Journal of Informetrics 4 (2010) 581–590

Contents lists available at ScienceDirect

Journal of Informetrics

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

A research impact indicator for institutions

E.S. Vieira, J.A.N.F. Gomes*

REQUIMTE/Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

ARTICLE INFO

Article history:
Received 9 October 2009
Received in revised form 21 June 2010
Accepted 22 June 2010

Keywords:
Research impact indicator
Institutional impact
h index
Size-dependence of the *h* index

ABSTRACT

This paper introduces a new impact indicator for the research effort of a university, $^nh_3$. The number of documents or the number of citations obtained by an institution are used frequently in international ranking of institutions. However, these are very dependent on the size and this is inducing mergers with the apparent sole goal of improving the research ranking. The alternative is to use the ratio of the two measures, the mean citation rate, that is size independent but it has been shown to fluctuate along the time as a consequence of its dependence on a very small number of documents with an extremely good citation performance. In the last few years, the popularity of the Hirsch index as an indicator of the research performance of individual researchers led to its application to journals and institutions. However, the original aim of this *h* index of giving a mixed measure of the number of documents published and their impact as measured by the citations collected along the time is totally undesirable for institutions as the overall size may be considered irrelevant for the impact evaluation of research. Furthermore, the *h* index when applied to institutions tends to retain a very small number of documents making all other research production irrelevant for this indicator. The $^nh_3$ index proposed here is designed to measure solely the impact of research in a way that is independent of the size of the institution and is made relatively stable by making a 20-year estimate of the citations of the documents produced in a single year.

1. Introduction

Research intensity of a university is frequently measured by the number of publications or by the number of citations it obtains in a given period. The average number of citations obtained (in a well defined period) by the publications it originates in the same (or in another) period is an indicator of research impact applied frequently. The rankings published by CWTS, SCImago and the Performance Ranking of Scientific Papers for World Universities (Taiwan) use this indicator to measure research impact (CWTS; HEEACT, 2009; SCImago). This indicator is frequently improved to compensate for the variability of the citation culture in different fields by a normalization technique (CWTS; SCImago; Vieira, Nouws, Albergaria, Matos, & Gomes, 2009). Another indicator based on citation performance is the number of highly cited papers. Hirsch proposed a new indicator, now called the “*h* index”, as a particularly simple and useful way to characterize the scientific output of a researcher. A scientist has *h* index if *h* of his or her N_p papers have at least *h* citations each and the other (N_p-*h*) papers have *h* or less citations each (Hirsch, 2005). The scientific community has shown great interest in this indicator as it has the advantage of combining a measure of quantity (number of publications) and impact (number of citations) in a single indicator. This indicator has been used in fields as different as information science (Cronin & Meho, 2006) and physics (Hirsch,

* Corresponding author. Tel.: +35 1220204507.
E-mail address: jfgomes@fc.up.pt (J.A.N.F. Gomes).

1751-1577/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2010.06.006

2005); it has been used for journal assessment (Braun, Glanzel, & Schubert, 2005; Braun, Glanzel, & Schubert, 2006; Schubert & Glanzel, 2007) to differentiate between successful and unsuccessful post-doctoral applications (Bornmann & Daniel, 2005) and in country assessment (Csajbok, Berhidi, Vasas, & Schubert, 2007).

The *h* index is frequently presented as a simple and easy to obtain indicator but it has some limitations (Bornmann & Daniel, 2007; Costas & Bordons, 2007; Egghe, 2006; Jin, Liang, Rousseau, & Egghe, 2007). In order to address these limitations, several changes have been suggested in the literature (Batista, Campiteli, Kinouchi, & Martinez, 2006; Egghe, 2006; Egghe, 2008; Jin, 2006; Jin et al., 2007; Liang, 2006; Molinari & Molinari, 2008a; Molinari & Molinari, 2008b; Prathap, 2006; Rousseau & Ye, 2008; Sidiropoulos, Katsaros, & Manolopoulos, 2007). The h_I proposed by Batista et al. (2006) is calculated as the ratio of the square of *h* index to the total number of authors of the documents in the *h*-core. If all publications in the *h*-core had a single author, then h_I equals *h*. According to the proposers, this has the advantage of being less sensitive to different research fields. To make scientists at different scientific age comparable, Liang (2006) constructed the *h* index sequence by calculating the usual *h* index in a time window of 1, 2 and more years from the current time to the past. The *g* index is intended as an improvement of the *h* index to measure the global citation performance of a set of articles in the *g*-core (Egghe, 2006). It is defined as the largest rank (where papers are arranged in decreasing order of the number of citations received) such that the first *g* papers have (together) at least g^2 citations. Prathap (2006) proposed a first order index, h_1, and a second order index, h_2. To calculate the h_1 index for institutions we rank the publications of a given period in decreasing order of the number of citations received. The institution has a h_1 index if h_1 of the N_p papers have at least h_1 citations each, this being the usual *h* index but for a limited period. The h_2 index is obtained using the *h* index of the researchers of the university. First the *h* index of each researcher is determined and then the researchers are ordered in decreasing order of the *h* index value. The institution has an h_2 index if h_2 of the researchers have at least an h_2 index each. For Prathap, these indices can be used to quantify the scientific performance of an institution and its researchers in a more robust way. The *A* index was developed by Jin (2006) to correct the fact that the original *h* index does not take into account the exact number of citations of articles retained in the *h*-core. This index is simply defined as the average number of citations received by the publications in the *h*-core. Recognizing some limitations of the *A* index, Jin et al. (2007) proposed two new indices, the *R* index and the *AR* index. The *R* index is the square root of the sum of the citations of articles included in the *h*-core. The *AR*, besides taking into account the number of citations, makes use of the age of the publications in the *h*-core. The normalized *h* index (h^n) is defined as the ratio between the *h* index and the number of articles (N_p) (Sidiropoulos et al., 2007). There are studies that defined the fractional *h* index and *g* index in two different ways. One method considers fractional citation counts where the citation count, y, of a *m*-authored paper, is divided by *m*. Another method leaves the citation counts unchanged but replaces the rank by the fractional paper count (Egghe, 2008). The dynamic h-type index (h_d) is a index dependent on the *h*-core, the number of citations received by the documents in the *h*-core and the recent increase on the *h* index. This index is defined as $R(T)$. $v_h(T)$. The $R(T)$ is the *R* index considering the sum of the citations received by articles included in the *h*-core at time T and the v_h is the *h* velocity at time *T*. This index allows the comparison of two researchers that have the same *h* index and the same number of citations on the *h*-core, but for one the *h* index is increasing over the time and for the other the *h* index is constant over the time (Rousseau & Ye, 2008).

