New Methods and Evaluation Criteria of Research Efficiency

Olga Popova

PhD in technical sciences, Associate Professor, Kuban State Technological University
Krasnodar, Russia; popova_ob@mail.ru

Dmitry Romanov

PhD in pedagogical sciences, Associate Professor, Kuban State Technological University
Krasnodar, Russia; romanovda1@rambler.ru

Boris Popov

PhD in technical sciences, Associate Professor, Kuban State Technological University
Krasnodar, Russia; pbk47@mail.ru

Vladimir Karandey

PhD in technical sciences, Associate Professor, Kuban State Technological University, Krasnodar, Russia

Svetlana Kobzeva

PhD in technical sciences, Associate Professor, Kuban State Technological University, Krasnodar, Russia

Marina Evseeva

PhD in philological sciences, Senior Lecturer, Kuban State University
Krasnodar, Russia; khizova2004@mail.ru

Abstract

This paper is aimed at designing new methods of assessing research efficiency.The well-established h-index is known to be used for assessing a scientist's productivity. Its introduction in 2005 was a considerable step forwardas compared to such indicator as the ratioof the number of references to researcher's publications and the number of publications themselves. However, h-index itself has got some shortcomings, the main of which is a weak differentiating capacity – the number of references to author's most quoted publications stops being a matter of importance upon reaching a certain level. There is an urgent need for designing a researcher's efficiency assessment method which will keep h-index strengths and make up for its weaknesses. This will allow reaching a more objective evaluation of research output. Methodology of the paper is based on the systemic approach (science is viewed as a social institution inseparable from society), meta-systemic approach (research output is viewed as a meta-system, i.e. a system with relatively independent components), statistically defined probability-based approach (research activities are viewed as a probabilistic process), synergetic approach (science is viewed as a self-organising system), qualimetric approach (research output is viewed as a latent variable resettable by a set of criteria)

Keywords: research activities, efficiency, objectivity, assessment, h-index.

1. Introduction

Scientometrics as one of the most important fields of Sociology of science is currently developing rapidly (Chebotarev, 2013, Egghe, 2008, Franceschini et al., 2010, Guns & Rousseay, 2009, Shtovba & Shtovba, 2013). The most important task of promoting scientometric parameters is motivation of both individual scientists and whole research groups towards regular and efficient activities. It is to do with the fact that the most significant function of science as socio-cultural phenomenon and social institute is to produce (generate) new knowledge and to translate (propagate and embed) it, to

ensure appropriate conditions for innovative development of society in general and specific fields of human activity in particular. All over the world in the last century and in the present one, the pursuit of objectivity in assessing research activities to establish adequate criteria of their measurement has been intense and unrelenting (Eck &Waltman, 2008, Egghe, 2006, Eichorn&Yankauer, 1987, Hirsch, 2005, Lehmann et. al., 2006).

Information society has provided good conditions for the adequate assessment of research activities of individuals and groups (Hirsch, 2005, Franceschini et al., 2010, Guns &Rousseay, 2009, Shtovba&Shtovba, 2013). Thanks to the application of modern IT, it has become possible to obtain primary monitoring information on the "materialized" results of researchers (Web of Science, Scopus, Agris, etc.). However, the assessment of research activities efficiency is still based on a limited set of parameters (although it must be multi-parametric, only then it can be considered objective). The issue of study is the question: what must the methods of a researcher's efficiency assessment be like, which possess a proper differentiating capacity? The aim of study is designing new assessment methods for determining a researcher's efficiency. The object of study are research activities of a scientist; the subject matter of study is the quality (productivity) of research activities.

Nowadays the globally recognized researcher's efficiency indicator is Hirsch index (h-index), according to which a researcher's efficiency equals h if there is no less than h references to no less than each of his h publications (the number of his/her less efficient publications and references to them do not matter). Before 2005 the following indicator was used: a ratio of number of references to author's publications to the number of publications themselves. Inadequacy of the latter is in the fact that the increase in number of publications decreases a researcher's efficiency. As opposed to it, the main advantage of h-index is that it settles the discrepancy between volume and efficiency of research activities (quantity and quality of publications). Thanks to h-index scientists do not tend to "split" their publications so as to increase their number, but to produce quality papers of significance. On the other hand, h-index does not "forbid" researchers publish new scientific papers. In other words, h-index stimulates generation of quality publications that tend to have great impact on science. It is also possible to analyze publication activity by such indicator as the number of references to the most cited publication.

