The geographical latitude index of the relevance of research results activity as scientometric index

Popova Olga, Romanov Dmitriy, Drozdov Alexander, Mironova Marina

***Abstract*— The purpose of the research is to suggest and the describe a new standard for productivity of the research activity. The Hirsh index based on scree method has become popular inperforming scientists` work productivity. This index is also applied to scientific groups (organisations). However, h-index and other scientometric indexes based on citations can be artificially increased with the help of scam models. Thus, there is a necessity to develop a new standard of assessment of a research activity efficiency that will be difficult (hard) to "improve". The authors suggest geographic latitude index of the relevance of citations. Despite the fact that suggested index as h-index based on citations, its big importance indicates that scientific union accepts the results of academic staff activity. Moreover, the given index could be a standard for identifying prospective (productive) scientific workers. To achieve the objectives and tasks the following correlated methods have been used: the analysis of problem citation, the analysis of scientific references and best practice of research activity management in scientific institutions, universities (benchmarking), cognitive, structure-functional and mathematic modeling, the method of graph theory, sets and relations; automated systematic cognitive analysis; method of qualimetry (theory of latent variables); method of mathematic statistics, method of analytic geometry. The research was carried out on the basis of high schools of Krasnodar Region. It was also used methodological base of the research: system approach, metasystem approach, probabilistic statistical method, sociological and qualimetric approaches.**

***Index Terms*—scientists team, research activity, results, references, geographic latitude, necessity, diagnosis.**

I. INTRODUCTION

The diagnosis of productivity of research activity of scientists and teams is one of the most actual but at the same time complicated scientometric tasks [1-23]. The Hirsh index and different scientometric activities based on citations have instantly become popular thanks to its objectiveness and humanistic potential. Unlike such activity as an average number for publication, h-index based on scree method (widely-used method of mathematical statistics) does not "forbid" a scientific worker to publish new scientific works (i.e. publications without citations that do not reduce complex scientometric index). However, research activity diagnosis based on the citation analysis have number of risks. One of such risks is connected with a desire and opportunity to increase artificially scientometric activity as well as the h-index [2,8-10,12] known as Goodhart`s law that states:

Popova Olga, Department of Information Systems & Programming at Kuban State Technological University, Russia, +79184414463.

Romanov Dmitriy, Department of Information Systems & Programming at Kuban State Technological University, Russia.

Drozdov Alexander, Deputy Director of the Center for Strategic Development at Kuban State Technological University, Russia.

Mironova Marina, Department of Business English at Kuban State Technological University, Russia.

"When a measure becomes a target, it ceases to be a good measure". Applied to scientometric, "Hirshmania" has become one of social disasters of new century [2, 8-10,12].

Thus, scientometric as the branch of science of science as well as management practice of research activity in scientific and educational institutions are in dire need in objective measures of productivity research activity that cannot be artificially improved. **The problem of study** is to find out the objective criterias of productivity of scientists not influenced by artificial improvement. The **research objective** is to present and explain new criteria for productive diagnosis of research activity. **The target of the current study** is scientific (research) activity of scientists. The **subject of the research** is the informative value and latitude index of citation relevance (citation ratio) in scientific research results (scientific publications). To achieve **the objective** the following **tasks** are important to set up:

1. To develop calculation model for a new scientometric index-latitude index of citation relevance (citation ratio) in scientific research results (scientific publications).
2. To determine connection between latitude index of citation relevancein scientific research results and financial support of scientists` research projects on the base of actual data.
3. To single out gradation of numerical values of a the new scientometric index - latitude index of citation relevance (citation ratio) in scientific research results (scientific publications).

The solution of the above-mentioned problem, the purpose achievement and tasks solution are actual with an increase in research activity in modern world and also to support W. von Humboldt`s ideas of Universities as the core of science.

II. THE ANALYSIS OF THE DEVELOPED PROBLEM

There are two correlated ways of struggling against social disasters connected with absolutization of "selected" indexes. The first one is polyvalent diagnosis and monitoring of scientists` research activity. The second way is to create adequate monitoring indexes not influenced by (or hardly influenced) artificial "improvement". The third way is the optimal combination of the 1st and the 2nd ones.

