Scientometrics (2013) 96:617–631
DOI 10.1007/s11192-013-0951-6

Two simple new bibliometric indexes to better evaluate research in disciplines where publications typically receive less citations

Giovanni Anania · Annarosa Caruso

Received: 11 October 2012 / Published online: 8 February 2013

Abstract The paper proposes two simple new indexes—*k* and *w*—to assess a scientist's publications record based on citations. The two indexes are superior to the widely used *h* index (Hirsch, 2005), as they preserve all its valuable characteristics and try to overcome one of its shortcomings, i.e. that it uses only a fraction of the information contained in a scientist's citations profile and, as a result, it is defined over the set of positive integers and does not show a sufficiently fine 'granularity' to allow a fully satisfactory ranking of scientists. This problem is particularly acute in many areas of Social Sciences and Humanities, where scientific productivity and citation practices typically yield fewer citations per paper and, as a consequence, are characterized by 'structurally' lower values of the *h* index. Both the indexes proposed are defined over R^+, their integer part is equal to the scientist's *h* index and they fall in the right-open interval $[h, h+1)$. While the *h* index is influenced only by part of the citations received by a scientist's most-cited publications, the *k* index takes into account all the citations received by her most-cited publications and the *w* index accounts for the citations received by the entire set of her publications. Variants of the *k* and *w* indexes are proposed which consider co-authorship. To show the extent to which the *h* index and the new indexes proposed may yield different results, they are calculated for 332 professors of economics in Italian universities and the results obtained used to rank Italian university departments.

Keywords Bibliometrics · Citation statistics · *h* index · Evaluating research in Social Sciences

JEL Classification A11

We wish to thank Giuseppe Rose and conference participants at the annual meetings of the Associazione Italiana di Economia Agraria ed Applicata (Trento, 4–5 June 2012) and the Società Italiana degli Economisti (Matera, 18–20 Oct 2012) for helpful comments on an earlier draft of the paper.

G. Anania (✉) · A. Caruso
Department of Economics, Statistics and Finance, University of Calabria, Rende, Italy
e-mail: ganania@unical.it

A. Caruso
e-mail: annarosa.caruso@tiscali.it

Introduction

Research performance is a multi-faceted endeavor and its evaluation should never be based on a single qualitative or quantitative indicator. Ideally, peer review should be the primary instrument for research evaluation and bibliometric indexes should be used as support tools to make peer review more objective and transparent.[1] Nevertheless, the use of quantitative indicators becomes unavoidable when the evaluation involves a very large number of individuals or institutions; however, even in such instances, one should never forget the evident risks and limitations of using quantitative indicators.[2] Since bibliometric indexes are more and more extensively used notwithstanding their limitations, refining them in order to improve their capacity to measure, albeit imperfectly, research performances would seem a goal worth pursuing.

This paper therefore proposes two simple new indexes—*k* and *w*—to assess a scientist's publications record based on citations. The two indexes are superior to the widely used *h* index (Hirsch 2005), as they not only preserve all its valuable characteristics but also overcome one of its known major shortcomings, i.e. that it uses only a fraction of the information contained in a scientist's citations profile and, as a result, does not show a sufficiently fine 'granularity' (the *h* index is defined over the set of positive integers) to allow a fully satisfactory ranking of scientists (many show the same value of the index). This problem is particularly acute in those disciplines, such as most of those in Social Sciences and Humanities, where scientific productivity and citation practices typically yield fewer citations per paper and, as a consequence, are characterized by 'structurally' lower values of scientists' *h* indexes.[3] Both the *k* and *w* indexes are defined over R^+, their integer part is equal to the scientist's *h* index and their fractional part is equal to the share of citations in excess of the minimum needed to hold her particular value of the *h* index, and, as a result, fall in the right-open interval $[h, h+1)$. While the *h* index is influenced only by some of the citations received by the most-cited papers in a scientist's publication record, the *k* index takes into account all the citations received by her most-cited publications and the *w* index accounts for the citations received by her entire set of published contributions. Variants of the *k* and *w* indexes are introduced which consider co-authorship.

The next section provides a brief review of the *h* index and some of its variants which have been proposed to try to overcome its limitations. The third section introduces the two new indexes. The fourth section shows the different indexes at work by comparing results obtained applying the *h* index and the two indexes proposed here to a group of 332 professors of Economics in Italian universities and then using the results obtained to rank Italian university departments. The final section concludes.

The *h* index and its variants

Among the indexes based on citations received by the publications of a specific scientist the one known as the *h* index, from the name of the author who introduced it (Hirsch 2005),

[1] In a non-ideal world, peer review too may have its limitations, including the possibility of highly subjective evaluations or conflicts of interest, and a bias against innovative ideas and approaches.

[2] A pertinent and thorough discussion of the use and misuse of citation statistics in quantitative indicators to assess scientific research performances is provided by Adler, Ewing and Taylor (2008).

[3] Nederhof (2006) provides a useful review of differences in publication and citation behaviors between many Social Sciences and Humanities and 'hard' sciences and discusses the implications of such differences for analyzing research performances based on bibliometric indexes.

is certainly the most popular. A scientist has a value of the *h* index equal to s if s of his n publications received each at least s citations and the remaining (n–s) received each at most s citations. The *h* index has several valuable properties, including: that it can be easily computed, it combines in a single index information on both ‘quantity’ (the number of publications) and ‘quality’ (their impact, measured through the citations received), and can be applied at different levels of aggregation (e.g. individuals, research institutions, or countries). At the same time the *h* index shows a number of equally evident limitations, some recognized by Hirsch himself, including the fact that it cannot be used to compare scientists in different disciplines (because the values it assumes are field-specific, due to systemic differences in productivity and citations patterns); moreover, it does not take into account the number of co-authors of each publication, nor does it account for citations received in excess of their minimum number (h^2) given its value, it creates an incentive for self-citation, and is dependent, at least to a certain extent, on the length of a scientist’s career.[4,5]

The indexes proposed in this paper try to address a specific disadvantage of the *h* index, i.e. that it depends on a limited portion only of the relevant information contained in the citations profile of a scientist’s publications and the information it does not consider can be used to obtain a finer assessment of the impact in terms of citations of a scientist’s research output. For example, in the case of two scientists both with an *h* index equal to 10, it could happen that the publications cited at least 10 times for the first one received 2,500 citations in total, and those for the second one only 120. Clearly the *h* index, by ignoring citations in excess of the minimum needed given its value (h^2), does not fairly reflect the evident difference between the two scientists’ citations profiles.