Nevertheless, h-index has disadvantages too. From the standpoint of the authors of this paper, the major drawback here is lack of differentiating capacity: two different researchers with the same h-index can have different numbers of references to their most cited publications. Apart from that, with any number of citations h-index cannot exceed the number of publications. To a certain extent, the h-index shortcomings are compensated by g-index: for a given set of publications sorted out in descending order of citations g-index is the largest number at which g of the most cited publications obtained (total) not less than g^2 of citations. Yet, even g-index does not always differentiate a researcher's efficiency. Besides, it is limited by the number of a researcher's publications. Modern scholars (Chebotarev, 2013, Franceschini et al., 2010, Guns and Rousseay, 2009, Shtovba & Shtovba, 2013) have carried out quite an impressive survey of existing indicators of a researcher's efficiency which are based on primary information on their citations.

Modern IT provide favourable conditions for solving scientometric problems: it is thanks to them that citations detecting and fixation are possible, as well as the analysis of a researcher's publishing activities. However, assessment and analysis of research activities are still based on a rather limited number of indicators which decreases objectivity and prevents from giving a full assessment of their input into a particular area of science. Everyone knows Goodhart's law: "When a measure becomes a target, it ceases to be a good measure." It is no doubt that research activities assessment must be multi-criterial.

2. Methodology of Study

2.1 Methods of study

The following methods were used within the study: cognitive and mathematical modeling, qualimetric methods, methods of mathematical statistics and automated systemic and cognitive analysis.

2.2 Research organization

The study was based on higher education institutions of the Krasnodar territory. The authors managed to specify new criteria of research efficiency assessment thanks to qualimetric methods (theory of latent variables). Owing to Russian Science Citation Index (on eLIBRARY platform) were obtained primary data on research activities of Kuban's university lecturers (n=420). Automated systemic and cognitive analysis made it possible to identify validity and differentiating capacity of the authors' indicators of research efficiency.

3. Results of Study and their Discussion

Without denying significance of h- and g-indices, the authors uphold the position that it is necessary to combine them with other indicators – only then it will be possible to obtain a full picture accurately displaying a researcher's efficiency. The authors suggest that the research activities assessment should be carried out in the following way. The first stage is sorting out publications in descending order of their citations and calculating the h-index. The second stage represents the authors' method, different from the generally recognized one. Let N be a total number of a researcher's publications, h is the number of the most cited ones (each one is referred to at least h times), f_i is the number of citations to the

i^{th} publication, then the total number of references to the most cited and to all publications will be $Z = \sum_{i=1}^{h} f_i$ and $S = \sum_{i=1}^{N} f_i$

respectively, and the authors' researcher's efficiency index is $r = \frac{Z}{h}$. The advantage of the authors' method is in lack of any principal limitations connected with a number of publications.

Other research efficiency indicators can also be calculated on the basis of primary information obtained from scientometric databases (scientometric systems). Let us consider them.

Let Q be a set of journals which published a researcher' papers, W be a set of his/her co-authors, then $\varsigma = P(Q)$ will be the number of journals, $\omega = P(W)$ will be the number of co-authors (P is the power of a set. Let the number of co-authors be a coordinating number of a researcher (in chemistry the coordinating number is the number of the nearest atom's neighbours in space). Let **w** be social valence of a researcher which represents such number when the researcher has no less than **w** publications with each of no less than **w** authors. The given indicator reflects firmness (steadiness) of a researcher's social ties with his/her co-authors. Another important indicator is productivity of social broadcasting of research results which equals **q**, on condition that no less than **q** of scientific papers are published in each of no less than **q** journals. The relevance of the last indicator is not only in implementing such researcher's function as broadcasting his/her own research results. It is very important to mention that at high values of this indicator it is highly unlikely that all publications of a researcher are inferior.