The correlation between the *h* index and the total number of citations, the total number of publications, the crown indicator and with peer judgment was investigated (van Raan, 2006). The results of this study showed that both the *h* index and the crown indicator correlate well with peer judgments. Another conclusion was that, for smaller groups in fields with 'less heavy citation traffic', the crown indicator appears to be a more appropriate measure of research performance. The relation of the *h* index with other bibliometric indicators was also analyzed at the micro-level for Spanish CSIC scientists in Natural Resources (Costas & Bordons, 2007). The findings suggest a good correlation, especially with the number of documents and citations received by scientists.

A study of the relationship between the *h* index and three standard bibliometric indicators and the peer assessments using a data set of applicants to the long-term fellowship and young researcher programmes of the European Molecular Biology Organization (EMBO) showed that the correlation between the *h* index and the number of publications and the number of citations are statistically significant. The main finding of this study is that the *h* index can be applied for researcher performance at micro- and meso-levels (Bornmann, Wallon, & Ledin, 2008).

The dependence of an institutional *h* index on the size of the institution as it shows up in the number of papers has been recognized by Molinari and Molinari (2008a) and Molinari and Molinari (2008b) in their attempt to compare research quality of different universities. Efforts have been made to construct a theoretical model of the dependence of the *h* index with other parameters (Glänzel, 2005; Schubert & Glanzel, 2007). The *h* index is found to depend on two fundamental scientometric indicators, the number of documents and the mean citation rate:

$$h = cn^a x^b$$

where *n* is the number of documents, *x* the mean citation rate, $a = 1/\alpha + 1$; $b = \alpha/(\alpha + 1)$; $\alpha = 2$, and *c* a positive constant.

The aim of this work is to develop a new indicator for the research impact production of an institution that is independent of its size and can be easily assessed for relatively short time spans. The theoretical framework of this research is presented in Section 2 below. In the following section, we test the theoretical model above using a very extensive set of data and considering, independently, the number of documents and the mean citation rate. At the end of Section 3, we propose the new indicator based on the Hirsch-index concept with corrections for the size of the institution and for the citation window, the *production corrected and citation projected*, ${}^{n}h_3$ index. The values of this new index for a selection of Brazilian, Portuguese and Spanish universities are calculated and compared with traditional indicators. The stability of the ${}^{n}h_3$ over the years is discussed. The final section presents the major arguments in favour of using this new indicator for institutional research impact.

2. Data and methods

The study is based on the analysis of more than 780 000 documents published in 2004 in journals indexed in the Web of Science (WoS) and classified according to the Essential Science Indicators (ESI) in the fields of Agriculture Sciences, Biology & Biochemistry, Chemistry, Clinical Medicine, Engineering, Material Sciences, Mathematics and Physics. Taken together, these fields are responsible for about 64% of the 2004 total scientific production originated in the 22 fields defined in the ESI. The ESI is a compilation of statistical information related with publications, citations and cites per paper for journals, scientists, institutions and countries referring to 10 years of Thomson Reuters data.

We started from Glänzel theoretical expression (Glänzel, 2005; Schubert & Glanzel, 2007), $h = cn^a x^b$.

