ISSN 0005-1055, Automatic Documentation and Mathematical Linguistics, 2019, Vol. 53, No. 4, pp. 179–188. © Allerton Press, Inc., 2019.
Russian Text © The Author(s), 2019, published in Nauchno-Tekhnicheskaya Informatsiya, Seriya 2: Informatsionnye Protsessy i Sistemy, 2019, No. 8, pp. 28–38.

Comparison of Research Results by Scientific Communities

P. A. Kalachikhin*
All-Russian Institute for Scientific and Technical Information, Russian Academy of Sciences (VINITI RAS), Moscow, 125190 Russia
**e-mail: pakalachikhin@viniti.ru*

Received May 5, 2019

Abstract—In the article, the issues of normalization of scientometric indicators of the publication level were studied. An integrated approach to the evaluation of research results was formalized. The similarities and differences between professional and scientific communities were considered. The concept of a professional scientific community was introduced. Local and network subtypes of scientific communities were identified. A method has been developed for obtaining the values of scientometric indicators normalized by both local and network scientific communities, the distinguishing feature of which is the use of linguistic variables and production rules of fuzzy logic. The testing of the proposed methodology was carried out based on the example of a scientometric database.

Keywords: altmetrics, bibliometric indicators, scientometric platforms, scientific community, professional community, network community

DOI: 10.3103/S0005105519040071

INTRODUCTION

In the context of global world processes, research centers are dispersed and financed unevenly. There is a tangible difference in the cost of funding research within individual countries, for example, in leading scientific powers and developing countries. This implies a significant difference in the levels of scientific potential in different parts of the globe. In order to reduce the influence of this factor on the development of science, when conducting a comparative analysis of the results of scientific research, in addition to the absolute indicators of the level of development of science, such relative indicators are used as the share of expenditure on science in the gross domestic product. Rationing of indicators for assessing the results of scientific research allows objective comparison of them.

A characteristic trend for modern science is its internationalism. New international projects are emerging, as well as projects executed with foreign support. While earlier, besides the scientific and educational organization, the state in which the research was conducted was considered as the business card of a publication, today, although country affiliation is still of great importance, new criteria are gradually being formed. Among them is the affiliation of the authors of the study to one or another community. Here the concept of "community" should be interpreted in the broadest sense, down to the group of users of social networks.

Reaching a consensus on the scientific significance of some work requires a mutual comparison of research results carried out in different environments and, as a rule, unequally scientifically recognized communities. Such a comparison is proposed to be carried out on the basis of a method whose distinctive feature is the use of fuzzy logic and fuzzy sets. The use of these mathematical tools should provide the adequacy and reliability of the comparison of research results by scientific communities.

Our proposed method of comparative analysis of research results, whose authors belong to different scientific communities, has methodological significance for selective examination of research works carried out in the framework of competitions, reporting, awarding prizes, and awarding academic ranks and degrees.

FORMALIZATION OF APPROACHES TO THE ASSESSMENT OF THE RESULTS OF SCIENTIFIC ACTIVITIES

The 20th century was characterized by the development of the metric sciences, which include scientometrics. Recent developments in this area include wikimetry, altmetry, and librametry [1].

In connection with requests from the state apparatus, as well as its various structures and departments responsible for the state of national science, an area of scientometry has been actively developing that explores the economic aspects of scientific activity.

Financing of scientific research requires the improvement of scientometric methods and criteria for the examination of research results. Such criteria should allow evaluation of the effectiveness of the expenditure of funds allocated to finance science.

A financial approach to evaluating research results implies:

$$\frac{In}{C} \to \max, \quad (1)$$

where *In* is income and *C* is costs.

However, many experts directly involved in scientific activities are guided by a bibliometric approach:

$$\begin{cases} NP \to \max \\ NC \to \max \end{cases}, \quad (2)$$

where *NP* is the number of publications and *NC* is the number of citations.

In turn, critics of the "pure" bibliometrics suggest evaluating the results of scientific research in a different way:

$$QR \to \max, \quad (3)$$

where *QR* is the quality of research.

Investors, as a rule, are guided by innovative considerations for evaluating the results of scientific research:

$$IP \to \max, \quad (4)$$

where *IP* is the innovative potential.

Thus, when evaluating the results of scientific activity, for managers, one set of indicators are important, for the researchers themselves, others are important; therefore, it is advisable to use an integrated approach, the essence of which can be expressed as follows:

$$Im \to \max, \quad (5)$$

where *Im* is the impact.