It is worthwhile mentioning that a researcher's efficiency is evaluated on the basis of quantity of citations to his/her publications. Yet, in modern information systems there are self-citations, citations by co-authors and citations by "third-party". No doubt that it is the latter that are genuine indicators of publications value for the academic community. Therefore, the authors suggest introducing the following indicator – citation index for a particular publication:

$N = N^{/} + \sum_{k=1}^{N^{//}} 0.75^k + \sum_{k=1}^{N^{///}} 0.5^k$. Here N$^{/}$, N$^{//}$, N$^{///}$ are numbers for "third-party" citations, co-authors' citations and self-citations,

respectively. General citation index for a particular researcher is then $H = \sum_{i=1}^{M} N_i$, where M is the total number of a researcher's publications, N$_i$ is citation index for the ith publication.

The advantage of the suggested method for calculating the citation index is in its fundamental restriction at the expense of citations by co-authors and self-citations (on the basis of the mathematical theory of limits it is easy to prove that the second and the third addenda in the formula will have final values even with the "infinite" number of citations by co-authors and self-citations). On the one hand, there is nothing bad in self-citing; moreover, the self-citations do indicate continuity of research activities (i.e. earlier works bear importance for later ones). On the other hand, it was inadmissible that the scientometric indicators should be raised artificially (even fraudulently). The suggested method will put an end to such attempts.

It is well-known that the aim of monitoring in any field is not sheer "criticism", but facilitation in correcting a situation by pointing to weaknesses there and suggesting ways of their eradication.

Let us introduce an indicator $Z = \frac{h^2}{N}$, where h is the well-known h-index, N is total number of references to the researcher's publications. The more this indicator tends to zero (especially, if it is less than 0,1), the more thought should be given to publications quality raise.

The analysis of publishing activity of lecturers at universities in the Krasnodar territory revealed a close relation between h-index, the widely recognized indicator, and the suggested indicators for the researchers with high efficiency degree. Thanks to automated systemic and cognitive analysis of evidence on research activities results of academics, we managed to identify clearly an experimental group (n1=33) of lecturers with high and more than average h-index (15 and up) and a control group (n2=387) of lecturers with average and low h-index (under 15). The analysis of Tables 1 and 2 reveals that at high values of h-index the researchers display close correlation between all scientometric indicators,

whereas for the researchers with low and medial h-index figures this correlation does not appear strictly unambiguous. Designations: P1 is h-index, P2 is a researcher's efficiency index, P3 is a researcher's coordination number, P4 is a researcher's social valence, P5 is the number of journals with researcher's publications, P6 is the social broadcasting index of researcher's publications, P7 is the authors' index of a researcher's science citation, P8 is the authors' science citation index of a researcher's most significant paper (the strength of a correlating group is an average value of correlating coefficient between pairs of different indicators).

Table 1. Correlating coefficients between scientometric indicators in the experimental group

Parameter	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	1	0,70	0,57	0,62	0,72	0,68	0,77	0,75
P_2	-	1	0,54	0,65	0,69	0,65	0,71	0,78
P_3	-	-	1	0,82	0,67	0,62	0,55	0,61
P_4	-	-	-	1	0,54	0,65	0,68	0,53
P_5	-	-	-	-	1	0,54	0,54	0,65
P_6	-	-	-	-	-	1	0,63	0,59
P_7	-	-	-	-	-	-	1	0,79
P_8	-	-	-	-	-	-	-	1
Strength of a correlating group: 0,65.								

Source: authors

Table 2. Correlating coefficients between scientometric indicators in the control group

Parameter	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	1	0,54	0,33	0,29	0,52	0,45	0,41	0,47
P_2	-	1	0,24	0,17	0,19	0,15	0,44	0,48
P_3	-	-	1	0,56	0,47	0,42	0,35	0,41
P_4	-	-	-	1	0,54	0,37	0,58	0,43
P_5	-	-	-	-	1	0,59	0,54	0,46
P_6	-	-	-	-	-	1	0,47	0,39
P_7	-	-	-	-	-	-	1	0,59
P_8	-	-	-	-	-	-	-	1
Strength of a correlating group: 0,42.								

Source: authors

The analysis of publishing activity of efficient researchers (with h-index 15 and up) revealed that the closest correlation with the h-index have the authors' index of a researcher's science citation and the authors' science citation index of a researcher's most significant paper. The h-index also closely correlates with the number of journals to which a researcher contributed his/her publications. It can be explained by the fact, that the bigger assortment of the journals is, the sooner the science community will know about the results of the research activities of a scientist. Also, there is quite a close correlation betweenthe authors' index of a researcher's science citation and his/her social valence. This results from the fact that the availability of a stable research group of colleagues is a significant factor for obtaining tangible results(realized in publications), recognized by science community.