The authors of the article and their colleagues have made an attempt to implement the 2nd way. Thus, for example, Romanov D.A. has offered citation index of specific publication of scientist:

$$Z = n_1 + \sum_{j=1}^{n_2} 0{,}75^j + \sum_{j=1}^{n_3} 0{,}5^j ,$$

where n_1 - number of external citations for a publication,

n_2 - number of citations on co-author`s publication (on scientometric basis)

n_3 - number of self-citations of a co-author.

Above-mentioned index [12,17] has been compiled on the basis of author`s approach [15] that states complex monitoring indexes are calculated on the basis of theory of limits aimed at loss of meaning of nonterminating increase of input monitoring information that can be artificially raised (both self-citation and co-author citation). Modified h-index is equal to Z. If Z scientist`s publications have citation index at least more than Z (citation index can be fractional).

Romanova M.L. has suggested in her scientific work [13] that scientific publication quality

$$K = (1+C)^2 \cdot (1+Z) \cdot s \cdot (1+D),$$

where C-impact factor of publication,

s – coefficient of publication status,

D-coefficient depended on extra information of publication.

The modified Hirsh index is equal to K. If K scientist publications have scientific quality at least more than K.

As we can see two above mentioned indexes are rather difficult to "improve" with scam models.

Undoubtedly, assessing recognition results of research activity (recognition is in citation publication of a researcher) Herfindahl index assessment (currently used in scientific journals) can be used to assess inequality citation of a scientist on the part of social environment. Herfindahl index is "harmful" index, i.e. the increase of its numerical value can deteriorate the situation. Evidently, if a scientific worker is cited by a rather limited number of people there will be high h-index. However, the authors of this article state that low h-index does not mean high recognition of a scientist`s publication (research activity results) as low h-index can be matched with low number of citations on scientists` works.

Complex scientometric index – h-index has some disadvantages. Taking its importance and relevance into account the authors underline its limitation in scientific worker`s publication assessment. Undoubtedly, the possibility in its artificial "improvement" is laid in the evils of society. However, h-index as the index based on citations cannot fully show real results of a research activity. This inference is supported by facts. What bibliometric indexes (ex. H-index) do Maxwell, Faradey, Lomonosov, Mendeleev, Archimedes (lack of the article volume doesn`t allow to give the full list) have? Can you even imagine a textbook on physics without D. Maxwell or A.S. Popov`s names, on chemistry- D. Mendeleev, M. Lomonosov or H. Cavendish, on mathematics-R. Decart or N. Lobachevskiy`sachievements?

Thus, it is important to create and describe a new objective criteria of scientists` research activity (including teachers of universities).

III. Research Organization and Methodology

Complementary methods of research, for example, the problematic situation analysis, analysis of scientific literature and advanced experience of managing research activities in scientific organizations – such as universities (benchmarking), cognitive, structurally-functional and mathematical modeling, graphs, relations and sets theory methods, automatized systematically-cognitive analysis, qualimetry methods (latent variable theory), mathematical statistics and analytical geometry methods were used in achieving our goal and problem solution. We should pay special attention to mathematical methods role in our research. Scientist's or team's research activity result importance is taken as integral index (latent variable), for which we can pick up particular criteria (indicator variables). Set, relation and graph theory methods let us create cognitive models of creative (research) team's activity, their influence on scientific community (scientific megaenvironment) in general. Analytical geometry methods, which were used to calculate the area of convex polygon (convex polygon represents the subset of cities, where the authors of citing publications live and work) turned out to be the most useful.