In formula (5), the *impact* should be understood as the cumulative effect obtained as a result of the influence produced or the effects on various spheres of life.

Why do statistics not work when assessing the effectiveness of scientific research? As a rule, these are macro-level indicators that are not suitable for the evaluation of smaller objects [2–5].

Some authors for each direction propose to use separate scientometric databases, citation standards, including those over time. These indicators can be used to compare similar assessment objects in more or less similar areas. The reduction of the indicator value to a value within a specified range is carried out using a conversion, which is usually called rationing or normalization. Normalized indicators can be used to determine whether an object shows a result above or below the average level in relevant areas [6]. The disadvantage of normalized indicators is that they involve the maintenance of special tables that need regular revision.

CRITERIA FOR THE EVALUATION OF RESEARCHERS, ORGANIZATIONS, AND JOURNALS BY THE PROFESSIONAL SCIENTIFIC COMMUNITY

A professional community should be understood as an association that has historically been established and institutionalized as a relatively self-regulating association of professionals with a specific ethical code that is relatively independent of public opinion or formal and informal norms that are strictly adhered to [7]. In the modern world, the professional community is a nonprofit and non-governmental organization that unites people based on the same interests or needs [8].

The scientific environment is not just a set of scientists, including those with the highest achievements in research activities. It includes scientific foundations that fund research projects and specialized publications, which should be considered as an arsenal of opportunities for a scientist to broadcast the results of their own research. It is possible to diagnose the involvement of a scientist in the scientific community based on the analysis of their publication activities. The use of the potential of a social environment by a scientist is a prerequisite for productive research and obtaining good scientific results [9].

A professional community can be informal, but only to a certain extent, since the usual audience of readers of scientific literature, as well as users of electronic scientific libraries and scientometric databases cannot be called informal. At the present stage of the development of science, the professional community is still an elite, including a narrow circle of people on whom the program, financial, and legal support of state policy in the scientific sphere directly depends [10]. As an example, these are various expert groups and commissions, including:

- The National Higher Attestation Commission;
- free associations (like the Russian Association for Artificial Intelligence—RAII);
- national scientific communities and organizations (such as the Royal Society of London for Improving Natural Knowledge);
- expert mega-groups.

That is how the project "The Corps of Experts in Natural Sciences" was initiated in 2007 to promote the use of bibliometric indicators as a private tool in a well-thought out system of searching for and selecting scientific experts, as well as organizing assessments and competitive procedures [11].

It seems more appropriate to separate professional and expert scientific communities, although they may merge or intersect somewhere. In the latter case, we should talk about the *professional scientific community*. The recognition of a researcher as a member of the professional scientific community can be expressed in the form of incentives and can take the form of:

- appointment of additional scholarships;

• selection as a corresponding member of the Academy of Sciences;

• state and international awards;

• assignment of honorary positions, titles, etc.

Alternative metrics (altmetrics) are non-traditional scientometric indicators, which are evaluated based on the functionality of web-platforms. Indicators of conventional altmetrics based on social networks and indicators derived from them cannot and should not be regarded as the opinion of the professional scientific community. A research front was formed around the questions about the use of altmetrics to determine the measure of social influence of academic information. In other words, altmetrics became an actual topic; publications on it are actively cited [12].

The attractiveness of the use of altmetrics tools at the management level is due to the possibility of taking the social effect of research results and popularization of scientific knowledge among the general public into account [13].

The main altmetrics generators are:

• *Altmetric*—a service that offers both paid services to corporate subscribers and free ones for individual users;

• *Impactstory*—a free service for individual scientists that collects information in *Google Scholar*, *ORCID*, *Figshare,* and *GitHub*;

• *Plum Analytics*—a service where the widest range of both assessment objects and altmetrics is presented [14].

It should be noted that not only researchers—authors of works, but also specific results of scientific research, scientific achievements, from applications that have passed the competitive selection to discoveries that received the Nobel Prize, are awarded. At the same time, it is important to know in which organization a particular researcher worked, which country they represented, etc. Accordingly, statistics on organizations, countries, and other summary features are calculated. At the same time, various prizes, awards, and regalia are not equivalent. Thus, in the Russian Federation, a doctoral degree is valued above the degree of a candidate of science and cash payments in the form of bonuses, awards, and scholarships differ in size. Despite the fact that most studies have a source of funding, not every result receives special recognition in the form of awards, prizes, etc. Apparently, the professional community is not able to respond to all publications published in a particular field over a certain period of time.