At a first stage, we keep relations $a = 1/(\alpha + 1)$ and $b = \alpha/(\alpha + 1)$ and allowed parameter α to vary and searched for the value that leads to the best linear relation between h and the product $n^a x^b$. The same procedure was then repeated for the separate optimization of a and b, now taken as independent parameters. This methodology was applied to the set of documents pertaining to the eight fields mentioned above. To study the dependence of the h index on the number of documents, the following procedure was repeated for each scientific field. Considering the whole set of documents referenced in the WoS for 2004 and published in journals associated with a particular field, we extracted randomly a subset of a given dimension and determined the h index of this subset using the method defined by Hirsch. This procedure was repeated seven times for each field, each time collecting a smaller percentage of the number of the documents in the set, 100%, 50%, 25%, and so on until the number of documents is about 60. This sampling was repeated again seven times to assess the statistical fluctuation of the h index on the random sample taken. In all, we have 49 data points for each of the 8 scientific fields. The optimization to estimate the best α value and then the best a and b values was performed using the Solver of Excel. The ${}^{n}h_3$ was calculated using the optimized values of a for the set of 4, 6 and 26 major public Portuguese, Brazilian and Spanish universities, respectively. These sets of universities were selected to account, taken together, for about 60% of the documents published in 2004 by their respective country. We considered as size reference the average number of documents published by these universities in 2004. In order to study the time stability of the proposed indicator we determined the mean citation rate, the h index, the h_3 index and the ${}^{n}h_3$ for a smaller set of universities in 2000, 2001, 2002, 2003 and 2004.

3. Results and discussion

In order to develop an impact indicator based on the h index concept but with a correction for the dimension of the university (number of publications from the university) and a projected citation window of about 20 years, the topics listed below were studied.

(1) Dependence of the h index with the number of documents.
(2) Temporal evolution of the citations obtained by one document.
(3) Number of documents in the h-core.

3.1. Dependence of the h index on the number of documents

The theoretical model developed for the dependence of the h index on the number of documents and on the mean citation rate was empirically tested for the data set of 2004 documents referenced in the WoS in the fields of Agricultural Sciences, Biology & Biochemistry, Chemistry, Clinical Medicine, Engineering, Material Sciences, Mathematics and Physics.

3.1.1. Optimization of α

In order to test the theoretical model, we first optimized the model using a single parameter, α, and then using the two independent parameters, a and b. As we assume a linear relation between h and $n^a x^b$, we optimized the values of α, and then those of a and b to produce the best regression line for the data set. Each field is assumed to have a characteristic mean citation rate, x, and seven points are constructed with different numbers of documents as described above (see Section 2).

The results of the optimization of α are shown in Table 1, for each field and for the set of all fields. The theoretical value $\alpha = 2$ appears to give a fair description of reality but allowing parameter α to vary will lower the unexplained variance up to more than 50% for some fields. When all fields are taken together, the unexplained variance is relatively high (3.0%) and it is improved by only 6.67% when α is allowed to vary.

Table 1
Optimization of parameter α.

	$(1-R^2)/1$	α_o	$(1-R_o^2)/1$
All fields	0.030	1.871	0.028
Agricultural Sciences	0.028	2.489	0.010
Biology and Biochemistry	0.019	2.418	0.007
Chemistry	0.005	2.194	0.002
Clinical Medicine	0.002	2.033	0.002
Engineering	0.007	2.187	0.004
Material Science	0.004	1.930	0.003
Mathematics	0.014	2.260	0.009
Physics	0.007	2.222	0.004

R^2, coefficient of determination for the regression line when α is 2. R_o^2, coefficient of determination for the regression line with the optimized α_o.

Table 2
Optimized values of a and b.

	$(1-R^2)/1$	a_0	b_0	$(1-R_o^2)/1$
All fields	0.030	0.3446	0.494	0.024

The optimized value obtained for α_0 when we consider the set of all scientific areas is 1.871 ± 0.006, below the value 2 that is considered in the theoretical model. The error bar is estimated by repeating the calculation for some of the seven samples taken at each point.

3.1.2. Optimization of a and b

The result of the independent optimization of the exponents a and b for the set of all scientific fields is shown in Table 2.

The optimized value of the exponent a is close to the theoretical value of 1/3 but the difference is significant as the error bar is rather small, 0.3446 ± 0.0007. The optimized value for the exponent b, 0.494 ± 0.003, is far from the theoretical value of 2/3. Again, the error bars were estimated by repeating the calculations with different random samples taken from the whole set of 2004 data for each field.

3.2. Number of documents in the h-core and temporal evolution of the citations

Another way of measuring institutional performance is based on the use of the h index. There are two problems with the direct use of the h index as an impact indicator. On the one hand, we have discussed above how it varies with the number of documents so that a large institution will rank higher just due to the larger number of publications produced in a certain period. The other problem is related with the rather small number of documents in the h-core so that the institutional h index will depend on the performance of a rather small set of top performing documents and this varies widely from year to year. This situation is aggravated by the fact that scientific fields with higher citation traffic will dominate the h index measure. We suggest below that this should be corrected and propose a new *production and citation projected h index* (${}^nh_3$) to achieve that.

Some studies showed that the percentage of documents in the h-core for single researchers along their careers, is as high as 20% in the most cases (Batista et al., 2006; Hirsch, 2007). If we consider groups of researchers in one specific area, the percentage of documents in the h-core for one institution decreases. These values can vary between 10% and 13% for research groups with about 100 documents published and a citation window of 3 years (van Raan, 2006).

In Table 3 we show the h index for a set of universities, considering all the scientific production referenced in the WoS. We can see that the percentage of documents in the h-core for these universities is very low (it varies from 0.7% to 1.2%). If we calculate the h index, but considering only the scientific production of one particular year, thenumber of documents in

Table 3
Total number of documents, h index and percentage of documents in h-core for the documents referenced in the WoS.