The research was based on Krasnodar Region's universities. With the help of qualimetry methods (latent variables theory), authors were able to pick out a new criteria of research activity rating – geographic demand (recognition) index. Thanks to Russian Science Citation Index, (website elibrary.ru) result data of primary research activity of Krasnodar Region's professors (n=420) was received. Financial support data (in form of scientific foundation grants) was received from university's scientific research managements (Kuban State Technology University), as well as open domain publications. Automatized systematic-cognitive analysis has let us find out validity and differentiating ability of author's research activity productivity index of scientists. Informational entropy is calculated by formula

$$H_I = -\sum_{j=1}^{C} \left(p_j * \ln\left(p_j\right)\right),$$

where p is a probability (relative frequency) of a scientist to appear in one or another cluster (clusterization criteria is an index, suggested by author),

C is an amount of clusters (there are 10 pointed out).

Our research methodological basis is systematic approach (takes research activity as a complex systematic process and integral part of university's functioning), metasystematic approach (takes citation publication complex of scientist as a system, which includes relatively independent components), probability-statistical (takes recognition of scientist's research activity as a stochastic process), sociological (takes science as sociocultural phenomenon and social institute and university as social system, which is the core of science, according to Humboldt's concept) and qualimetrical approach (announces the necessity of multi-criteria diagnosis of scientist's research activity productivity).

IV. Research Result

From author's point of view, we can take geographic result demand as an objective and resilient to 'perfection' criteria of research activity productivity. This means, the wider the geographical distribution of links (citations) of some scientist, the more surely we can speak of his recognition (demand) by science community. For example, if works of one scientist is cited in ten cities, separated with hundreds of kilometers, and another is cited only in one city, then undoubtedly, the first one's work is more recognized than the second one's (even if both have the same publication rate, citing amount and h-index). Obviously, geographic latitude of links (citations) should be measured not by science magazine assortment, in which the works (citation sources) are published, but by author's workplaces. Because it's possible that one author may cite another from different magazines, but we can't talk

of geographic latitude in that case (even more, false index increase may happen).

Let *q* be a set of analyzed scientist publications and *w* is a set of citations to these works, *w* is a subset of citations, which are not self-citations, then an amount of publications Q=P(q), amount of citations (not self-citing) W=P(w), where P is cardinality (further in the article we'll only mention links, which are not self-citing) Also obviously, that $w = U_{j=1}^{P(q)} w_j$, where w_j is a link set (citing publications, but not self-citing), according to a number j of analyzed scientist work. Then result demand index of analyzed scientist by geographic latitude is λ= S * η , where η=P(N). Here, P is cardinality, N is a set of localities, where the authors of citing publications live (work). S is an area of the **convex** polygon, surrounding the localities of N set (we should have such polygon, that localities of N set would be on polygon's edge or inside of it, but not outside). If we have two localities, then we take distance between them as a polygon's area. If we only have one locality, than the polygon's area is equal to locality area (For example, the area of Krasnodar is more than one hundred square kilometers).

Obviously, $N = U_{j=1}^{P(w)} n_j$, where n_j is a locality set, which consists of places where the authors of citing publication number j work and live. For example, one author of citing publication lives in Kaliningrad, another lives in Volgograd, and the third one lives in Krasnoyarsk. From author's point of view, λ index can objectively represent the scientist's publication recognition by scientific community. On one hand, if the area of imagined polygon takes millions of square kilometers, then it's unlikely, that scientist had a "deal" with all of the citing publications authors. On other hand, this index makes artificial "citing deal" of scientists, who live in the same locality, useless.

With all that, while objectively representing the results of scientist's research activity, this index doesn't take productiveness in account. For example, we can achieve the same value of λ by writing 10 works, as well as 20. By simplified scheme, the productivity index (or scientific community influence index) can be measured as $\mu = \frac{\lambda}{z}$, where z is minimal amount of scientist's publications (not the citing publications), to provide λ index. We can calculate z with this algorithm.

Step 1. Take z equal to Q.

Step 2. Reduce z by 1.

Step 3. Check if λ has reduced and go back to step 2 if it hasn't.

For example, some scientist's index λ (publications demand index by geographic latitude) is $5*10^7$, all in all he has published 120 works, but only 40 of them provide the λ value. Then the scientist's productiveness index is $\mu = 5*10^7/40 = 1.25*10^6$.