Researchers use journal impact factors as tools in planning a publication to evaluate the likelihood of gaining benefits from placing it in a journal with a high impact factor without the risk of rejection [15].

The so-called "gray" scientific literature transmits some of the knowledge that is not only complementary to international journals, but has an independent value. The content of "gray" literature finds its way into the mainstream of knowledge through secondary and highly specialized sources, as well as through new search methods generated by computer technology [16]. Some of the studies published in the "gray" literature, cause public interest, which can be traced using altmetrics.

The *Altmetric Attention Score* rating of the *ResearchGate* scientometric platform is calculated using an automated algorithm and is a weighted total of the number of citations of a research result. This rating is useful in that it makes it possible to evaluate the results of research on the basis of attention to them, but it does not allow us to say anything about the quality of the results, although the discussions associated with it can help [17].

Due to the heterogeneity of groups of users of social networks, altmetrics assess the social, applied, or educational significance of a publication, but not the fundamental significance [18].

All other research results are only reviewed and are mentioned in the research reports. Clearly, such results constitute the majority, but their importance also needs to be somehow assessed using scientometric indicators, such as Hirsch-like indicators of citations, as well as indicators derived from them [19].

A METHOD OF COMPARATIVE ANALYSIS OF RESEARCH RESULTS

The metadata of the results of scientific work are characterized in accordance with the field of science and research format. Fields of science are grouped according to the classifier built into the scientometric platform. There are various ways to systematize the sciences; however, it is problematic to choose a single standard that is generally accepted by the professional scientific community among them. The research format is a type of result of intellectual activity supported by scientometric platform. The result of a study may be in the form of a publication specifying existing subtypes, they may be a patent, and they may also take a different form depending on the capabilities of the platform, as well as the structure of the database and its contents.

The external environment in which, due to circumstances, a researcher has to work has a significant influence on the choice of the community. This may be a scientific school, localized in a certain place, or some kind of infrastructure that creates opportunities for effective scientific activity in a certain territory. In addition to the geographical location, the community can be located at the network address of a resource that provides services for scientific activities, for example, at the network address of scientometric platforms that operate on the principle of "*Web 2.0*" (Science 2.0). Researchers are beginning to master *wiki*, blogs, and other *Web 2.0* technologies, going

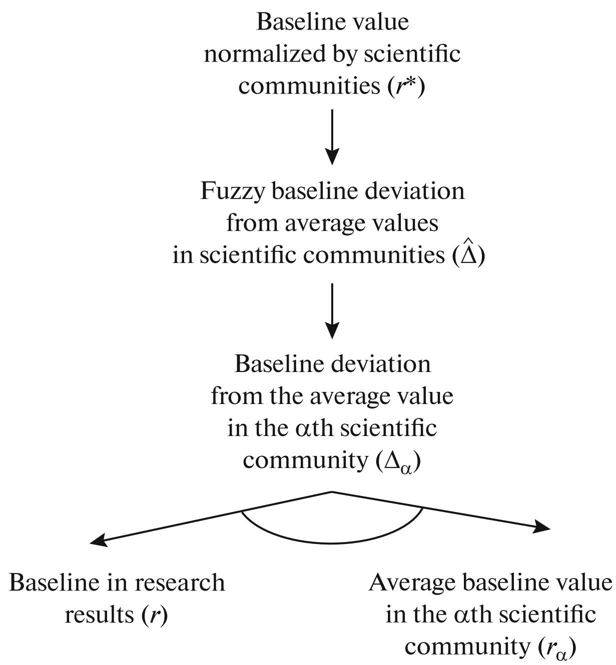

Fig. 1. An objective tree for comparing indicators by scientific communities.

beyond the usual research experience [20]. The scientific contribution of a study is expressed through an assessment of the significance of the results, taking the amendments to the environment in which these studies were carried out into account. Thus, the characteristics of a scientific community can be determined by the task of the science field and a certain address in the form of geographical coordinates or an electronic link.

Let us introduce an additional dimension z, the attitude of some scientific community, as borders of which any indicators can be taken, that is, logical ones, ordered lists, etc. To assess the results of research, a universal set R of all existing indicators is set, equipped with methods for performing calculations, describing the requirements for input and output data formats, as well as other explanatory information.

