric indicators. Luckily, these "suspected" authors have been detected quite easily, and their citation statistics have been manually revised.

For the analysis to be as much exhaustive as possible, five analysis dimensions are identified (summarised in Table 1).

Here is the list of these dimensions, introducing and describing in detail the indicators selected for representing them:

Dimension 1: average and variability of the productivity/impact of one research group's members.

Indicator: *h-spectrum*, defined as the distribution representing the *h* values associated to a group of researchers. *h*-spectrum gives a "snapshot" of the population of a specific research group. In some past works [Franceschini & Maisano, 2010b, in pressa] *h*-spectrum was used to compare scientific journals on the basis of the *h* values associated to their (co-)authors. We can distinguish between *local* *h*-spectra – i.e. those related to researchers of the same group – and a *global* national *h*-spectrum, constructed considering the *h* values of all the researchers at national level. The *h*-spectrum *average value* ($\bar{h}$) is used as a synthetic indicator to perform quick comparisons among local *h*-spectra.

Dimension 2: overall impact of the publications of each group on the scientific community.

Indicator: the *h-index of a research group* (h_{GROUP}), that is to say the *h* of the union of the publications associated to the researchers of the same research group. Thus, a local research group has index h_{GROUP} if h_{GROUP} of the publications associated to the group's researchers have at least h_{GROUP} citations each and the other publications have $\leq h_{GROUP}$ citations each. h_{GROUP} depicts the impact of a research group on the scientific community but favours large groups – in terms of staff number (N) – since they are likely to have more publications than the others.

To make h_{GROUP} values comparable and obtain an indication on the average performance of the research group, complementary to the one provided by $\bar{h}$, a normalization has to be introduced. A possible way is to multiply the h_{GROUP} values by the inverse of the square root of the group size ($\sqrt{N}$). This normalization is quite consistent with other models in the literature, in which the relationship between h_{GROUP} and N is governed by the power law $h_{GROUP} \propto N^{\beta}$, with exponent β around 0.4–0.5 [Batista, Campiteli, Kinouchi, & Martinez, 2006; Molinari & Molinari, 2008a, 2008b]. Here follows a simplified justification of this normalization. Assuming that a group consists of researchers with the same *h*-value and considering the empirical relationship observed by Hirsch – i.e. on the average, the total number of citations (C) is approximately proportional to h^2 [Hirsch, 2005]:

$$C \propto h^2 \tag{1}$$

it can be said that:

$$h^2_{GROUP} \propto C_{GROUP} = \sum_{i=1}^{N} C_{(i)} = C_{(1)} + C_{(2)} + \ldots + C_{(N)} \tag{2}$$

being C_{GROUP} the total number of citations received by the group's publications and $C_{(i)}$ is the total number of citations received by the publications of the *i*th group member. In this case:

$$h^2_{GROUP} \propto h^2_{(1)} + h^2_{(2)} + \ldots + h^2_{(N)} \tag{3}$$

and finally:

$$h_{GROUP} \propto \sqrt{h^2_{(1)} + h^2_{(2)} + \ldots + h^2_{(N)}} \rightarrow \sqrt{N}h \tag{4}$$

So, in this specific case, h_{GROUP} is proportional to the group members' *h*-indices by a factor $\sqrt{N}$. The normalized e. $h_{GROUP,norm} = h_{GROUP}/\sqrt{N}$) is therefore insensitive to N. Of course, this result is based on the fact that $h_{(1)} \approx h_{(2)} \approx \ldots \approx h_{(N)} \approx h$ and that there is no overlap between the publications of the (N) members of the same group.

Dimension 3: size of the most productive core of researchers within each group.

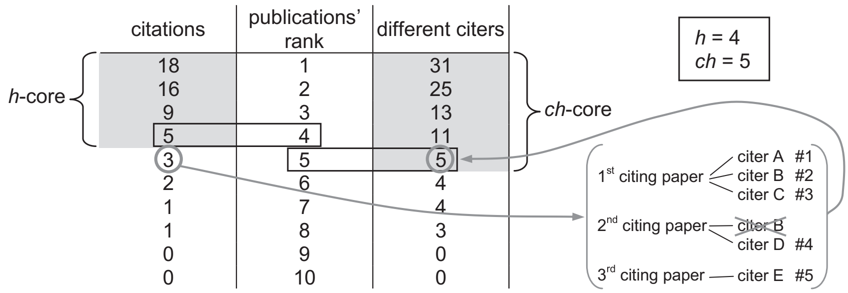

Fig. 1. Example of calculation of *h* and *ch* using the same (fictitious) input data. The inset on the right side of the figure shows the count of the different citing (co-)authors for the 5th publication. Publications are ordered decreasingly with respect to the number of received citations. For simplicity, this ranking is assumed to coincide with the one obtained ordering the publications decreasingly with respect to the number of different citers. Adapted from Franceschini et al. (2010).

Indicator: the *group's successive h-index* (h_2). h_2 is defined by Prathap (2006) and Schubert (2007) in this way: a research group has index h_2 if it has h_2 members with an *h*-index of at least h_2 [Ruanea & Tol, 2007]. h_2 indicates the portion of members that "keep the show going" for one group.