University	Total number of documents	h index	% of documents in the h-core
Universidad de Córdoba	4028	47	1.2
Universidade do Minho	6014	58	1.0
Universidad de Extremadura	7044	75	1.1
Universidad de Vigo	7589	67	0.9
Universidade Nova de Lisboa	8306	83	1.0
Universidad la Laguna	8335	73	0.9
Universidad de Alcalá	8362	74	0.9
Universidade de Aveiro	8396	58	0.7
Universidad de Málaga	8760	75	0.9
Universidad de Valladolid	9705	81	0.8

Table 4
Number of documents, h index and percentage of documents in h-core for the documents referenced in the WoS in 2004 using a 5 years citation window.

University	Number of documents (2004)	h index	% of documents in the h-core
Universidad la Laguna	420	30	7.1
Universidad de Alcalá	421	24	5.7
Universidad de Extremadura	422	24	5.7
Universidad de Córdoba	459	27	5.9
Universidad de Valladolid	491	25	5.1
Universidade do Minho	521	24	4.6
Universidad de Málaga	538	23	4.3
Universidade Nova de Lisboa	581	30	5.2
Universidad de Vigo	624	26	4.2
Universidade de Aveiro	697	29	4.2
Universidade do Porto	1235	39	3.2
Universidade Técnica de Lisboa	1291	43	3.3
Universidad Autónoma de Barcelona	1543	41	2.7
Universidad Autónoma de Madrid	1558	55	3.5
Universidad de Barcelona	2497	57	2.3

the h-core is still small (between 2.3% and 7.1%). These values were influenced by the small citation window (5 years). The results can be observed in Table 4.

In order to attenuate the limitations of the h index we propose a new h_3 index. This is based on the estimation of the h index associated with the same set of documents but for a longer observation window to count citations. An institution has an h_3 index if h_3 of the documents (P) have at least $h_3/3$ citations and the other (P-h_3) documents have $h_3/3$ or less citations each. We consider the P documents published in 1 year by the university or institute and a 5-year citation window. We rank the documents published by the university in that year by the number of citations received and then determine the value of h_3 index applying the definition above.

To clarify further the calculation of the h_3 index we give the following example. In 2004, the “Universidad de Alcalá” published 421 documents and these are ranked in Table 5 in descending order of the number of citations obtained by each document in the 5 years (2004–2008).

Using Hirsch’s definition, the h index is 24. Using our definition above, $h_3 = 51$ as 51 of the 421 published have at least $h_3/3 = 17$ citations and all other 370 documents have 17 or less citations each.

In order to determine the projection of the h_3 index related with the citation window we extracted all the referenced documents in the WoS in 1989 for a set of universities and determined the relation between the total number of citations received by these documents in 5 years and 3 different citation windows. Each point in Fig. 1 represents one university. Here we used the Portuguese and Spanish universities with at least 100 documents referenced in the 1989 WoS.

The results suggest that the number of citations obtained in 10, 15 and 20 years by the set of documents published in a given year is ca. 2, 2.5 and 3 times the number of citations obtained by the same set of documents in 5 years. So we can say that using the h_3 index we are considering a projection of citations of about 20 years.

Fig. 2 shows the relation observed between the h_3 index and the h index for the documents referenced in the WoS in 1989. The h index was obtained using a 20 years citation window.

Table 5
Example of calculation of the h index and the h_3 index for documents published in 2004 by the “Universidad de Alcalá”: $h = 24$ and $h_3 = 51$.

Ranked documents	Total of citations (2004–2008)
1	75
2	63
3	59
...	...
20	25
21	24
22	24
23	24
24	**24**
25	23
...	...
47	18
48	17
49	17
50	17
51	**17**
52	17
...	...

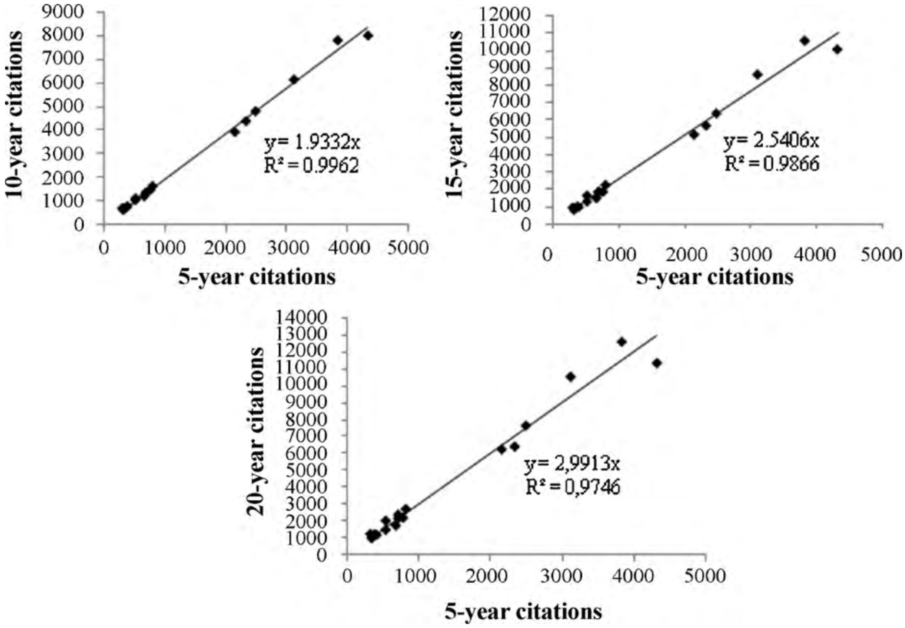

Fig. 1. Relation between the total number of citations in 5 years and three different citations windows (10 years; 15 years and 20 years) for the documents referenced in the WoS in 1989. Each point represents a set of documents referenced in the WoS for one university.