It is more convenient to use local communities Z_{geo}, built on the geographical principle: city–region–country. In the conditions of the information society, scientific activities go beyond the borders of one state and often becoming international. Articles published in collaboration with representatives of an international scientific team have the greatest chances of being highly appreciated by the professional community [21].

The local scientific community, as a rule, is represented by many researchers whose affiliation of publications points to the same locality or territorial entity. In the era of the development of scientific communications, a local community Z_{geo} can be replaced by an "online community" Z_{net}, i.e., a group of users of a web resource united by a common interest; this resource is endowed with the necessary functionality for social communication. The network scientific community is an association of users of a web resource built in accordance with the concept of *Web 2.0* and the ideology of Science 2.0. Researchers can communicate on the basis of political, moral, cultural, economic, or other interests [22].

The indicator $r \in R$ selected as baseline may be very different. However, we must understand what objects are subject to assessment. In this case, the results of scientific research are evaluated, i.e., publications. It is possible to build indicators R based on rankings for bibliometric indicators (for local communities), altmetrics (for network communities), and various indicators derived from them, which are widely studied and reflected in the literature on scientometrics [23]. If we add some "equalizing" ("rationing" or "normalizing") corrections to them, then we obtain a new indicator r^*, on the basis of which it is possible to draw conclusions about the effectiveness of the research in relation to the level of the nearest scientific community.

We will build the methodology for comparing scientometric indicators for scientific communities guided by the scheme shown in Fig. 1, starting at the bottom elements.

We propose to calculate the deviations of the values of indicators of a scientific research result from the average value of this indicator in the scientific community z:

$$\Delta = r - r_z, \tag{6}$$

where Δ is the baseline deviation; r is the baseline value relative to the research result; and r_z is the average value of baseline r in the scientific community z.

In this case, it does not matter whether the scientific community is built on a geographic basis or is virtual. Deviation can be positive ($\Delta > 0$) or negative ($\Delta < 0$), i.e., better or worse.

There may be several near scientific communities. If we consider the scientific community as a multitude of individuals joining it, then the intersection of these communities is possible. In addition, in some cases, one community may be nested in another, which is indicated by the sign $\subset$. Due to the fact that scientific communities are interconnected, they need to be somehow united.

In particular, the "part of the world" of a community z_{pow} is represented by nested country z_{coun}, regional z_{reg} and city z_{city} communities; here, $z_{\text{city}} \subset z_{\text{reg}} \subset z_{\text{count}} \subset z_{\text{pow}}$. In turn, the network community z_{soc} of all users of a social science network include limited user groups by interest z_{com} for this resource and $z_{\text{com}} \subset z_{\text{soc}}$.

Using the mathematical apparatus of fuzzy sets, it is possible to formally define inexact and ambiguous concepts. The theory of fuzzy sets includes the concept of a linguistic variable introduced by Lotfi Askar