More labor-intensive (in sense calculations) is another scientist's productiveness rating algorithm. We analyze all possible combinations of scientist's publications (the amount of these combinations is huge) and we calculate the μ_j value every j numbered transaction (more specific, publication combination) (obviously, not only z, but λ would vary), then $\mu = \max\{\mu_j\}$.

The presented indexes are also fit for rating the demand (recognition) of research activity result of research teams (including educational research teams of universities). However, instead of one scientist's publication set we analyze research team members publication set. Research team publication set is $q' = U_{j=1}^{M} q_j$, where M is amount of research team members, q_j is publication set of j team member, U is set union symbol.

For rating the informational content and validity of suggested scientometric index of demand (citation) of research activity results by geographic latitude, the authors researched the correlation between its numeric value and funding of scientific researches (research projects). From author's point of view, scientist's or team's research project funding (for example, scientific foundation grant, crowdfunding, etc.) represents the appropriate level of their research activity. Undoubtedly, funding system (especially the granting one) has drawbacks (as well as scientific prize system), but it allows to preserve the traditions of research (scientific) activities.

By analyzing the indexes, which represent the recognition of research activity results of Krasnodar Region's scientists (faculty of universities), we were able to differentiate scientists by index of demand (citation) of research activity results by geographic latitude λ (chart 1).The research was done according to April 2016 data. By taking together relative frequencies of scientists (probability of scientist being in one or another cluster) we have an entropy equal to 2,176; The most effective scientist (by authors criteria) is $1,73*10^5$ times more effective than the least effective one.

Chart 1. Demand index of research activity results by geographic latitude of Krasnodar Region scientists.

Cluster №	Demand index	Number of scientists	Their percentage
1.	More than 10^8	9	2.142857
2.	From 10^7 to 10^8	19	4.52381
3.	From $5*10^6$ to 10^7	27	6.428571
4.	From 106 to $5*10^6$	32	7.619048
5.	From $5*10^5$ to 10^6	37	8.809524
6.	From 10^5 to $5*10^5$	40	9.52381
7.	From $5*10^4$ to 10^5	51	12.14286
8.	From 10^4 to $5*10^4$	67	15.95238
9.	From $5*10^3$ to 10^4	79	18.80952
10.	Less than $5*10^3$	59	14.04762

All in all, there were 23 financed (by one or another foundation) research projects, where 127 scientists of Krasnodar Region were taking part, each of them had different demand of research activities index, that's why a geographic latitude rating for a team is needed. Actual data analysis showed that from 23 teams, 5 had the highest demand of research activities result index level (more than 10^8), 16 had very high level (from 107 to 10^8) and 2 had high level (from $5*10^6$ to 10^7).

By analyzing the results of this research, we can see that a high value of demand index of research activity results by geographic latitude is a necessary, but not sufficient condition of research project (scientific research) funding. In other words, if we take scientific community as a megaexpert and

funding committee as a macroexpert, we should note that only those results, that are recognized both by megaexpert and macroexpert may be considered high quality.

Geographic latitude method allows us also to determine recognition of specific scientist's publication by scientific community: $\chi = 10^{-6} * Z * S_{pub} * \eta_{pub}$, where Z is publication's citing index (according to Romanov's method), S_{pub} is area of imagined convex polygon, which connects localities of where the citing publications (link source of analyzed publication of the current scientist) authors live and work. η_{pub} is an amount of localities (they can also be inside of imagined polygon). Correction factor is introduced to normalize recognition index. Then value L of productivity index is equal to χ, if no less than χ publications of analyzed scientist have recognition index not less than χ each.

We should also note that, magazines and scientific community do not always share the same opinion. It is possible that an article may be published in a prestigious magazine with high ratings (impact factor), but it doesn't get (according to citing) much recognition by scientific community.

It is also possible that the article is published in a magazine with not so high impact factor, but it gets recognized by scientific community, gets cited from different country regions (if we mean national scientometric basis of a large country) or countries (if we mean international scientometric basis, such as Web of Science, Scopus, Agris, DOAJ, EBSCO, etc.).