Dimension 4: portion of the scientific community "infected" by the publications of one research group, according to the Goffman's (1966) epidemic model of information transmission.

Indicator: the *ch-index of each researcher*. *ch*, closely akin to *h*, is defined as the number such that, for a general group of scientific publications, *ch* publications are cited by at least *ch* different citers, while the other publications are cited by no more than *ch* different citers [Ajiferuke & Wolfram, 2010; Franceschini et al., 2010]. According to the definition of *ch*, if the same citing author cites a publication more than one time, then he has to be counted only once. *Ch*-index can be calculated similarly to *h*, that is to say by ranking the publications in decreasing order with respect to the number of different citers and identifying the break-even point between publications' rank and number of citers (see the example in Fig. 1).

The most important benefit of *ch* is to be almost insensitive to self-citations and/or citations made by recurrent citers. Furthermore, *ch* gives credit to all the co-authors of a citing paper (for instance, not just the first one). In disciplines where *hyperauthorship* is an issue – e.g. biomedicine, where it can be found citing papers with literally hundreds of authors – this may seriously inflate this indicator [Cronin, 2001; Franceschini et al., 2010]. Fortunately, in the chosen sample area we have not found any case of hyperauthorship and therefore our empirical analysis is uninfluenced by this effect. This indicator can be used to construct other indicators (*ch*-spectrum, ch_{GROUP} and ch_2) analogous to the three above mentioned, but based on *ch* instead of *h*.

Dimension 5: citation reputation of the "masterpiece" paper of each research group.

Indicator: the *h-index of a single publication* (h_{SINGLE}), which is defined as the *h*-index of the set of papers citing it. In other terms, not more than *h* of the papers citing the publication of interest should receive not less than *h* citations. h_{SINGLE} takes into account the direct citation influence of a publication – i.e. the amount of citations received – and the indirect one – i.e. the presence of the publication in the reference lists of other highly cited publications [Schubert, 2009]. Since h_{SINGLE} has been devised to evaluate individual highly cited publications, in this specific case it is calculated considering the most-cited paper of each research group.

The previous five indicators are used to provide a quite complete bibliometric positioning of local research groups and are calculated taking into account the publications/citations accumulated up to the moment of the analysis (April 2010).

In Italy, the total researchers in the scientific sector of *Production Technology and Manufacturing Systems* are 173, affiliated to 33 universities (see Table 2). It can be seen that research groups have a very variable staff number (*N*), i.e. from 1 to 21 members. The average *N* value is 5.2, while the median is 3.We remark that the use of robust indicators, like *h* and the *h*-based ones, is a form of protection from distortions consequent upon database errors [Henzinger, Sunol, & Weber, 2010]. In addition, data accuracy has been improved by manually checking and correcting citations statistics related to each researcher.

3. Analysis results

Citation statistics resulting from our analysis are synthesised in Table 3 and deeply discussed in the five following sub-sections.

3.1. h-Spectrum

Fig. 2 shows the (global) *h*-spectrum related to the 173 researchers. As expected, the distribution is right-skewed [Franceschini & Maisano, 2010b]. Average value is $\bar{h} = 4.8$, median is $h_{\mathrm{MEDIAN}} = 4$ and standard deviation is $s = 3.1$. *h*-Indices related to the individual members of each group are reported in Table 4.

Table 2
University affiliation, name abbreviation and staff number (N) of the Italian research groups in the scientific sector of *Production Technology and Manufacturing Systems*. Universities are sorted in alphabetical order according to their name.

University	Abbrev	N	University	Abbrev.	N
Basilicata	U1	1	Palermo	U18	15
Bergamo	U2	5	Parma	U19	1
Bologna	U3	4	Perugia	U20	1
Brescia	U4	5	Pisa	U21	11
Cagliari	U5	3	Politecnica delle Marche	U22	3
Cassino	U6	4	Politecnico di Bari	U23	11
Catania	U7	3	Politecnico di Milano	U24	13
Della Calabria	U8	3	Politecnico di Torino	U25	21
Ferrara	U9	1	Roma “La Sapienza”	U26	3
Firenze	U10	2	Roma “Tor Vergata”	U27	5
Genova	U11	7	Salento	U28	6
L’Aquila	U12	3	Salerno	U29	7
Libera Università di Bolzano	U13	3	Seconda Univ. Napoli	U30	1
Messina	U14	1	Trento	U31	2
Modena e Reggio Emilia	U15	3	Trieste	U32	2
Napoli “Federico II”	U16	14	Udine	U33	2
Padova	U17	7			
Total staff number					173

Table 3
Synthesis of analysis results. For each academic research group the following indicators are reported: staff number (N), number of papers with at least one citation (P), total number of citations (C), citations of the group’s most-cited publication ($C_{\text{most-cited}}$), average number of co-authors (*avg co-authors*), h-index average value ($\bar{h}$), h-index median value (h_{MEDIAN}), h-index standard deviation (s), h-index of the group (h_{GROUP}), $h_{\text{GROUP,norm}}$, group’s successive h-index (h_2), ch-index average value ($\overline{ch}$), ch-index of the group (ch_{GROUP}), group’s successive ch-index (ch_2), and h-index of the group’s most-cited publication (h_{SINGLE}). Research groups are sorted alphabetically according to the university abbreviation (Table 2).