The results in Fig. 2 show that the h_3 index corresponds to a good estimation of the h index of a set of documents published in 1 year with a citation window of 20 years.

To assess the validity of the new index h_3, we studied the relation between the h_3 index and the h index, the relation between the h_3 index and the mean citation rate, and the relation between h_3 index and the number of citations. The mean citation rate is the average number of citations obtained by the documents published in 2004 in the period between 2004 and 2008 divided by the number of documents published in 2004. In Fig. 3, are presented the results obtained using the 2004 WoS documents for Brazilian, Portuguese and Spanish universities that produced ca. 60% of the scientific production of the respective country and a citation window of 5 years (2004–2008).

As expected from the discussion above, the results show a good relation between the h_3 index and the h index obtained for the documents published in a given year and using a citation window of 5 years. The value of h_3 is about twice that of the standard h index well within the theoretical interval $h < h_3 < 3h$. The second plot in Fig. 3 confirms that the h_3 index and the mean citation rate are two rather different measures of research impact. The mean citation rate considers all documents and it may be influenced by a small number of documents with an extremely high number of citations. The h_3 index is immune to these exceptional documents. The third plot suggests that the h_3 index may be considered to be a predictor of the total number of citations however, poor as pointed out by the low coefficient of the determination.

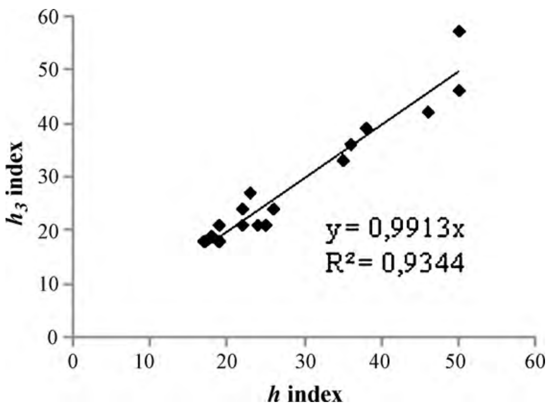

Fig. 2. Relation between h_3 index obtained for the documents referenced in the WoS in 1989 and 5 years citation window and the h index determined for the same set of documents but using a 20 years citation window. Each point represents a single university.

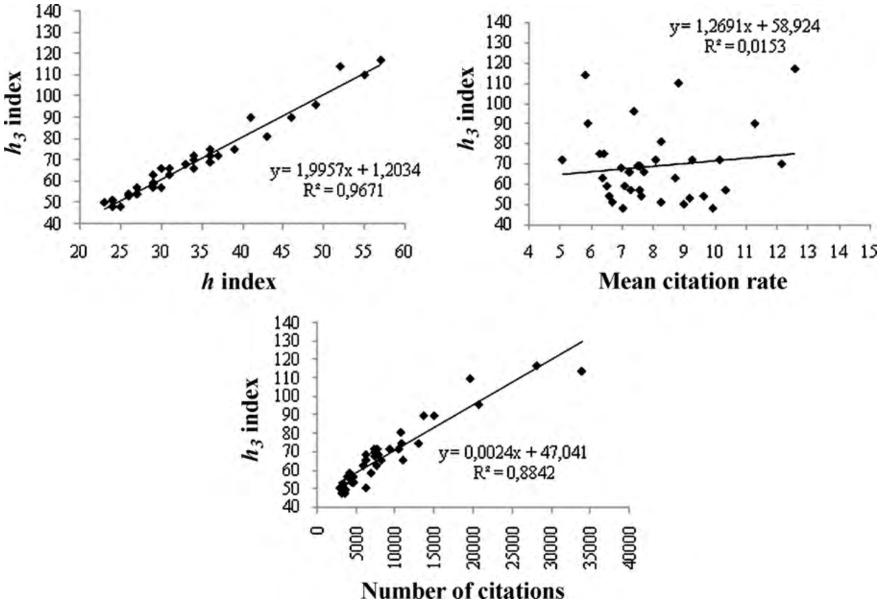

Fig. 3. Relation between the h_3 index and the h index, the mean citation rate and the number of citations for the universities analyzed.

3.3. *Definition of the production and citation projected h index ($^nh_3$)*

Using the results obtained in the empirical study of the dependence of the h index with the number of documents and findings obtained with the application of the h_3 index for universities we defined the new index as:

$$^nh_3 = h_3 \times \left(\frac{n_0}{n}\right)^a$$

where n is the number of documents published by the university in a given year and n_0 is a standard value. Here we used the mean number of documents published by the universities analyzed in 2004 and referenced in the WoS. We would like to clarify that the calculation of n_0 as standard is dependent on the type of comparison that we want to do. If we want to compare universities from the same country we should use the n_0 as the mean number of documents published by all the universities of the country in the period analyzed. If we want to compare universities from two countries we should use n_0 as the mean number of documents published by the universities of the two countries. The constant a was determined empirically and its value is 0.3446 ± 0.0007.