Many indexes have been proposed to overcome this limitation of the *h* index.

Among those proposed which take into account citations in excess of h^2, i.e. those accounted for by the *h* index, are the *g* (Egghe 2006), α[6] (Jin 2006), *R* (Jin et al. 2007), *e* (Zhang 2009), *h(2)* (Kosmulski 2006), *m* (Bornmann, Mutz and Daniel 2008), *hg* (Alonso et al. 2010), h_T (Anderson, Hankin and Killworth 2008) and h^{Δ} (Ruane and Tol 2008) indexes.[7]

The *g*, α, *R*, *e*, *h(2)*, *m* and *hg* indexes all take into account more information on citations received by the researcher’s most highly cited papers, i.e. those contained in the ‘*h* core’,[8] than the *h* index. A researcher has an index *g* equal to z if, after ranking her publications in descending order with respect to the number of citations received, z is the largest number such that the first z publications together received at least z^2 citations ($g \geq h$). The *e* index is defined as the square root of the citations received by the publications in the *h* core in excess of h^2 (*e* is a positive real number). The α index is the average number of citations received by the publications in the *h* core, while *R* is the square root of the total number of citations received by publications in the *h* core; α and *R* are positive

4 Any publication has a probability of receiving a given number of citations which increases with time since it appeared.

5 Alonso et al. (2009) and Todeschini (2011) provide useful discussions of the pros and cons of using the *h* index and a review of variants to the original index which have been proposed to overcome some of its shortcomings.

6 This is sometimes also referred to as the *A* index.

7 Many more additional variants of the *h* index have been proposed, including those by Garcìa-Pérez (2009), Panaretos and Chrisovaladis (2009), Todeschini (2011) and Tol (2009).

8 For a scientist whose *h* index equals k the *h* core is defined, after ranking his publications in descending order with respect to the number of citations received, as the subset of the first k of his publications which each received at least k citations, while the remaining ones each received at most k citations.

real numbers and α, $R \geq$ h; these three indexes are linked by the relation $R = \sqrt{h \cdot \alpha}$. The m index is the median number of citations received by papers in the h core ($m \geq h$). The $h(2)$ index is defined as the highest natural number such that each of the h(2) most cited papers received at least $[h(2)]^2$ citations ($h(2) \leq h$). Finally, the hg index is given by the geometric average (the square root of the product) of the h and g indexes ($g \geq hg \geq h$).

Unlike the indexes mentioned above, the h_T (Anderson, Hankin and Killworth 2008) and the h^{Δ} (Ruane and Tol 2008) indexes take into account also citations received by publications outside the h core. The 'tapered h index' (h_T) uses a Ferrers graph to account for citations received by all publications ($h_T \geq h$). The h^{Δ} index considers part of the citations of the publications in the h core in excess of h^2 and those of the publication 'adjacent' to those in the h core; it is defined as (h+1) − m/(2 h+1), where m is the number of additional citations the scientist needs in order to increase her h index by one, to h+1 ($h \leq h^{\Delta} \leq$ h+1).

The indexes proposed

We propose two new indexes which use more information from the citations of a scientist's publications than the h index. The goal of the two indexes is to make use of the additional information to move from the discrete metric of the h index to a continuous metric, thus allowing for a ranking of those scientists who show the same value of the h index, while preserving the information provided by the h index untouched.

A distribution of scientists by their h index showing a high concentration in few (low) values of the index is common for those disciplines where publications typically receive fewer citations. The median impact factor of the journals by 'category' as listed in the 2011 ISI-WoK Journal Citation Reports can be used as a quick indicator of 'systemic' differences in citations patterns by discipline. Typically, journals in most disciplines in Social Sciences and Humanities tend to show significantly lower median impact factors than those in 'hard' sciences. While, for example, journals in Anthropology, Economics, History, International Relations, Law, Linguistics, Political Sciences and Sociology all show a median impact factor lower than 0.8, for those in Astronomy and Astrophysics, Biology, Chemistry and Ecology, typically, this is well above 1.5 and journals in most categories in the area of medical research have a median impact factor above two. In disciplines where scientists usually show relatively low values of the h index, it may well be worthwhile to use more of the information contained in their citations profiles to rank those with the same value of the h index. This is the case, for example, when the need arises to compare individual publication records in the context of hiring or promotion decisions when the number of candidates exceeds the number of available posts and a choice has to be made, as commonly occurs in the Italian university system. The nation-wide centralized rules governing hiring and promotions in Italy have been modified in recent years to impose on selection and promotion committees the requirement, '*for those disciplines for which their use is accepted by the international community*', to take into account in the evaluation of candidates a list of bibliometric measures based on citation statistics, including the h index.[9] In this and similar frameworks, the possibility to exploit more of the information contained in the citations received by each candidate's publications in order to rank those with the same value of the h index seems a definite improvement.

[9] Italian Ministry of Education, University and Research, Ministerial Decree no. 89/2009, 28 July, 2009.

The first of the two indexes—the k index—maintains all the desirable properties of the h index and takes into account the citations received by a scientist's most cited publications.

The second index—the w index—differs from k insofar as it takes into account the citations received by the entire set of a scientist's publications record.

Let h be a scientist's Hirsch index computed on the basis of the citations received by her publications and cit_j be the number of citations received by the j-th publication included in the h core. The k index is defined as:

$$k = h + [1 - (h^2/\Sigma_{j=1,2,\ldots,h}\, cit_j)], \quad \forall\, h > 0 \tag{1a}$$

$$\text{and} \quad k = 0, \quad \text{if} \quad h = 0. \tag{1b}$$

The expression in square brackets in (1a) is nothing other than the share of citations received by the scientist's publications contained in the h core in excess of their minimum number, h^2, given the value of her h index. When h equals zero, k is set equal to zero; in fact, in this case the h core of the scientist's publications record is empty and, as a result, citations of these publications in excess of h^2 (zero under these circumstances) cannot exist.