Univ.	N	P	C	$C_{\text{most-cited}}$	*Avg co-authors*	h-spectrum			h_{GROUP}		h_2	ch-index		ch_2	h_{SINGLE}
						$\bar{h}$	h_{MEDIAN}	s	h_{GROUP}	$h_{\text{GROUP.norm}}$		($\overline{ch}$)	ch_{GROUP}		
U1	1	11	51	13	4.5	5.0	5.0	–	5	5.0	1	5.0	5	1	2
U2	5	77	303	25	3.0	3.8	2.0	2.5	7	3.1	2	5.2	11	3	3
U3	4	82	360	25	3.5	4.8	4.5	1.0	7	3.5	4	7.8	11	4	4
U4	5	80	611	86	2.6	4.2	5.0	3.2	12	5.4	3	5.6	16	3	21
U5	3	33	185	27	3.7	3.3	3.0	2.5	6	3.5	2	4.7	8	2	8
U6	4	81	411	27	2.9	5.8	5.0	2.2	9	4.5	4	7.8	11	4	5
U7	3	43	260	25	2.3	4.3	6.0	3.8	6	3.5	2	5.0	8	2	7
U8	3	125	768	46	3.8	8.3	9.0	1.2	11	6.4	3	12.3	16	3	8
U9	1	5	11	5	2.8	2.0	2.0	–	2	2.0	1	3.0	3	1	1
U10	2	10	32	6	1.1	2.0	2.0	2.8	4	2.8	1	2.5	5	1	2
U11	7	84	553	48	2.6	3.4	2.0	3.5	12	4.5	3	5.1	16	3	6
U12	3	34	209	37	3.3	4.3	3.0	3.2	8	4.6	2	5.0	9	3	12
U13	3	10	71	12	1.1	2.0	0.0	3.5	6	3.5	1	3.0	9	1	6
U14	1	13	25	7	3.7	2.0	2.0	–	2	2.0	1	5.0	5	1	3
U15	3	41	300	40	3.7	4.3	2.0	4.0	9	5.2	2	7.0	13	3	7
U16	14	358	2557	65	3.3	6.8	6.5	3.9	19	5.1	7	9.1	30	7	15
U17	7	142	734	55	3.5	4.9	5.0	2.0	11	4.2	4	7.9	18	6	7
U18	15	439	2511	65	3.4	5.8	6.0	3.0	19	4.9	6	8.1	28	7	7
U19	1	36	199	42	3.2	7.0	7.0	–	7	7.0	1	9.0	9	1	8
U20	1	25	428	131	3.2	10.0	10.0	–	10	10.0	1	13.0	13	1	20
U21	11	146	1434	86	2.9	4.8	5.0	3.7	18	5.4	5	6.5	24	6	14
U22	3	49	356	39	3.2	6.3	8.0	2.9	9	5.2	3	8.7	13	3	9
U23	11	194	940	33	3.2	4.5	4.0	2.3	12	3.6	5	6.7	20	6	8
U24	13	308	1851	157	2.8	5.2	5.0	3.1	16	4.4	5	7.2	22	7	24
U25	21	408	3019	144	2.5	5.0	5.0	3.8	21	4.6	7	6.4	28	8	23
U26	3	21	102	34	2.9	2.3	2.0	1.5	4	2.3	2	3.0	6	2	11
U27	5	154	767	62	3.4	5.6	4.0	4.0	12	5.4	3	8.4	19	4	15
U28	6	84	521	47	3.8	4.8	5.0	2.6	9	3.7	4	6.0	13	4	5
U29	7	61	271	15	3.3	3.1	3.0	1.2	8	3.0	3	3.6	11	3	2
U30	1	1	1	0	1.0	1.0	1.0	–	1	1.0	1	1.0	1	1	0
U31	2	54	231	23	3.8	5.0	5.0	4.2	8	5.7	2	6.5	12	2	4
U32	2	6	45	14	4.1	2.5	2.5	2.1	5	3.5	1	3.0	6	1	3
U33	2	38	146	27	2.3	4.5	4.5	2.1	7	4.9	2	5.5	8	2	3
Mean	5.2	98.6	614.0	44.5	3.0	4.5	4.3	2.8	9.2	4.3	2.8	6.2	12.9	3.2	8.3
Total	173	3253	20263	–	–	–	–	–	–	–	–	–	–	–	–