In Table 6 are presented the results obtained when the impact indicator proposed is applied to universities. The values obtained for the h index, h_3 index are also presented in order to present the differences related with ranking. The values were determined using the number of documents, referenced in 2004 WoS for the Brazilian, Portuguese and Spanish universities that produced ca. 60% of the scientific production of the respective country.

As expected, the ranking obtained from the h index varies considerably from that obtained from the mean citation rate. It is well known that the h index is directly influenced only by those documents that reached a number of citations around the value of h. We find in these examples that the percentage of documents in the h-core varies widely from less than 1% to around 7%. h_3 is designed to increase the number of documents retained in the core, to about twice the previous value. In fact, the value of h_3 will depend on the citation performance of those documents lying in between h and $3h$ in the ranking of documents by their citation performance, especially around $2h$. Some events of change in the ranking when we prefer h_3 to h are examples of this. The "Universidad de Alcalá", "Universidad de Extremadura" and "Universidad de Málaga" are cases where the application of the h_3 index puts them in a lower position compared with the ranking obtained for the h index. In the cases of "Universidad Autónoma de Barcelona", "Universidad del País Vasco" or "Universidade Estadual de Campinas" the value obtained for the h_3 index puts them in a better position. Introducing the size correction, the institutional ranking based on the $^nh_3$ shows significant differences from those based on the usual h index, h_3 index or the mean citation rate. Compared with the institutional ranking obtained for the mean citation rate, the ranking obtained for the $^nh_3$ index puts some universities in a lower position. This may be explained by a small number of documents with a very high impact (the documents in the h-core for these universities produce about 50% of the total number of citations obtained by the documents published in that year). These will have a marked effect upon the mean citation rate while not affecting the $^nh_3$.

For this set of 36 universities, the relative standard deviation is 26% for the h index and 12% for the $^nh_3$ index. The larger value of the relative standard deviation of the h index may be explained by the dependence of this indicator on the size of the university or, more rigorously, the number of documents published by each university. The $^nh_3$ index depends on the

Table 6
Values obtained for the h index, h_3 index and $^nh_3$ index for the 2004 referenced documents in the WoS.

University	Total documents	Mean citation rate	h index	h_3 index	$^nh_3$	% documents	
						h_3-core	h index
U. de Alcalá	421	6.67 (29)	24 (12)	51 (18)	71.7 (19)	12%	6%
U. de Alicante	451	9.12(9)	29 (14)	57 (15)	78.3 (8)	13%	6%
U. Autónoma de Barcelona	1543	8.81(11)	41 (7)	90 (5)	80.9 (5)	6%	3%
U. Autónoma de Madrid	1558	12.57(1)	55 (2)	110 (3)	98.6 (1)	7%	4%
U. de Barcelona	2497	11.27(3)	57 (1)	117 (1)	89.1 (2)	5%	2%
U. de Castilla-La Mancha	429	7.61(18)	26 (14)	53 (17)	74.1 (14)	12%	6%
U. Complutense de Madrid	1978	7.55 (20)	46 (5)	90 (5)	74.3 (13)	5%	2%
U. de Córdoba	459	9.92(6)	27 (15)	54 (16)	73.7 (15)	12%	6%
U. de Extremadura	422	7.21(25)	24 (12)	48 (20)	67.5 (25)	11%	6%
U. de Granada	1058	10.33(4)	34 (11)	66 (12)	67.6 (24)	6%	3%
U. de La Laguna	420	8.98(10)	30 (13)	57 (15)	80.2 (7)	14%	7%
U. de Málaga	538	6.57(30)	23 (11)	50 (19)	64.6 (29)	9%	4%
U. de Murcia	566	7.69(17)	27 (15)	54 (16)	68.6 (22)	10%	5%
U. de Oviedo	803	10.13(5)	30 (13)	66 (12)	74.3 (13)	8%	4%
U. del País Vasco	967	7.49(21)	34 (11)	72 (8)	76.0 (10)	7%	4%
U. Politécnica Cataluña	1240	5.88(34)	33 (12)	68 (11)	65.9 (28)	5%	3%
U. Politécnica de Madrid	836	7.37(22)	24 (12)	51 (18)	56.6 (34)	6%	3%
U. Politécnica de Valencia	878	7.01(27)	36 (10)	69 (10)	75.3 (12)	8%	4%
U. Rovira i Virgili	420	9.62(7)	29 (14)	59 (14)	83.0 (4)	14%	7%
U. de Salamanca	650	6.96(28)	27 (15)	57 (15)	69.0 (21)	9%	4%
U. de Santiago de Compostela	1118	8.24(15)	37 (9)	72 (8)	72.3 (18)	6%	3%
U. de Sevilla	1027	7.56(19)	36 (10)	69 (10)	71.4 (20)	7%	4%
U. Valencia	1705	12.14(2)	49 (4)	96 (4)	83.4 (3)	6%	3%
U. de Valladolid	491	7.07(26)	25 (13)	48 (20)	64.0 (30)	10%	5%
U. de Vigo	624	7.27(23)	26 (14)	54 (16)	66.3 (27)	9%	4%
U. de Zaragoza	900	8.07(16)	34 (11)	70 (9)	75.8 (11)	8%	4%
U. de Coimbra	852	7.23(24)	31 (12)	66 (12)	72.8 (16)	8%	4%
U. de Lisboa	823	9.25(8)	36 (10)	72 (8)	80.4 (6)	9%	4%
U. do Porto	1235	8.70(12)	39 (8)	75 (7)	72.8 (17)	6%	3%
U. Técnica de Lisboa	1291	8.24(14)	43 (6)	81 (6)	77.4 (9)	6%	3%
U. Federal de Minas Gerais	894	6.49(31)	31 (12)	63 (13)	68.4 (23)	7%	3%
U. Federal Rio Grande do Sul	1202	8.65(13)	29 (14)	63 (13)	61.7 (31)	5%	2%
U. Federal do Rio de Janeiro	1792	5.79(35)	36 (10)	72 (8)	61.5 (32)	4%	2%
U. Estadual Paulista	1350	5.05(36)	29 (14)	59 (14)	55.5 (35)	4%	2%
U. de São Paulo	5301	6.39(32)	52 (3)	114 (2)	67.0(26)	2%	1%
U. Estadual de Campinas	2035	6.36(33)	36 (10)	75 (7)	61.3 (33)	4%	2%