Conveniently the k index varies between h (included) and h+1; it equals h when the number of citations received by the publications in the h core equals h^2 and tends to h+1 as the number of citations of the publications in the h core increases.

Let *totcit* be the number of citations received by the entire set of a scientist's publications record. The w index is defined as:

$$w = h + \left[1 - h^2/totcit\right], \quad \forall\, h > 0 \tag{2a}$$

$$\text{and} \quad w = 0, \quad \text{if} \quad h = 0. \tag{2b}$$

The expression in square brackets in (2a) is the share of citations received by a scientist's entire publications record in excess of their minimum number, h^2, given the value of his h index.

The w index also varies between h and h+1; it equals h when the overall number of citations received by a scientist's publications equals h^2 and tends to h+1 as the number of citations increases.

To help understand the differences between the h, k and w indexes, in Fig. 1 the citations profiles of two hypothetical scientists with the same h index are represented. Scientist A published 16 papers; the most cited one received 20 citations, the next 16 citations, two papers received one citation each and two papers have not been cited yet. Scientist B published 12 papers; the most cited one received 9 citations, while, at the other extreme, two papers received one citation each and one no citation. For both scientists the h index equals 6, so their ranking would be the same in a comparative evaluation of their publication records solely based on this index. However, scientist A shows a larger number of citations for both the publications contained in her h core (91 vs. 45) and for the full set of her publications (117 vs. 54). The k and the w indexes take these differences into account, k by considering the relative weight of the citations received by the publications in the h core in excess of the minimum (36 citations, the area of the white square in each of the two profiles), w by considering the relative weight of the citations received by all publications in excess of the same minimum. The values of the k and w indexes for scientist A equal 6.60 and 6.69, respectively; those for scientist B 6.20 and 6.33. While both indexes provide more information than the h index, the choice between them depends

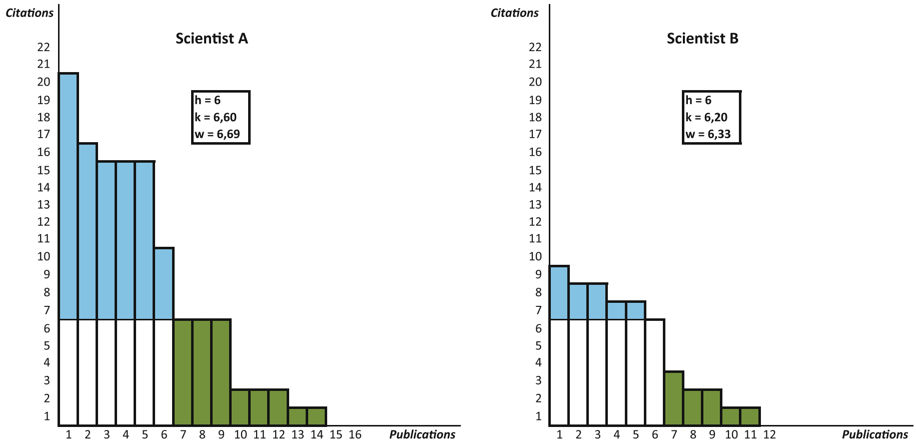

Fig. 1 Citation profiles for two hypothetical scientists with the same *h* index

on the assessment criteria of the specific evaluation: citations received by the most cited publications only, or by the entire publications record.[10]

The *k* and *w* indexes maintain all desirable properties of the *h* index and offer a few additional ones. In particular, as for the *h* index, they can be easily computed, combine in a single index information on both 'quantity' (the number of publications) and 'quality' (their impact, measured through citations), and can be applied at different levels of aggregation; in addition, they give credit for citations in excess of h^2, their interpretation is straightforward (the integer part equals the value of the *h* index and the fractional one is the proportion of citations in excess of h^2), they generate a quantitative assessment of a scientist's publications record which can assume any value over the set of the positive real numbers, thus allowing for the ranking of scientists with the same value of the *h* index, and they always increase with citations (of those in the *h* core in the case of *k*, of all publications in the case of *w*). At the same time, *k* and *w* suffer from some of the same drawbacks as the *h* index, including the fact that they do not take into account the existence of co-authors, they cannot be used as such to compare scientists from different disciplines, they create an incentive for self-citation and, possibly more than the *h* index, are to some extent biased in favor of scientists with a longer career.

The *g* (Egghe 2006), *m* (Bornmann, Mutz and Daniel 2008), *h(2)* (Kosmulski 2006) and *hg* (Alonso et al. 2010) indexes are computed using more information on the citations of publications in the *h* core than the *h* index, but do not use all of it. On the contrary, as is the case for the *k* index proposed in this paper, the *e* (Zhang 2009), α (Jin 2006) and *R* (Jin et al. 2007) indexes all take into account the total number of citations received by publications in the *h* core. However, while the *k* index preserves the property of the *h* index of combining in a single index information on both 'quantity' and 'quality' of a researcher's publications, this is not the case for these indexes. *e* is an index of the 'quality' of the publications in the *h* core in excess of the minimum associated to value of the *h* index, but does not give information on the overall 'quality' (the total number of citations) nor on the number of publications in the *h* core (*h*). The α index provides information on the average 'quality' of the publications in the *h* core, but not on their 'quantity' (*h*). For example,

[10] If *w* is preferred, scientists with a career extending longer into the past may have been given an advantage.

e assumes the same value (2) for two researchers, one with only one publication which received 5 citations, the other with an *h* index equal to 10 and a total number of citations of his publications in the *h* core equal to 104 (the *k* index for the two researchers equals 1.8 and 10.04, respectively). Analogously, α assumes the same value (12) for two researchers, one with only one publication which received 12 citations, the other with an *h* index equal to 10 and a total number of citations of his publications in the *h* core equal to 120 (the *k* index for the two researchers equals 1.92 and 10.17). The *R* index merges information on both ‘quantity’ and ‘quality’ of a researcher’s publications falling in his *h* core, but, again, does not give information on the number of publications this contains (*h*). For example, *R* assumes the same value (10) for two researchers, one with only one publication which received 100 citations, the other with an *h* index equal 7 and the 7 publications in the *h* core having received 100 citations in total (the *k* index for the two researchers equals 1.99 and 7.30). In fact, for these reasons, indexes *e*, α and *R* have not been proposed as possible replacements of the *h* index, but their use is suggested in conjunction with *h*. Because its integer part equals index *h*, index *k* can be used instead of *h* and, by using some of the information on the citations of a researcher’s publications record which is ignored in the calculation of *h*, it provides additional relevant information.