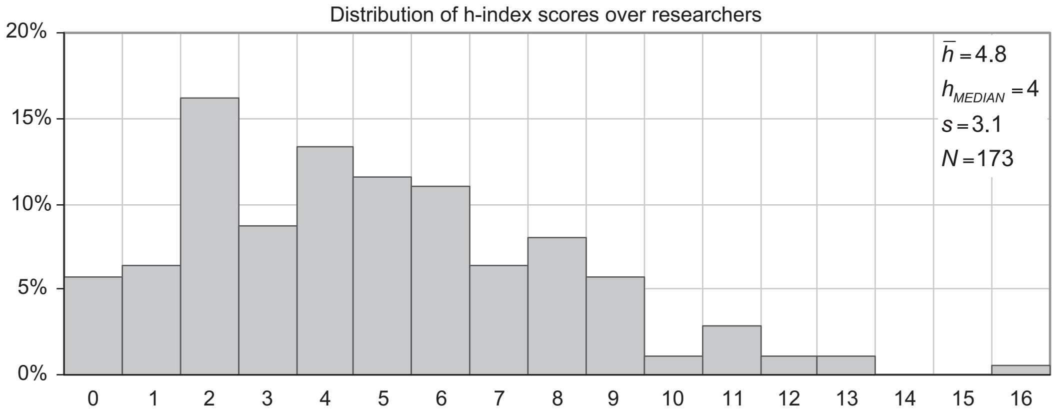

Fig. 2. Global h-spectrum (i.e. distribution of h-index scores) of the 173 researchers of interest.

Global h-spectrum may represent a national reference for individual researchers within the area of interest. For example, a researcher with $h = 3$ will fall on the 37th percentile. Analogous (local) h-spectra can be constructed for each of the 33 research groups.

Consistently with Lazaridis (2010), $\bar{h}$ is used as a synthetic indicator to perform quick evaluations and comparisons among the local h-spectra, even if – from a conceptual point of view – it would be more correct to use h_{MEDIAN} [Fra]. The reason is that h is defined on an ordinal scale with only equivalence and ordinal properties and h-spectra are highly skewed [Bornmann et al., 2008; Franceschini et al., 2010]. Unfortunately, the fact that h_{MEDIAN} is insensitive to extreme values may give results that are not well representative of the research group's average performance. This is particularly evident for small sized research groups. For example, the group of U2 consists of 5 scientists with h-values of 2, 2, 2, 6 and 7. In this case, $\bar{h} = 3.8$ is quite twice as large as $h_{\text{MEDIAN}} = 2$. Since, a significant portion of the research groups are small sized, we decided to use both $\bar{h}$ and h_{MEDIAN}.

3.2. h_{GROUP} index

h_{GROUP} gives an indication of the impact of a research group on the scientific community. Of course, large groups are favoured, since they generally have a larger number of publications. For example, U25 is the group with both the highest h_{GROUP} and N value. Thus, this indicator cannot be used to make direct comparisons among groups with different size. Another problem is that h_{GROUP} can be dominated by the contribution of one very productive group member. This is particularly evident when there is a great difference between the researcher with the highest h and the remaining ones [Lazaridis, 2010]. In our specific case, this condition does not frequently occur since researchers of the same group have not very dissimilar h-values (see Table 4).

The approximate proportionality between h_{GROUP} and $\sqrt{N}$, at the basis of the normalization for obtaining $h_{\text{GROUP,norm}}$, is empirically confirmed by the graph in Fig. 3.

Table 4
h-Indices related to the individual members of each research group. Groups are sorted alphabetically according to the university abbreviation (Table 2 h-indices are sorted in ascending order).

Univ.	N	h-Indices of staff members	Univ.	N	h-indices of staff members
U1	1	5	U18	15	2, 3, 3, 3, 4, 5, 5, 6, 6, 6, 6, 6, 6, 11, 13
U2	5	2, 2, 2, 6, 7	U19	1	7
U3	4	4, 4, 5, 6	U20	1	10
U4	5	0, 2, 5, 6, 8	U21	11	1, 1, 1, 2, 4, 5, 5, 5, 7, 11, 11
U5	3	1, 3, 6	U22	3	3, 8, 8
U6	4	4, 5, 5, 9	U23	11	1, 2, 3, 4, 4, 4, 5, 5, 5, 8, 9
U7	3	0, 6, 7	U24	13	0, 1, 4, 4, 4, 4, 5, 5, 6, 6, 8, 8, 12
U8	3	7, 9, 9	U25	21	0, 0, 1, 1, 2, 2, 2, 2, 3, 4, 5, 5, 6, 7, 7, 8, 9, 9, 9, 11, 13
U9	1	2	U26	3	1, 2, 4
U10	2	0, 4	U27	5	2, 3, 4, 7, 12
U11	7	0, 0, 2, 2, 4, 7, 9	U28	6	2, 2, 4, 6, 7, 8
U12	3	2, 3, 8	U29	7	2, 2, 2, 3, 4, 4, 5
U13	3	0, 0, 6	U30	1	1
U14	1	2	U31	2	2, 8
U15	3	2, 2, 9	U32	2	1, 4
U16	14	2, 2, 3, 4, 4, 5, 6, 7, 8, 8, 9, 10, 11, 16	U33	2	3, 6
U17	7	3, 3, 3, 5, 6, 6, 8			