impact of the documents published by each university and not on their number. As we can see in Table 6 the number of documents varies substantially among universities. The fact that $^nh_3$ includes a size correction as described above explains the lower value of its relative standard deviation when compared with that of the h index.

The "Universidade de São Paulo" has an h_3 index that is 73% higher that the h_3 index of the "Universidad de Granada" and the total number of documents is 5 times higher than the total number of documents of the "Universidad de Granada". The $^nh_3$ index for the "Universidade de São Paulo" is similar to those obtained by the "Universidad de Granada". If we consider the number of documents with zero citations we found that for "Universidade de São Paulo" these represent 30% of the total of documents published in 2004 and 20% for the "Universidad de Granada". This let us to conclude that the indicator does not favour those universities with larger number of documents and low impact. On the other hand, if the number of documents published by the university is low and the percentage of documents with zero citations is larger this university will be privileged. The application of other indicators in order to complement the information given by the $^nh_3$ should be considered.

In order to assess the stability of the $^nh_3$ index we determined this index in 2000, 2001, 2002, 2003 and 2004 for a small set of universities. The mean citation rate, the h index and the h_3 index were also calculated and the stability of these indicators discussed below (Fig. 4, Table 7).

Table 7
Relative standard deviation calculated for the mean citation rate, h index, h_3 index and $^nh_3$ index.

Institution	Relative standard deviation along 5 years			
	Mean citation rate	h index	h_3 index	$^nh_3$ index
Universidad Autónoma de Madrid	7	7	6	4
Universidad de Barcelona	11	11	8	6
Universidade de Coimbra	15	14	17	11
Universidade de Lisboa	22	17	16	11

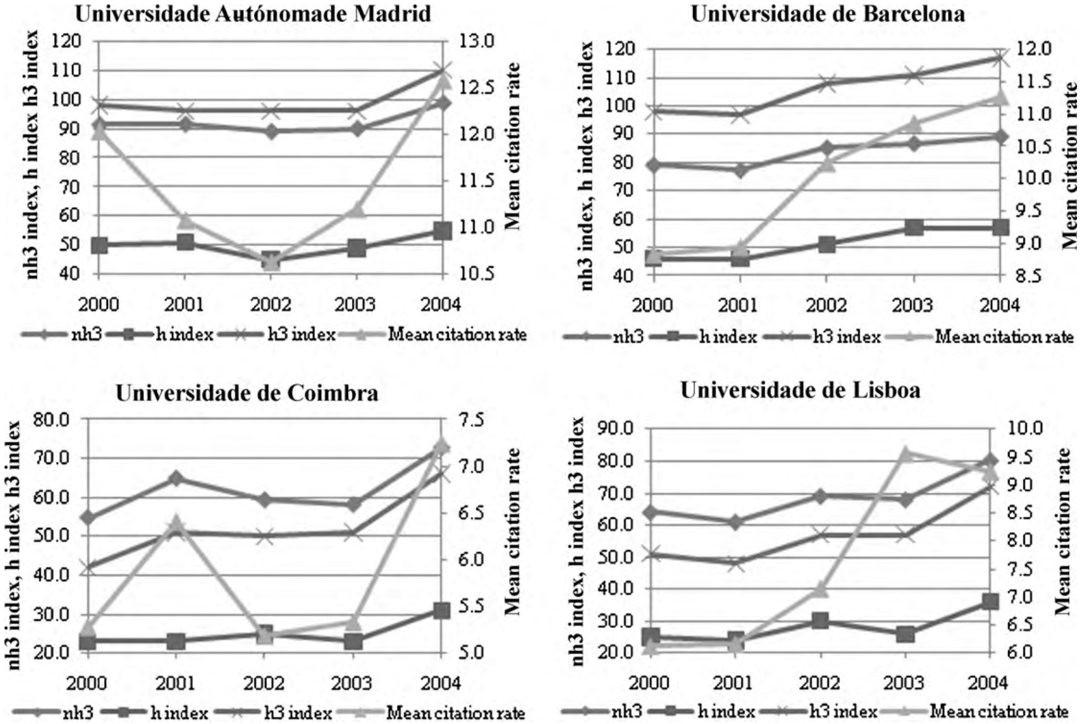

Fig. 4. $^nh_3$ index, h index, h_3 index and mean citation rate determined for 2000, 2001, 2002, 2003 and 2004 using a 5 years citation window.