Like index *w*, also the h_T (Anderson, Hankin and Killworth 2008) and the h^{Δ} (Ruane and Tol 2008) indexes take into account citations received by publications outside the *h* core. However, unlike *w*, the h_T index does not give information on the number of publications contained in the *h* core (*h*), while h^{Δ} does not take into account all the citations received by a scientist’s publications. On the contrary, index *w* provides information on the number of publications contained in the *h* core (its integer part equals *h*) and considers all citations received by a scientist’s publications.

The *h*, *k* and *w* indexes do not take into account the existence of co-authors. In the case of two scientists with the same number of publications and identical citations profiles, the publications of the first scientist each the result of collaboration with 10 co-authors and those of the other all the result of her own work only, the *h*, *k* and *w* indexes would all return the same value for both. For a bibliometric index to be used to compare the publication records of social scientists it is a severe limitation to ignore the existence of co-authors. Because of the marked differences in publication and citation patterns (Nederhof 2006), this becomes an even more severe shortcoming when the index is used to compare research performances of individuals or institutions across different fields.

For all indexes variants have been, or may be developed to take into account the existence of co-authors by normalizing, in one way or another, the number of citations received by each publication. The ‘normalized individual *h*-index’, one of the bibliometric statistics generated by the Publish-Or-Perish software (Harzing, www.harzing.com/pop.htm), is defined as the *h* index calculated after having normalized citations received by each publication by dividing them by the number of the co-authors of that publication. Alternative proposals of indexes which take into account the number of co-authors include those by Batista et al. (2006) and Schreiber (2008).

The *k-norm* and *w-norm* indexes modify the *k* and *w* indexes to take into account the number of co-authors by considering normalized, rather than absolute, citations, i.e. the number of citations received by each publication divided by the number of its co-authors.

The *k-norm* index, is defined as:

$$k\text{-}norm = h\text{-}norm + [1 - (h\text{-}norm^2/\Sigma_{j=1,2,\ldots,h-norm}\ citnorm_j)], \quad \forall\ h\text{-}norm > 0 \qquad (3a)$$

$$\text{and} \quad k\text{-}norm = 0, \quad \text{if} \quad h\text{-}norm = 0, \tag{3b}$$

where *h-norm* is the 'normalized individual *h*-index' and $citnorm_j$ is the number of normalized citations received by the *j*-th publication included in the scientist's *h-norm* core. When *h-norm* equals zero, *k-norm* is set equal to zero, for the same reason given above for the *k* index.

The *k-norm* index varies between *h-norm* and *h-norm* + 1; it equals *h-norm* when the number of normalized citations received by the publications in the *h-norm* core equals $h\text{-}norm^2$ and tends to *h-norm* + 1 as the number of normalized citations of publications in the *h-norm* core increases.

The index analogous to *w* based on normalized citations, the *w-norm* index, is defined as:

$$w\text{-}norm = h\text{-}norm + \left[1 - h\text{-}norm^2/totcit\text{-}norm\right], \quad \forall\ h\text{-}norm > 0 \tag{4a}$$

$$\text{and} \quad w\text{-}norm = totcit\text{-}norm/(1 + totcit\text{-}norm), \quad \text{if} \quad h\text{-}norm = 0, \tag{4b}$$

where *totcit-norm* is the total number of normalized citations received by a scientist's publications. (4b) guarantees that for a scientist whose *h-norm* index equals zero and, nevertheless, is the co-author of publications which received citations, the *w-norm* index is greater than zero (giving credit for the citations received) and less than one.

The *w-norm* index varies between *h-norm* and *h-norm*+1; it equals *h-norm* when the number of normalized citations received by a scientist's publications equals $h\text{-}norm^2$ and tends to *h-norm*+1 as the number of normalized citations increases.

For any scientist the following relations hold: $h\text{-}norm \le k\text{-}norm \le w\text{-}norm$ and $h \le k \le w$.

Finally, indexes *k*, *w*, *k-norm* and *w-norm* can all be easily modified to weed out self-citations and/or to take into account, when necessary, differences in the length of the scientists' careers by normalizing citations based on the 'age' of each publication.

An application

To show how results obtained using the indexes proposed in this paper compare with those obtained using the *h* and the *h-norm* indexes, we use in this section the different indexes to evaluate the publication records of the 332 assistant, associate and full professors of 'Economic Policy' in Italian universities.[11] The choice of this particular set of scientists is motivated by the fact that Economics is among the disciplines where publications tend to receive a significantly lower number of citations than in other fields, such as Biochemistry or Medicine, and, as a result, the distribution of economists by the value of their *h* index shows a large number of them holding the same (low) value of the index.[12] When this is the

[11] They include 'ricercatori', 'ricercatori non confermati' and 'ricercatori a tempo determinato'; 'professori associati' 'confermati' e 'non confermati'; 'professori ordinari' e 'straordinari' active in the 64 Italian - public and officially accredited by Italian Ministry of Education, University and Research (MIUR) private universities—in which at least one professor of 'Economic Policy' is present. MIUR classifies professors in Economics in six groups: 'Economics', 'Economic Policy', 'Public Finance', 'History of Economic Theory', 'Econometrics' and 'Applied Economics'. The set was extracted from the MIUR database (http://cercauniversita.cineca.it/php5/docenti/cerca.php) on 28th Feb 2012.

[12] Studies related to the evaluation of the research performance of professors of economics in Italian universities using bibliometric indexes include Checchi and Jappelli (2009), Corsi and Defrancesco (2012), Corsi, D'Ippoliti and Lucidi (2010, 2011), Lippi and Peracchi (2007), and Reichlin (2008). Checchi and

case, the need may arise to make a fuller use of the information associated with their citations profiles to obtain a ranking of those with the same value of the *h* index.