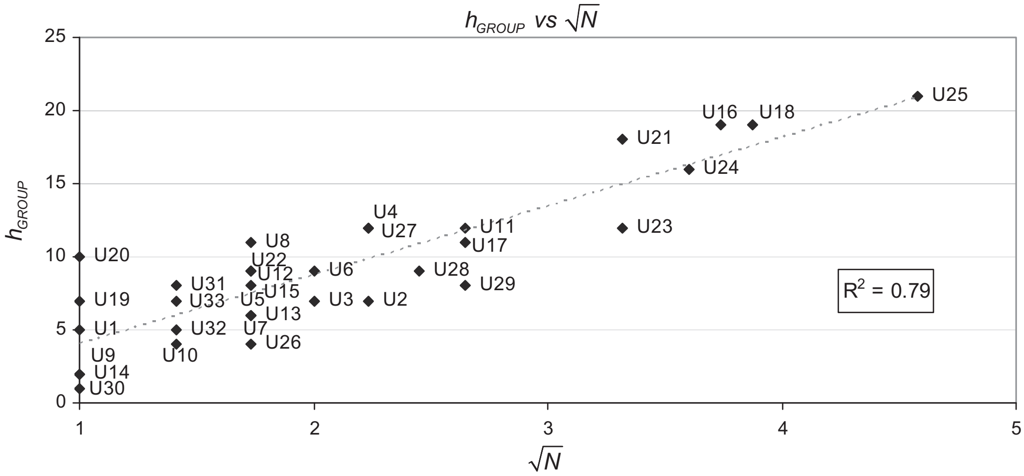

Fig. 3. Relationship between h_{GROUP} and $\sqrt{N}$, using the empirical data related to the 33 research groups investigated.

The advantage of $h_{\text{GROUP,norm}}$ with respect to $\bar{h}$ is that it cannot be inflated by the co-authorship among members of the same research group. For example, in case of systematic co-authorship, the h-indices of the individual researchers would artificially increase, with a resulting increase in $\bar{h}$. However, it can be seen that in our analysis the bibliometric positioning according to $h_{\text{GROUP,norm}}$ is not so different from that one according to $\bar{h}$ (see Fig. 4). This is probably due to the relatively homogeneous distribution of co-authorship within research groups. Also, there is not any "critical mass" effect, meaning that big research groups do not necessarily perform better than small ones [Moed, 2006].

3.3. h_2 index

The h_2 values are reported in Table 3. Two problems can arise with this indicator: (1) it is influenced by N and (2) it is low discerning when N values are quite small. In fact, the synthesis provided by this indicator becomes relevant when the number of the group members and the corresponding h-values have roughly the same order of magnitude, so – despite their different nature – they can be compared [Franceschini & Maisano, 2010a].

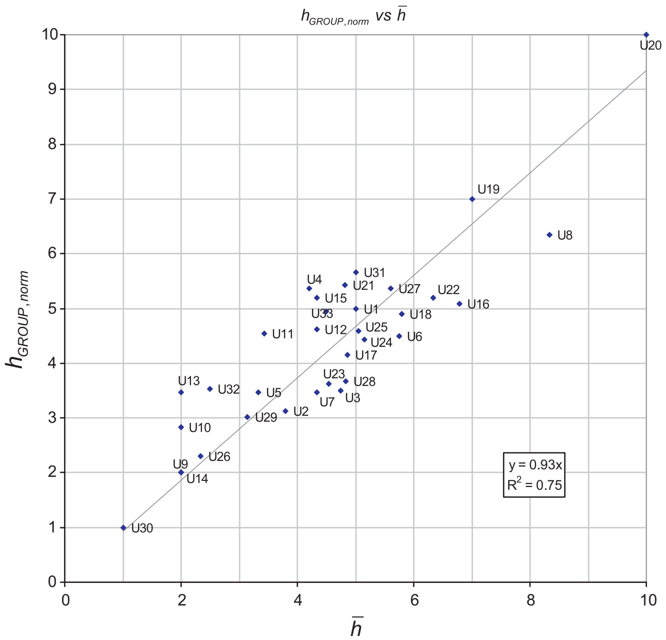

Fig. 4. Relationship between the $h_{\text{GROUP,norm}}$ and $\bar{h}$. Differences in the relevant rankings are not very large ($R^2 = 0.75$).

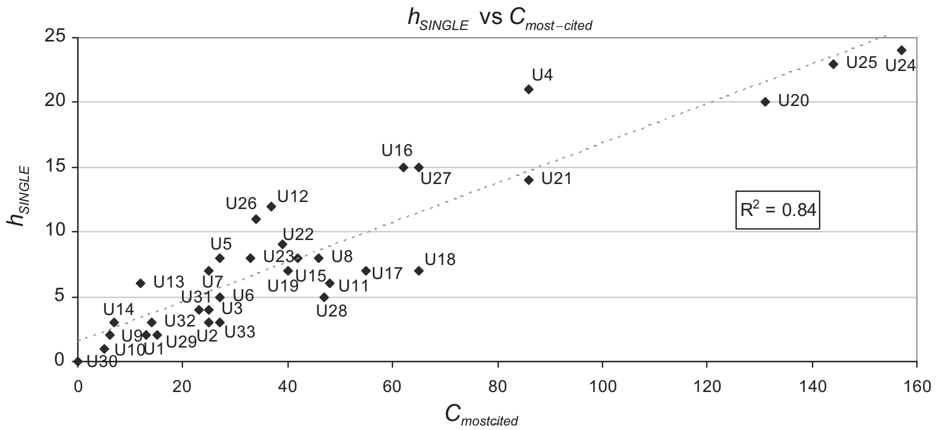

Fig. 5. Relationship between the *h*-index for single publications (h_{SINGLE}), applied to the most-cited paper of each research group, and the corresponding received citations ($C_{most\text{-}cited}$). Correlation is rather strong (R^2 = 0.84) probably because, on average, citing papers associated to the most-cited paper of each research group have a quite homogeneous diffusion.