The values obtained for the relative standard deviation show that the mean citation rate is the less stable indicator. This may be explained by a small set of documents with a high number of citations that have a large influence on the value obtained for the mean citation rate. This suggests that the mean citation rate can be used as an indicator of impact for a university since applied warily. The h index tend to be more stable than the mean citation rate, and this can be explained by the fact that the h index is not influenced by documents with high number of citations. The h_3 index is more stable than the mean citation rate and the h index because it considers in the calculation a larger number of documents than the h index and it is not affected by documents with high number of citations as the mean citation rate. The $^nh_3$ index is the most stable indicator. One possible explanation for this was previously given.

4. Summary and conclusions

The proposed indicator $^nh_3$ is determined using the h_3 index introduced here and a size correction. The h_3 index is determined using the concept of the h index. When we state that the h_3 index is determined using the concept of the h index we are considering that both are not influenced by a small set of documents with high number of citations or by a larger group of documents with low number of citations or without citations. The size correction is applied using the ratio between the mean number of documents published by the particular universities, under consideration, in a given period (used as standard) and the number of documents published by a university in the same period.

The indicator proposed here should be used carefully since it is not field normalized with respect of the citations culture. The number of citations received by a document published in health science tends to be higher than the number of citations received by a document published in engineering or applied sciences, so the indicator proposed should be used carefully for comparison between universities with different research profiles.

The field normalized citation rate must be used as complementary indicator to the $^nh_3$ index. This new indicator has the disadvantage of not being field normalized. Using global averages, the more common field normalized citation rate may be strongly influenced by the extreme behaviour of a small set of documents with a high number of citations. The index $^nh_3$ is immune to these influences and use a projection to about 20 years of the citations counts in a 5-year window.

The results obtained on the study presented here allow us to draw the following findings:

(a) The $^nh_3$ index can be used as a research impact indicator for institutions;
(b) To calculate the $^nh_3$ one considers a single year of publications and their citations along 5 years to estimate the number of citations in about 20 years.
(c) Compared with the h index the $^nh_3$:
  (c.1) retains a larger core of documents, what may be expected to give a better description of a complex university;

(c.2) is independent of the size of the institution as measured by the number of documents produced yearly;
(c.3) is shown to be slightly more stable along the time.

(d) Compared with the mean citation rate (one of the indicators used frequently in ranking of institutions) the ${}^{n}h_3$:
(d.4) does not suffer the exceedingly large effect of a very small number of high impact documents or a high number of documents with low number of citations;
(d.5) is shown to be more stable along the time.

(e) The institutional ranking based on the ${}^{n}h_3$ shows significant differences from those based on the usual h index or the mean citation rate.

(f) The results obtained for the ${}^{n}h_3$ index suggest that this indicator does not favor universities with large number of documents and low impact.

(g) The dependence of the h index on the number of documents and the mean citation rate follows the theoretical model developed by Glänzel; the best value for exponent a (=0.3446 ± 0.0007) is close to the theoretical prediction while the value obtained for exponent b (=0.494 ± 0.003) is significantly lower.

(h) The number of documents in the h_3-core is about twice that in the h-core.

(i) The h_3 index should not be applied to entities for which the number of documents published is low. The application of the h_3 index to review journals with low number of reviews published each year or small teams also with a small number of documents published each year is not reasonable.

(j) The h_3 index does not consider the total number of citations of the documents in the h_3-core, the use of other indicators to complement the information being suggested. (The same is true for the h index applied to individual researchers.)

References

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.

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

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

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

Braun, T., Glanzel, W., & Schubert, A. (2005). A Hirsch-type index for journals. *Scientist*, *19*(22), 8–18.

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

Center for Science and Technology Studies (CWTS). (2009). The Leiden ranking. Retrieved in November from http://www.cwts.nl/ranking/LeidenRankingWebSite.html.

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

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.

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

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

Egghe, L. (2008). 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.

Glanzel, W. (2005 Jun 02–04). On the h-index—A mathematical approach to a new measure of publication activity and citation impact. In *Paper presented at the 33rd annual conference of the Canadian-Association-for-Information-Science* (pp. 315–321). London, Canada: Springer.

Higher Education Evaluation & Accreditation Council of Taiwan. (2009). 2009 Performance *ranking* of scientific papers for world universities. Retrieved in November from http://ranking.heeact.edu.tw/en-us/2009/Page/Methodology.

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*(46), 16569–16572.

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

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

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

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

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

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

Prathap, G. (2006). Hirsch-type indices for ranking institutions' scientific research output. *Current Science*, *91*(11), 1439–11439.

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

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

Scimago. (2009). Scimago institutions rankings 2009 world report. Retrieved in November from http://www.scimagoir.com/pdf/sir_2009_world_report.pdf.

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

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

Vieira, E. S., Nouws, H. P. A., Albergaria, J. T., Matos, C. D., & Gomes, J. A. N. F. (2009). Research quality indicators for Brazilian, Portuguese and Spanish Universities. In *12th international conference on scientometrics and informetrics* 14–17 July, Rio de Janeiro, Brazil,