Thus, a high degree of strength of a correlating group for the experimental group of researchers proves the validity of h-index, as well as the fact that a highly efficient researcher is characterized by high values of all scientometric parameters. At the same time, a low degree of strength of a correlating group for the experimental group of researchers shows evidence of need for application of a whole complex of scientometric indicators (including the authors'ones) to assess research activities of those with modest experience in this area.

4. Findings

Objective assessment of a researcher's efficiency is one of the most important tasks of our time. In compliance with principles of systemic analysis and qualimetric requirements, the research efficiency should be viewed as an integral

indicator (latent variable) which can be defined by a set of criteria (indicating variables).

A set of scientometric indicators reflecting quality and efficiency of a researcher's activity is the most significant component of criteria and assessment apparatus of scientific research monitoring.

Modern information technologies provide primary information on the research results of individuals and organizations which can be used for calculating (assessment) all scientometric indicators (both traditional, and those suggested by the authors of this study). Calculation (assessment) of the authors' indicators can be implemented by PC.

The widely-recognized scientometric indicator, "Hirsch-index", enables the revealing of efficient researchers. The researchers with high and higher than average values of h-index do show the same high values in other authors' scientometric indicators, whereas the results for the researchers with medial and low values of h-index cannot demonstrate such unambiguous correlation.

5. Conclusion

Objective assessment of research activity is an extremely difficult task. And it is with good reason, that for decades now there have been ongoing attempts to resolve it. Nevertheless, within the new context of information society (informatization of all spheres of human activities) it is surely possible to face any challenges of today. The authors have no doubt that for monitoring objectification (assessment in particular) of the research activity it is necessary to develop high-tech methods of processing primary information on it. It should be based on qualimetric approach.

The authors of the article uphold the position that any sphere of activity is spoilt not by the very idea of applying quantitative indicators, but by inadequate assessment methods and systems of indicators. The historical achievement of Jorge Eduardo Hirsch, an Argentine American professor of physics, is that he was the first who succeeded in settling the discrepancy between "quantity and quality"; however, scientometrics should go further. Research activities monitoring is inconceivable without scientometric parameters, but they must be operational (individually), and the system of indicators must be functionally complete (this is the most important requirement to the system of criteria).

The results of this study enable us to produce practical advice on improving managerial effectiveness in the sphere of research activities at higher education institutions, such as: to establish at universities and their departments (faculties, chairs) databases (on the basis of Internet technologies) of most important and sound publications by outstanding colleagues among lecturers so as to outline patterns of research activities, as well as to specify career targets of young lecturing researchers.

6. Acknowledgements

References

Chebotarev, P.Y.(2013)Naukometriya: kakseyopomoshchyulechit, anekalechit? [Scientometrics: how to make use of it, but not abuse it?]. *Upravlenie bolshimi sistemami*, 44, 14-31.
Eck, N.V.,Waltman, L. (2008).Generalizing the h- and g-indices. *Journal of Informetrics*, 2(4), 263-267.
Egghe, L. (2006).Theory and practice of the g-index. *Scientometrics*, 69(1), 131-152.
Egghe, L. (2008).Mathematical theory of the h- and g-indices. *Journal of Informetrics*, 2(4), 263-271.
Eichorn, P., Yankauer, A. (1987).Do authors check their references? A survey of accuracy of references in three public health journals. *American Journal of Public Health*, 77(8), 1011-1012.
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, 203-217.
Guns, R., Rousseau, R. (2009). Real and rational variants of the h-index and the g-index.*Journal of Informetrics*, 3(11), 64-71.
Hirsch, J.E. (2005). An index to quantity an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572.
Lehmann, S., Jackson, A.D., Lautrup, B.E. (2006).Measures for measures.*Nature*, 444(21), 1003-1004.
Shtovba, S.D., Shtovba, E.V. (2013).Obzor naukometricheskih pokazatelej dlya ocenki publikacionnoj deyatelnosti uchenogo. [Scientometric parameters survey for the researcher's publishing activity assessment]. *Upravlenie bolshimi sistemami*, 44, 262-278.