According to authors` opinion L possible gradations quantity are the following: less than 5 is the lowest level, from 5 to 9-low level, from 10 to 19- medium, from 15 to 19 is the right level (above average), from 20 to 24 is high level, from 25 to 29-very high level, above 29 is the highest level.

Within the given article the authors consider an appropriate to give practical recommendations that can increase the effectiveness of research activity. First of all, it is important to build information portal for keeping publication data of high quality (publications on pdf-format as a sample!) as well as the most productive scientific workers. The importance of making such portals is to provide targets for professional growth and sample results of research activity of scientific workers. But more important thing is the availability of information resources that can allow scientific-pedagogic officers to use research activity results of their colleagues (among with their own results!) in education content.

V. Conclusion

The geographical latitude index of the relevance of citations is a standard that allows to single out high-effective scientific workers. This index is difficult to improve artificially.

VI. Summary

1. It is necessary to introduce a new scientometric index for scientific productivity of researchers to avoid its artificial improvement.

2. The geographical latitude index of the relevance of citation suggested by the authors is a standard of research productivity assessment of a scientific worker that considers new recognition aspects of its results. A metasystem approach is the most important methodological basis for assessing this index as the authors of cited publications live and work in areas which represent a metasystem. A system with independent components, mathematical framework – the theory of sets and analytical geometry as well.

3. There is clear connection between the author's recognition of the research and the funding of the scientific projects that proves the validity (authenticity) and the informative character of the author's index.

The high level of geographical latitude index and relevance of the research results is a reflection of the proper level of scientists research work and groups and this, in turn, is a main factor for funding research projects, grants and etc.

4. The assessment of the scientists research work efficiency must be versatile taking into account different aspects of how these research work result are recognized by the Scientific Community.

The Hirsh index and the latitude index of the relevance of research work results are interrelated. They are used to monitor the scientists' research projects.

Acknowledgment

The work is carried out with the financial support of Russian Scientific Humanitarian Foundation as part of the research projects 13-06-00350 "The modern information educational environment" and 16-03-00382 "The research work monitoring of educational institutions in information society.

References

[1] Balykhin, M. G., Generalova, A. V. Crowdfunding as a platform for supporting research and development. *Sotsiologicheskie Issledovaniya* – 2015, *Is.*9, pp. 57-61.

[2] Gavrilova, E. V., Ushakov, D. V., Yurevich, A. V. Translation of scientific experience and tacit knowledge. *Sotsiologicheskie Issledovaniya – 2015, Is.*9, pp. 28-35

[3] Geidarov, P. S. The scoring system evaluation of academic works and electronic academic seminar. *Sotsiologicheskie Issledovaniya - 2015, Is.*4, pp. 162-167.

[4] Grebennikova, V.M. Education in modern society: the communicative aspect. *Izvestiya Southern Federal University. Pedagogical sciences* - 2010, *Is.*3. - pp. 54-59.

[5] Dudina, V. I. Sociological knowledge in the context of information technologies development. *Sotsiologicheskie Issledovaniya – 2015, Is.*6, pp. 13-22.

[6] Covey, S.R. Seven skills of highly effective people: Powerful tools for personal development. - Moscow: Alpina Pablisher, 2016. - 396 p.

[7] Kovshov, D.K. Building a socio-psychological map of the team according to sociometric research. *Sotsiologicheskie Issledovaniya – 2015, Is.*10, pp. 195-209.

[8] Lazarev, V.S. Research of the pedagogical collective as a subject of innovation activity. Questions of Psychology - 2015, *Is.*1, pp. 87-97.

[9] Lebedeva, I. P. Soft models as a form of the mathematization of social knowledge. *Sotsiologicheskie Issledovaniya – 2015, Is.*1, pp. 79-84

[10] Loiko, V.I. Modern models and methods for diagnosing the research activities of scientific and pedagogical teams. *Polytechnical network electronic scientific journal of the Kuban State Agrarian University* –2015, *Is.*112, pp. 1906-1933.