3.4. *ch-index*

ch-Indices of all the researchers are calculated. *ch*-spectrum, ch_{GROUP} and ch_2 can be constructed considering *ch* instead of *h* as a reference indicator. Results are reported in Table 3. It can be noticed that *ch*-based indicators are not very different from those obtained by the corresponding *h*-based. This is an empirical proof that self-citations and recurrent citers are distributed relatively uniformly over the publications of the different research groups.

3.5. h_{SINGLE} *index*

h_{SINGLE} is representative of the most cited paper of the research group. It was noticed a rather strong empirical correlation between the value of h_{SINGLE} and the number of citations received by the corresponding publication ($C_{most\text{-}cited}$). This means that – on average – citing papers associated to the most cited paper of each research group have a quite homogeneous diffusion in this specific scientific sector (see Fig. 5).

4. Further comments

4.1. *Remarks on the suggested methodology*

The proposed methodology has several strengths:

- structured multidimensional analysis of academic research groups;
- easy comparison among different groups;
- easy implementation and periodic refreshing;
- intuitive meaning (all indicators are *h*-based);
- limited effort for data elaboration;
- global overview of the academic groups' "bibliometric positioning".

Focusing the attention on the proposed indicators, they have different pros and cons, which partially emerged in the previous sections. A synthetic description is reported in Table 5. Being based on *h*-index – these indicators could be subjected to the benefits but also criticisms made to *h*-index itself (e.g. they are insensitive to co-authorship, age of publications, type of publications, journal quality, and – apart from *ch* – they can be influenced by self citations) [Franceschini & Maisano, 2010a]. Also, *h* is not perfectly suitable to compare scholars with different seniority, being in favour of those with long careers. To focus on the impact of recent work and thus on current research performance, the same analysis could be repeated restricting citation period to a 4–5 years window instead of "life time counts".

4.2. *Bibliometric profiles*

As highlighted, bibliometric comparison among the performances of different research groups is a complex issue. A useful supporting tool for this purpose can be a graphical representation by bibliometric profiles. Each profile, associated to each research group, is constructed on the basis of the values of the proposed indicators. Fig. 6 shows the comparison among

Table 5
Summary of the major pros and cons of the proposed indicators.

Indicator	Pros	Cons
h-spectrum ($\bar{h}$) [Franceschini & Maisano, 2010b]	• easy calculation; • it can be used to construct local and global references for individual researchers within the area of interest.	• $\bar{h}$ can be inflated by co-authorship among members of the same research group.
h_{GROUP} and $h_{GROUP,norm}$: [Moed, 2005]	• calculation is easy but a bit more complex than for $\bar{h}$, since the publications of the research group's members have to be joined together; • it is insensitive to co-authorship among members of the same research group.	• it favours large research groups, even if this drawback can be attenuated by a normalization.
h_2: [Schubert, 2007]	• calculation is easy; • very immediate meaning.	• it may be not highly discerning when the group size is quite small; • it favours large sized research groups.
ch-based indicators: [Ajiferuke & Wolfram, 2010; Franceschini et al., 2010]	• they are insensitive to self-citations or citations made by recurrent citing authors; • apart from the ease of calculation, they have the same benefits of the corresponding *h*-based indicators.	• complex calculation; • hyper-authorship: they can be inflated by citing papers with a huge number of co-authors.
h_{SINGLE}: [Schubert, 2009]	• for highly cited publications, this indicator may contribute to a more refined picture of their diffusion/impact.	• complex calculation. However, a Web application to calculate h_{SINGLE} by using Google Scholar data can be found in Thor and Bornmann (2010); • in case of publications receiving only a few citations, this indicator has practically no additional value.

some bibliometric profiles (i.e. those related to U6, U16, U23 and U30). It can be seen that the profile of U16 outranks the others. Conversely, the profile of U30 is outranked by the others. Moreover, U16 and U23 profiles look rather conflicting.

This representation might be used:

- to establish a quick benchmarking among research groups, from different analysis perspectives. Graph readability is favoured by the fact that indicators are all adimensional natural numbers;
- to define some reference regions (or stripes) in which "good" profiles are expected to fall within;
- to support the planning of strategies for future improvement.