For each Italian professor of ‘Economic Policy’ the citations profile has been obtained from the Thomson-Reuters ‘ISI Web of Knowledge’ (ISI-WoK) data base, which was accessed between 13th and 17th March, 2012. ISI-WoK includes a large set of scientific journals, mostly in English, selected on criteria set and implemented by Thomson-Reuters itself. Although this data base remains the most widely used in quantitative evaluations of scientific research outputs, many would argue that the selection of journals it includes is biased. If this is true, both the identification of the papers published by each author and the number of citations each of these publications received would suffer from the same bias. However, because our aim is not to perform a comparative assessment, but only to provide evidence of the different results which could be obtained by using the different indexes, the possible bias of the ISI-WoK data base does not constitute a problem here.

Publications by each author have been identified on the basis of the ‘Social Sciences Citation Index’ (SSCI) data base. A very careful assessment of the results obtained has been performed with the aim of limiting as much as possible problems due to homonymy. The results of the search for each scientist have been analyzed making use of the ‘refine results’ option of the ISI-WoK web-based facility; this allows for a screening of the search results obtained for a specific author’s name based on a number of filters, including ‘categories’ (discipline sub-sections), ‘subject areas’, and author affiliation (‘institutions’). The first refinement of the ‘gross’ results obtained was the exclusion of papers in a ‘subject category’ manifestly different from economics, such as geriatrics or psychiatry. Then each author’s publications record was checked to ensure that the process had not failed to identify publications which were, without any doubt, authored by a different scientist bearing the same name.[13] Once the identification of the publications extracted from the ISI-WoK data base for each of the 332 scientists considered was completed, the ‘create citation report’ facility was used; this generates a report which includes a number of citation statistics for the specific author. This information constituted the base for calculating the values of indexes *h*, *k* and *w* and *h-norm*, *k-norm* and *w-norm* for each of the university professors considered.[14]

The cumulative distributions of the 332 professors of ‘Economic Policy’ in Italian universities by the value of their *h* and *h-norm* indexes are provided in Fig. 2. Both distributions show a high number of professors with the same value of the index. 136 professors have an *h* index equal to zero (for 104 of them no publication in the ISI-WoK data base was found, while the publications of the remaining 32 have not been cited yet), 88 have an *h* index equal to one, 43 equal to two, 33 equal to three; only for 32 of them is the *h* index higher than three. The number of professors with an *h-norm* index equal to zero

Footnote 12 continued
Jappelli (2009) used the *h* index calculated using information from the Google Scholar data base to evaluate the 696 full professors of Economics. Corsi and Defrancesco (2012) analyzed the publication records of the 311 full, associate and assistant professors of ‘Agricultural Economics and Rural Appraisal’ calculating the *h* and *g* indexes based on information from the SCOPUS, ISI-WoK and Google Scholar (via Publish or Perish) data bases.

[13] No filtering based on the author’s affiliation has been performed, as it is impossible to control for the mobility of an author from one institution to another.

[14] The data base with the publications record from ISI-WoK, the citations profile and the bibliometric indexes for each of the 332 university professors of ’Economic Policy’ of Italian universities, and a table with the values of the six indexes calculated for each of them are available at http://www.ecostat.unical.it/anania/AnaniaandCaruso.htm.

is 158; 95 have an *h-norm* index equal to one, 38 equal to 2 and only 41 have a value of the index equal to three or higher. These distributions signal that for this group of scientists the use of the *h* or the *h-norm* index is clearly unsatisfactory if one has to evaluate them comparatively and generate a rank.