[11] Lutsenko, E.V. Hirschmania in assessing the results of scientific activity, its negative consequences and the attempt to overcome them with the application of a multicriteria approach and information theory. *Polytechnical network electronic scientific journal of the Kuban State Agrarian University* –2015, *Is.*108, pp. 812-832.

[12] Romanov DA, Popova OB, Nosova Yu.S. Modern methods for assessing the productivity of research activities. *Polytechnical network electronic scientific journal of the Kuban State Agrarian University* - 2015, *Is.*108, pp. 725-739.

[13] Romanova, M.L. Reflection of scientific knowledge in the content of higher education. *Scientific works of the Kuban State Technological University* - 2016, *Is.*3, pp. 372-393.

[14] Tolstova, Y. N., Voronina, I. D. On the necessity to widen the notion of sociological measurement. *Sotsiologicheskie Issledovaniya - 2012, Is.*7, pp. 67-77.
[15] Fedorova NP, Tyupenkova GE, Kiseleva ES, Romanov DA, Nikulina O.N. Modern ways of forming monitoring indicators. *Scientific works of the Kuban State Technological University* – 2015, *Is.*11, pp. 266-292.
[16] Christiansen J.A. Building the innovative organization: Management systems that encourage innovation. – New York: St. Martin's Press, 2000. – 357 p.
[17] Popova O.B., Popov B.K., KarandeiV.Yu., Romanov D.A., Kobzeva S.A., Evseeva M.A. (2015) "New methods and evaluation criteria of research efficiency" Mediterranean Journal of Social Sciences. 2015. T. 6. № 6.C. 212-217.
[18] Eck N.V., Waltman L. Generalizing the h- and g-indices. *Journal of Informetrics*. – 2008, Vol. 2, No 4, pp. 263-267.
[19] Egghe L. Mathematical theory of the h- and g-indices. Journal of Informetrics – 2008, Vol. 2, No 4, pp. 263-271.
[20] Franceschini F., Maisano D., Perotti A., Proto A. Analysis of the ch-index: an indicator to evaluate the diffusion of scientific research output by cites. Scientometrics – 2010, Vol. 85, pp. 203-217.
[21] Guns R., Rousseau R. Real and rational variants of the h-index and the g-index. *Journal of Informetrics* – 2009, Vol. 3, No 11, pp. 64-71.
[22] Hirsch J.E. An index to quantity an individual's scientific research output. Proc. Nat. Acad. Sci. USA – 2005, Vol. 102, No 46, pp. 16569-16572.
[23] Jonash R.S., Sommerlatte T. The innovation Premium: How next generation companies are achieving peak performance and profitability. – Cambridge, Massachusetts, 2000. – 151 p.

First Author Olga Popova is an Assistant Professor in the Department of Information Systems & Programming at Kuban State Technological University in Krasnodar, where she has been a faculty member for 16 years. All this time, along with teaching, she has been doing extensive research and writing a Doctorate Thesis. Olga is the author of 90 publications on various issues in system analysis which are rated in Russian Science Citation Index (RSCI) where her h-index is 11. Completed her undergraduate studies at Kuban State Technological University with honors, receiving a diploma in electrical engineering, being the top graduate in her class, and in 2002 she received Ph.D. in technical sciences from the same university. Her thesis covered mathematical modeling and optimization of special electromechanical systems. Research interests lie in the area of complex system analysis with focus on Decision Support Systems (DSSs) and information structuring. In recent years, she has been exploring artificial intelligence enhancement methods approximated to those used by natural intelligence.

Romanov Dmitriy At the present time he is completing his doctoral thesis on "Methodological and pedagogical system for monitoring the quality of vocational education". He took part in two research projects supported by the Russian Humanitarian Scientific Foundation (from 2011 to 2015). He is deputy chairman of the section of fundamental disciplines and independent work of students of the Methodological Council of Kuban State Technical University, as well as an expert of educational publications of the Federal Institute of Education Development. It differs in breadth of scientific interests: the theory and methodology of vocational education, science of science, modeling of socio-economic processes, etc. The motto of the research activity is the famous words of I. Kant: "In every science there is as much truth as there are mathematicians in it".