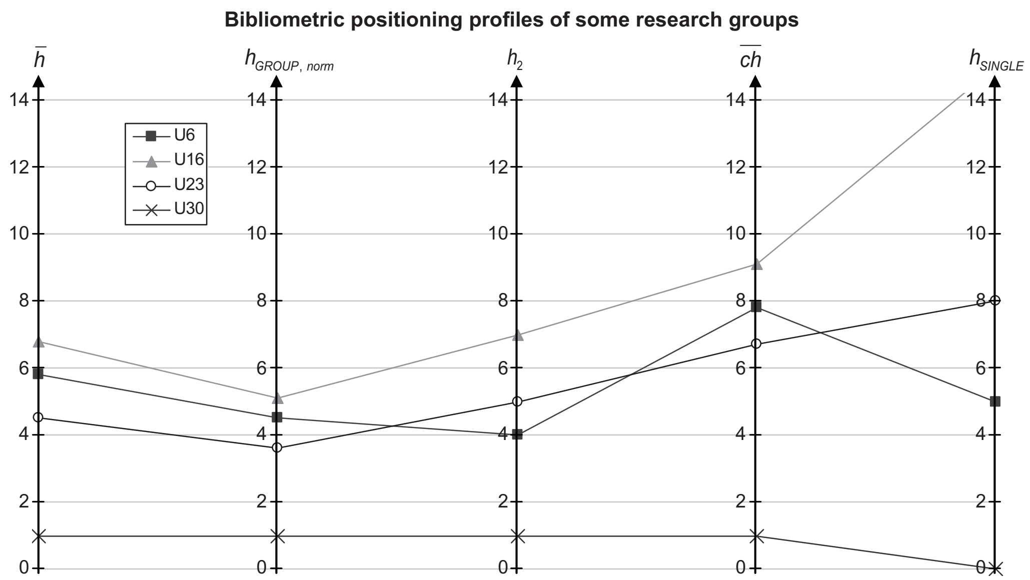

Fig. 6. Bibliometric profiles related to some research groups. In some cases (e.g. U16) h_{SINGLE} value can be significantly higher than the other indicators'.

In this case, $\bar{h}$, $h_{GROUP,norm}$ and *ch* are not so dissimilar, probably because they are all insensitive to *N*. Also, h_2 – which is influenced by *N* – is less discerning than the other indicators, while h_{SINGLE} – being constructed considering just the most cited paper of the group – seems to have a higher dissimilarity with respect to the others.

5. Conclusions

This paper presented a methodology for comparing homologous research groups on the basis of their scientific production. Precisely, the attention was focused on the Italian academic groups within the scientific sector of *Production Technology and Manufacturing Systems*.

Five different *h*-based indicators (i.e. *h*-spectrum, h_{GROUP}, h_2, *ch*, and h_{SINGLE}) were used in order to have a global point of view of the scientific production. Of particular interest is *h*-spectrum, since it is helpful for constructing local and global references for individual researchers within the area of interest. Furthermore, h_{GROUP} gives a more accurate indication of the overall impact of a research group on the scientific community and – by means of a normalization – it allows comparisons among research groups with different size. A graphical representation by bibliometric profiles is furthermore suggested for facilitating the reading of analysis results.

Regarding the future, our analysis will be extended to other scientific sectors at national or even international level. Also, we will evaluate the opportunity of aggregating results related to different scientific sectors into a general ranking, able to reflect the overall scientific production of universities or research institutions.

References

Ajiferuke, I., & Wolfram, D. (2010). Citer analysis as a measure of research impact: Library and information science as a case study. *Scientometrics*, *83*(3), 623–638.

Alonso, S., Cabrerizo, F., Herrera-Viedma, E., & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*(4), 273–289.

Bach, L., & Llerena, P. (2007). Indicators of higher-education institutes and public-research organizations technology transfer activities: Insights from France. *Science and Public Policy*, *34*(10), 709–721.

Bar-Ilan, J. (2010). Citations to the “Introduction to informetrics” indexed by WOS, Scopus and Google Scholar. *Scientometrics*, *82*(3), 495–506.

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

Billaut, J. C., Bouyssou, D., & Vincke, P. (2010). Should you believe in the Shanghai ranking? An MCDM view. *Scientometrics*, *84*(1), 237–263.

Bornmann, L., Mutz, R., & Daniel, H. D. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, *59*(5), 830–837.

Bornmann, L. (2010). Towards the ideal method of measuring research performance: Some comments to the Opthof and Laydesdorff (2010) paper. *Journal of Informetrics*, *4*(3), 441–443.

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M. P., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics*, *71*(3), 349–365.

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

Cai Liu, N., Cheng, Y., & Liu, L. (2005). Academic ranking of world universities using scientometrics—A comment to the “Fatal Attraction”. *Scientometrics*, *64*(1), 101–109.

Cai Liu, N., Cheng, Y., & Liu, L. (2006). A first approach to the classification of the top 500 world universities by their disciplinary characteristics using scientometrics. *Scientometrics*, *68*(1), 135–150.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, *52*(7), 558–569.

Da Luz, M. P, Marques-Portella, C., Mendlowicz, M., Gleiser, S., Coutinho, E. S. F., & Figueira, I. (2008). Institutional h-index: The performance of a new metric in the evaluation of Brazilian Psychiatric Post-graduation Programs. *Scientometrics*, *77*(2), 361–368.

Egghe, L. (2010). The Hirsch-index and related impact measures. *Annual Review of Information Science and Technology (ARIST)*, *44*, 64–114.

Franceschini, F., & Maisano, D. (2010a). Analysis of the Hirsch index’s operational properties. *European Journal of Operational Research*, *203*(2), 494–504. doi:10.1016/j.ejor.2009.08.001

Franceschini, F., & Maisano, D. (2010b). The Hirsch spectrum: A novel tool for analysing scientific journals. *Journal of Informetrics*, *4*(1), 64–73. doi:10.1016/j.joi.2009.08.003

Franceschini, F., & Maisano, D. (2010c). The citation triad: an overview of a scientist’s publication output based on Ferrers diagrams. *Journal of Informetrics*, *4*(4), 503–511.