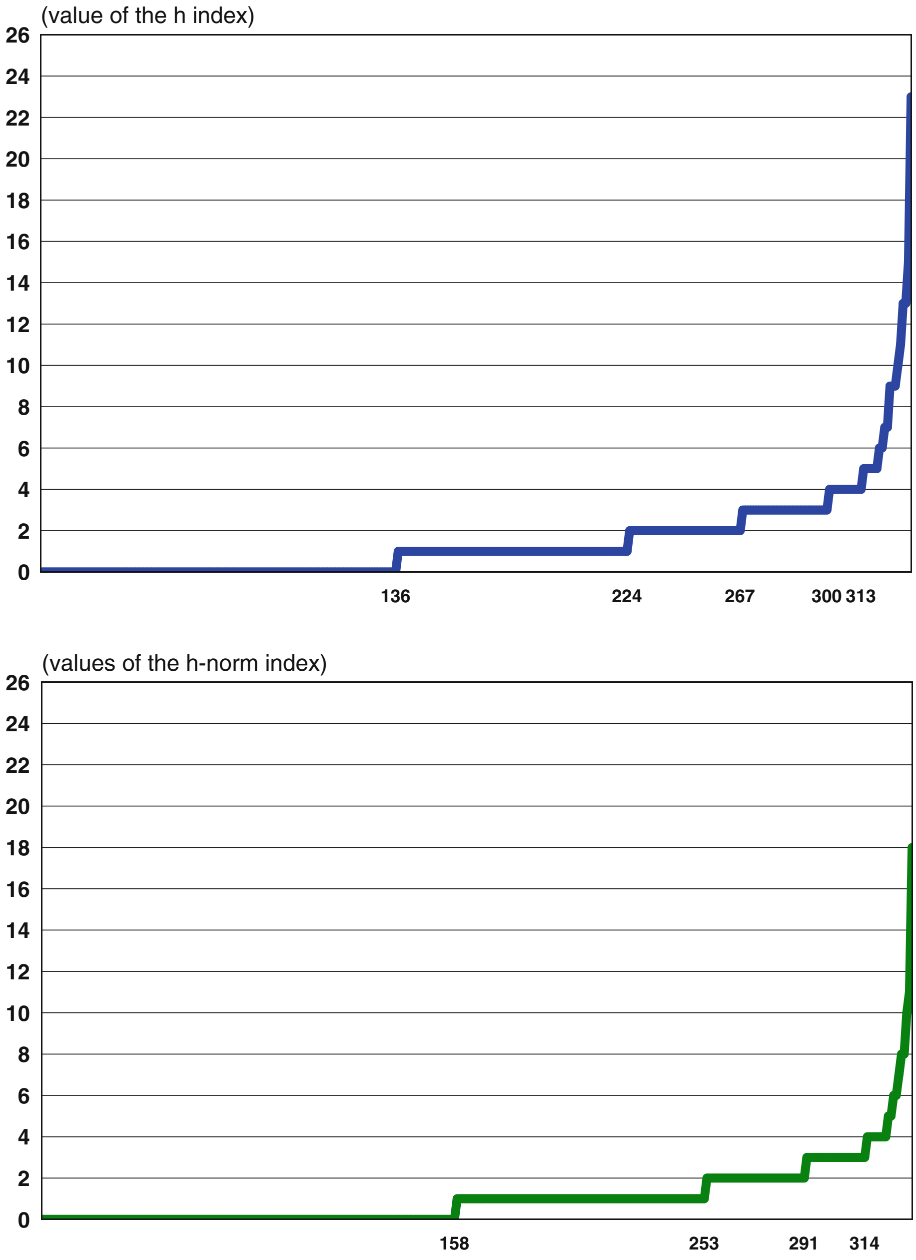

Fig. 2 Cumulative distributions of the professors of 'Economic Policy' in Italian universities by the value of their *h* and *h-norm* indexes

Figure 3 offers the cumulative distribution of the professors of 'Economic Policy' in Italian universities by the value of their *h* index along with those by the *k* and *w* indexes. When the scientist has a value of his *h* index equal to zero, the *k* and *w* indexes equal zero as well and no ranking can be generated based on citation statistics. When the *k* index—the one which takes into account citations received by publications contained in the *h* core in excess of their minimum—is considered, the same is true for the 30 professors with the *h* and *k* indexes both equal to one (these are scientists whose single publication contained in the *h* core received only one citation). The *k* index allows ranking the remaining 57 professors with a *h* index equal to one as well as those with a larger *h* index. Analogously, when the use of the citations received by the entire set of an author's publications is found more appropriate, the *w* index ranks the 64 professors with a *h* index equal to one and a number of citations received by their scientific production greater than one, as well as those with a higher value of the *h* index.

Figure 4 presents the cumulative distribution of the same group of professors by the value of their *h-norm* index along with those by the *k-norm* and *w-norm* indexes. When the use, for each publication, of the number of citations normalized by the number of co-authors is preferred, again, the indexes proposed in this paper—one based on the total number of normalized citations received by the publications contained in the *h-norm* core (the *k-norm* index), the other based on the total number of normalized citations received by all the scientist's publications (the *w-norm* index)—allow us to exploit citational information disregarded in the calculation of the *h-norm* index to rank individuals with the same value of this index. In the case of the *w-norm* index, it even ranks some of the professors with a *h-norm* index equal to zero; in fact, it returns a positive value for those 22 with co-authored publications which received citations, but not enough to yield them a *h-norm* index equal to one.

Bibliometric indexes based on citations are often used to rank research institutions. In order to assess how the choice of a specific index may affect the ranking obtained, we used the average values calculated over their members of the *h*, *k*, *w*, *h-norm*, *k-norm* and *w-norm* indexes to generate rankings of the 109 departments of the 64 Italian universities in which at least one professor of 'Economic Policy' is present.[15]

Although this particular issue is outside the scope of this paper, it is interesting to note that the choice whether to account for the number of co-authors of each publication or not, does make a difference. We calculated for each of the 109 departments the differences, in absolute value, between the rankings obtained using the *h* and the *h-norm* indexes. For 12 out of the 109 departments the two rankings differ by 20 positions or more, for 34 departments by 10 or more. At the other end of the distribution, rankings are equal or differ by 5 positions or less only for 21 departments.

Basing the comparative evaluation of research institutions on quantitative citation indexes which take into account more information contained in the citations profiles of their members than that used by the *h* and *h-norm* indexes can indeed make a difference. When rankings are based on the *k* index instead of the *h* index for 20 departments out of 109 the ranking change by 10 positions or more; the same happens for 15 departments

[15] An alternative way of measuring the performance of research institutions is the one based on the 'successive *h*-index' (h_2) proposed by Schubert (2007). An institution has an index h_2 equal to s if s of its n researchers each have a *h* index at least equal to s, and the remaining (n–s) each have a *h* index which does not exceed s.

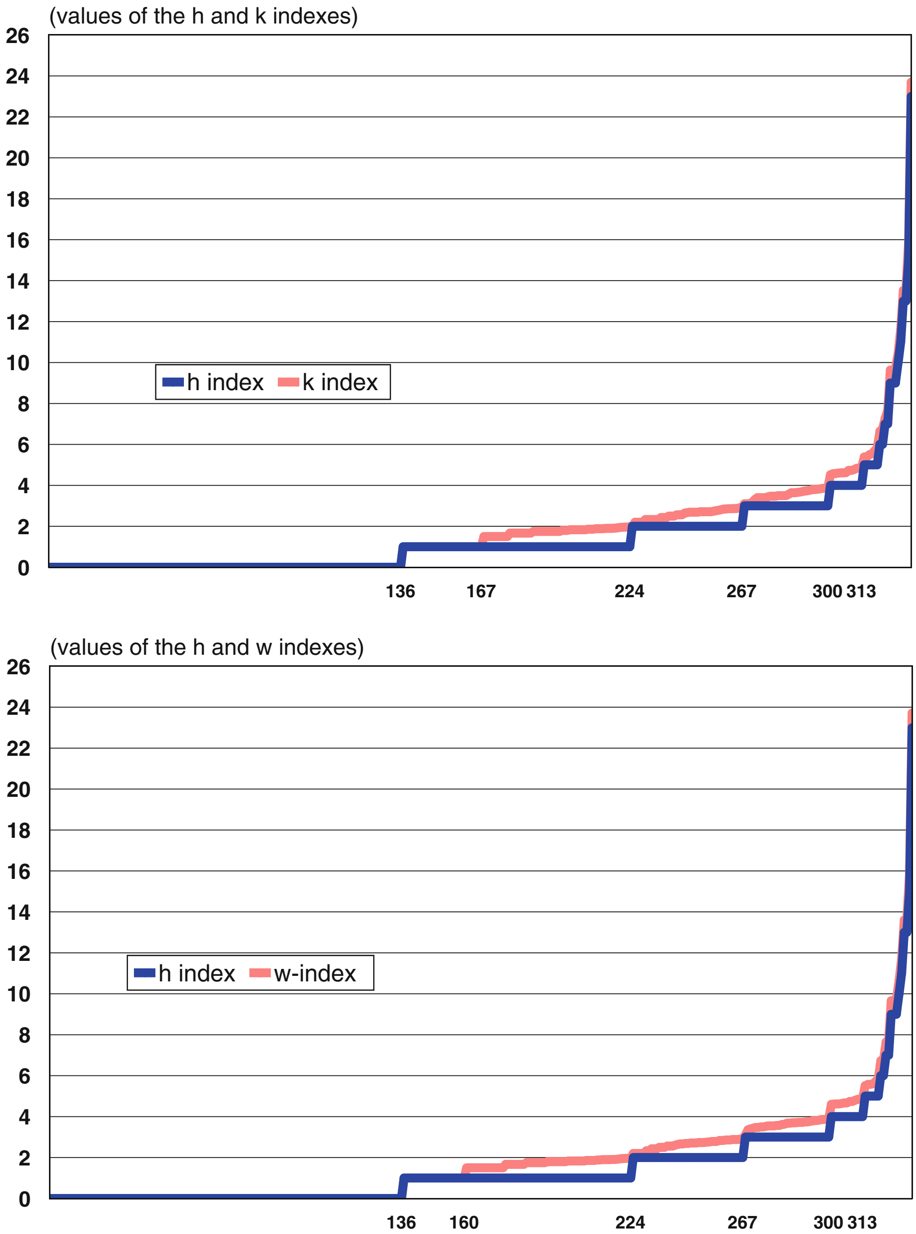

Fig. 3 Cumulative distributions of the professors of 'Economic Policy' in Italian universities by the value of their *h*, *k* and *w* indexes

when the *w* index is used, while differences are much less pronounced when the rankings based on the *k* and *w* indexes are compared. Similar results emerge when rankings based on the indexes which use normalized citations are considered. If the *k-norm* index is used instead of the *h-norm* one, the ranking changes by 10 positions or more for 13 of the 109 departments; the same is true for 9 departments if the *w-norm* index is used, while only one

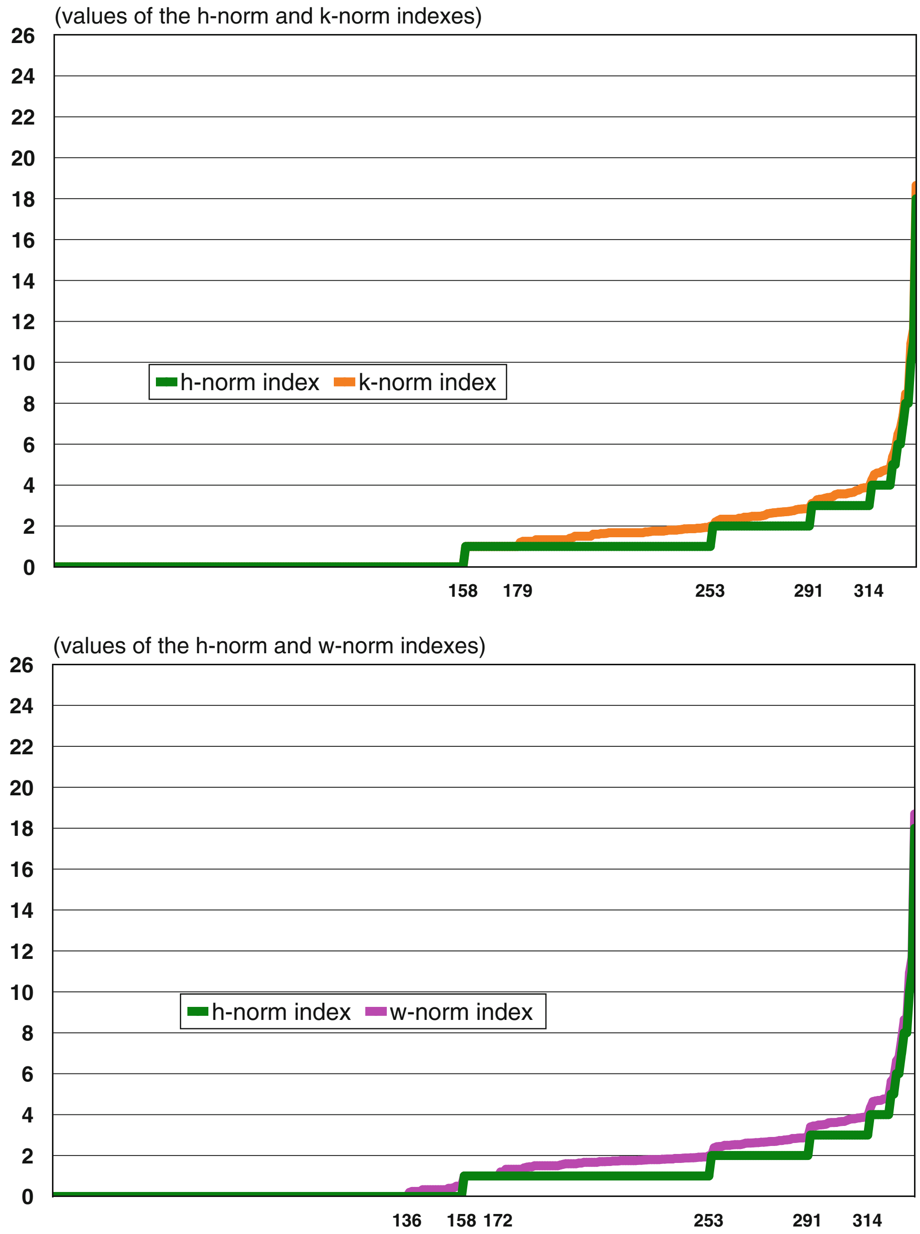

Fig. 4 Cumulative distributions of the professors of 'Economic Policy' in Italian universities by the value of their *h-norm*, *k-norm* and *w-norm* indexes

department sees its ranking change by 10 positions or more when those based on the *k-norm* and *w-norm* indexes are considered.[16]

[16] A table containing the average values of the *h*, *k*, *w*, *h-norm*, *k-norm* and *w-norm* indexes for the 109 departments, the rankings based on each of these indexes and the differences between these rankings is available at http://www.ecostat.unical.it/anania/AnaniaandCaruso.htm.