Franceschini, F., & Maisano, D. (in pressa). Bibliometric positioning of scientific Manufacturing journals: A comparative analysis, *Scientometrics*.

Franceschini, F., & Maisano, D. (in pressb). Influence of database mistakes on journal citation analysis: Remarks on Franceschini and Maisano QREI (2010), *Quality and Reliability Engineering International*.

Franceschini, F., Maisano, D., Perotti, A., & Proto, A. (2010). Analysis of the ch-index, an indicator to evaluate the diffusion of scientific research output by citers. *Scientometrics*, *85*(1), 203–217. doi:10.1007/s11192-010-0165-0

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

Goffman, W. (1966). Mathematical approach to the spread of scientific ideas—The history of mast cell research. *Nature*, *212*, 449–452.

Harzing, A. W., & van der Wal, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*, *8*(11), 61–73.

Henzinger, M., Sunol, J., & Weber, I. (2010). The stability of the h-index. *Scientometrics*, *84*, 465–479.

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.

Jacso, P. (2006). Deflated, inflated and phantom citation counts. *Online Information Review*, *30*(3), 297–309.

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

Lazaridis, T. (2010). Ranking university departments using the mean h-index. *Scientometrics*, *82*(2), 211–216.

Lindsey, D. (1991). The relationship between performance indicators for academic research and funding—Developing a measure of return on investment in science. *Scientometrics*, *20*(1), 221–234.

Lombardi, J. V., Craig, D. D., Capaldi, E. D., & Gater, D. S. (2002). *The myth of number one: Indicators of research university performance—The top American research universities*. , retrieved from http://mup.asu.edu/MythNumberOne.pdf.

MIUR (2010). Ministero dell’Istruzione dell’Università e della Ricerca http://cercauniversita.cineca.it/php5/settori/index.php.

Moed, H. F. (2005). Hirsch index is a creative and appealing construct but be cautious when using it to evaluate individual scholars, retrieved from http://www.cwts.nl/hm/Comments_on_Hirsch_Index_2005_12_16.pdf.
Moed, H. F. (2006). Bibliometric rankings of world universities, CWTS Report 2006-01, available at http://www.cwts.nl/hm/bibl_rnk_wrld_univ_full.pdf.
Molinari, A., & Molinari, J. F. (2008a). A new methodology for ranking scientific institutions. *Scientometrics*, *75*(1), 163–174.
Molinari, A., & Molinari, J. F. (2008b). Mathematical aspects of a new criterion for ranking scientific institutions based on the h-index. *Scientometrics*, *75*(2), 339–356.
Noyons, E., & Van Raan, A. F. J. (2002). In J. Meij (Ed.), *Dealing with the data flood. Mining data, text and multimedia* (pp. 64–72). Beweton: The Hague: STT.
Opthof, T., & Laydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluation of research performance. *Journal of Informetrics*, *4*(3), 423–430.
Orr, D. (2004). Research assessment as an instrument for steering higher education—A comparative study. *Journal of Higher Education Policy and Management*, *26*(3), 345–362.
Prathap, G. (2006). Hirsch-type indices for ranking institutions' scientific research output. *Current Science*, *91*, 1439.
Rousseau, R. (2008). Reflections on recent developments of the h-index and h-type indices. In *Proceedings of WIS 2008, Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting* Berlin, Germany.
Ruanea, F., & Tol, R. S. J. (2007). Rational (successive) h-indices: An application to economics in the Republic of Ireland. *Scientometrics*, *75*(2), 395–405.
Schubert, A. (2007). Successive h-indices. *Scientometrics*, *70*(1), 201–205.
Schubert, A. (2009). Using the h-index for assessing single publications. *Scientometrics*, *78*(3), 559–565.
Thor, A., & Bornmann, L. (2010). Web application to calculate the single publication h index (and further metrics) based on Google Scholar, http://labs.dbs.uni-leipzig.de/gsh/.
Van den Berghe, H., Houben, J. A., de Bruin, R. A., Moed, H. F., Kint, A., Luwel, M., & Spruyt, E. H. J. (1998). Bibliometric indicators of university research performance in Flanders. *Journal of the American Society for Information Science*, *49*(1), 59–67.
Van Raan, A. F. J. (2000). The Pandora's box of citation analysis: measuring scientific excellence, the last evil? In B. Cronin, & H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield, ASIS Monograph Series* (pp. 301–319). Medford, NJ: Information Today Inc.
Van Raan, A. F. J. (2005). Fatal attraction: Ranking of universities by bibliometric methods. *Scientometrics*, *62*(1), 133–143.
Van Raan, A. F. J. (2006a). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.
Van Raan, A. F. J. (2006b). statistical properties of bibliometric indicators: Research group indicator distributions and correlations. *Journal of the American Society for Information Science and Technology*, *57*(3), 408–430.
Zachos, G. (1991). Research output evaluation of two university departments in Greece with the use of bibliometric indicators. *Scientometrics*, *21*(2), 195–221.