Conclusions

Despite their widely recognised limitations, quantitative indicators are more and more extensively used to measure and compare research performances of individuals and research institutions. This paper introduces two indexes which try to improve on the *h* index, by far the most widely used bibliometric index. Two simple variants of the *h* index are proposed: one—the *k* index—based on the assumption that in assessing a scientist's citations profile the total number of citations received by her most-cited contributions is relevant; the other—the *w* index—to be used when the evaluator believes that citations received by all the publications by a scientist should be considered to evaluate her performance.

Although substantially different from the *h* index in the implicit assumption made about what is important and should be considered in evaluating publication records using citations, the two indexes proposed are evident extensions of this index which use more of the information contained in a scientist's citations profile. In fact, (a) the integer part of both indexes conveniently equals the value of his *h* index and (b) the fractional one equals the share of citations (of those in the *h* core for *k*, of citations received by any of the scientist's publications in the case of *w*) in excess of the minimum needed (h^2) for him to hold a value of the index equal to h; (c) they are both smaller than the value of the scientist's *h* index augmented by one, and (d) they strictly increase with the number of citations (of those in the *h* core for *k*, of citations received by any of the scientist's publications in the case of *w*). As for the *h* index, *k* and *w* are easy to compute and can be applied at different levels of aggregation.

k and *w* may be particularly useful when used to comparatively evaluate research performances in many disciplines in Social Sciences and Humanities, where citations per paper tend to be lower, yielding lower values of the *h* index and a strong concentration of scientists in very few (low) values of the index. In this case the use of the two indexes proposed in this paper allows us to exploit more of the information on the citations received by their papers than that contained in the *h* index and move from a discrete measure of their performance to a continuous one, thereby making a 'finer' ranking possible. This can be useful when bibliometric indexes are used to evaluate individuals or research institutions in the framework of hiring selections, promotion decisions or funding allocation exercises.

Being extensions of the *h* index, the two indexes share some of its limitations, including the fact that their values are discipline-specific, which implies they cannot be used as such to make comparisons across different fields, they are influenced by self-citations, and are to some extent biased in favour of scientists with a longer career. However, to overcome these limitations, variants of the *k* and *w* indexes along the lines of those proposed for the *h* index can be easily developed.

Finally, whether or not we normalize citations to take into account co-authorship does make a difference. When bibliometric indexes are used, co-authorships should be accounted for, especially in the case of Social Sciences and Humanities. A version of the two indexes which use normalized citations to take into account the number of co-authors of each publication has been proposed.

References

Adler, R., Ewing, J., & Taylor, P. (2008). *Citation Statistics. A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS).* (http://mfa-national.org/yahoo_site_admin/assets/docs/Documents/CitationStatistics-FINAL-1.pdf).

Alonso, S., Cabrerizo, F. J., 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.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2010). *Hg*-index: A New Index to Characterize the Scientific Output of Researchers Based on the *h*- and *g*- Indices. *Scientometrics, 82*(2), 391–400.

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

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., 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.

Checchi, D., & Jappelli, T. (2008). Ricerca per indice h. *La Voce*, 16 dicembre (www.lavoce.info).

Corsi, M., D'Ippoliti, C., & Lucidi, F. (2010). Pluralism at Risk? Heterodox Economic Approaches and the Evaluation of Economic Research in Italy. *American Journal of Economics and Sociology, 69*(5), 1495–1529.

Corsi, M., D'Ippoliti, C., & Lucidi, F. (2011). On the Evaluation of Economic Research: The Case of Italy. *Economia Politica, 28*(3), 369–402.

Corsi, A., & Defrancesco, E. (2012). La valutazione basata su *peer-review* e indicatori bibliometrici: esperienze in atto e riflessioni sul settore AGR/01. *Agriregionieuropa 8*(28) 31–36 (http://agriregionieuropa.univpm.it/).

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

Garcìa-Pérez, M. A. (2009). A multidimensional extension to Hirsch's h-index. *Scientometrics, 81*(3), 779–785.

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.

Jin, B. H. (2006). h index: An evaluation indicator proposed by scientist. *Science Focus (in Chinese), 1*(1), 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.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI Newsletter, 2*(3), 4–6.

Lippi, M., & Peracchi, F. (2007). Il primo esercizio italiano di valutazione della ricerca: una prima valutazione. *Rivista Italiana degli Economisti, 12*(2), 267–276.

Nederhof, A. M. (2006). Bibliometric monitoring of research performance in the Social Sciences and Humanities: A review. *Scientometrics, 66*(1), 81–100.

Panaretos, J., & Chrisovaladis, M. (2009). Assessing scientific research performance and impact with single indices. *Scientometrics, 81*(3), 635–670.

Reichlin, P. (2008). La valutazione della ricerca in economia. *Statistica & Società, 6* (special issue), 33–38.

Ruane, F. P., & Tol, R. S. J. (2008). Rational (successive) *h*-indices: An application to economics in the Republic of Ireland. *Scientometrics, 75*(2), 395–405.

Schreiber, M. (2008). To share the fame in a fair way, h_m modifies h for multi-authored manuscripts. *New Journal of Physics 10*(4) (open access, Doi: 10.1088/1367-2630/10/4/040201).

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

Todeschini, R. (2011). The *j*-index: a new bibliometric index and multivariate comparisons between other common indices. *Scientometrics, 87*(3), 621–639.

Tol, R. S. J. (2009). The h-index and its alternatives: An application to the 100 most prolific economists. *Scientometrics, 80*(2), 317–324.

Zhang, C.-T. (2009). The *e*-index, Complementing the *h-index* for Excess Citations. *Plos One 4*(5), 1–4 May. (www.plosone.org).