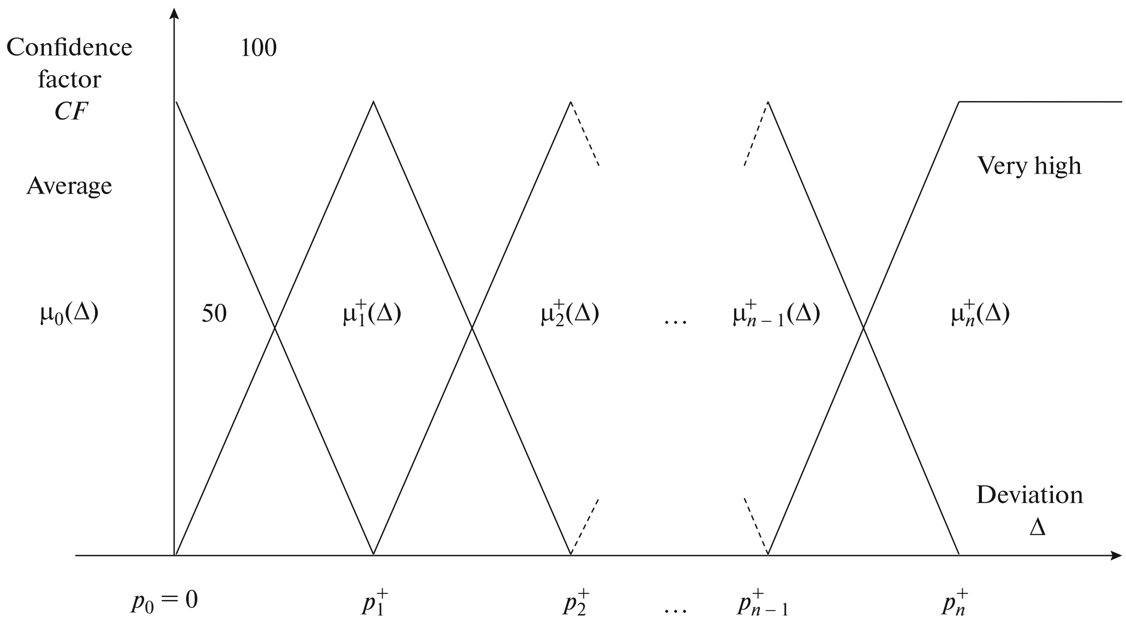

Fig. 2. A fuzzification diagram for conversion of φ_+.

Zadeh. Its values words or sentences of a natural language containing fuzzy values [24].

Rating scale of the linguistic variable $\hat{\Delta}$ of deviation of value of scientometric indicator r from the average level r_z in the community z in general form consists of n levels:

$$\hat{\Delta} = \{A_1 CF_{A_1}, A_2 CF_{A_2}, \ldots, A_n CF_{A_n}\}, \quad (7)$$

where A_i is the name of the i-th rating scale; CF_{A_i} is the confidence factor of belonging to the i-th scale of assessment of the linguistic variable of deviation of the scientometric indicator r from the average level r_z for the community z; $i = 1, \ldots, n$

To obtain the linguistic variable $\hat{\Delta}$, first the deviations Δ_z in the z-th communities are fuzzified into linguistic variables:

$$\hat{\Delta}_z = \begin{cases} \Delta_z \geq 0, \varphi_+(\Delta_z) \\ \Delta_z < 0, \varphi_-(\Delta_z) \end{cases}, \quad (8)$$

where $\hat{\Delta}_z$ is the linguistic variable of deviation of the scientometric indicator r from the average value r_z in the community z; Δ_z is the deviation of the scientometric indicator r from the average value r_z in the community z; φ_+ and φ_- are the *fuzzification* conversion.

Fuzzification φ is carried out on the basis of production rules, where the membership functions IF $\Delta_z \geq p_i \wedge \Delta_z < p_{i+1}$, THEN $\hat{\Delta}_z = \{A_i\mu_i(\Delta_z), A_{i+1}\mu_{i+1}(\Delta_z)\}$, have a triangular form. In addition, production rules may look like μ_i. At the same time, for the linguistic variable estimated by IF $\Delta_z \geq p_n$, THEN $\hat{\Delta}_z = \{A_n CF^{\max}\}$, scales, it is necessary to set n parameters $p_1, p_2, \ldots, p_n$ (by default believing that $p_0 = 0$), which act as boundaries of intervals in fuzzification diagrams, as can be seen in Figs. 2 and 3. Thus, for the boundaries of the intervals $p_1^+, p_2^+, \ldots, p_n^+, p_{+\infty}^+, p_1^-, p_2^-, \ldots, p_n^-, p_{-\infty}^-$, we obtain $(n_+ + 1) + (n_- + 1) = n_+ + n_- + 2$ production rules of fuzzy logic.

Positive values of Δ corresponds to fuzzification conversion φ_+, set in a positive direction.

To obtain the boundaries p_i^+ of the intervals of fuzzification diagram in Fig. 2, one can use the universal formula:

$$p_i^+ = d^{i-1}, \quad (9)$$

where p_i^+ is the i-th interval boundary of the fuzzification diagram in the positive direction; $d \in \mathbb{N}$; $i = 1, \ldots, n_+$.

Thus, at the output of the conversion of φ_+, we obtain the following linguistic variable:

$$\varphi_+(\Delta_z) = \{\ldots, A_0\mu_0(\Delta_z), A_1\mu_1(\Delta_z), A_2\mu_2(\Delta_z), \ldots, A_{n_+}\mu_{n_+}(\Delta_z)\}, \quad (10)$$

where φ_+ is the fuzzification conversion in the positive direction; Δ_z is the baseline deviation from the average value in the community z; A_i is the i-th level of the scale of linguistic variable values; μ_i is the membership function of the i-th level of the linguistic variable values; n_+ is the number of intervals in the fuzzification conversion diagram φ_+; and $i = 0, \ldots, n_+$

Negative values of Δ corresponds to fuzzification conversion φ_-, set in the negative direction. Accord-

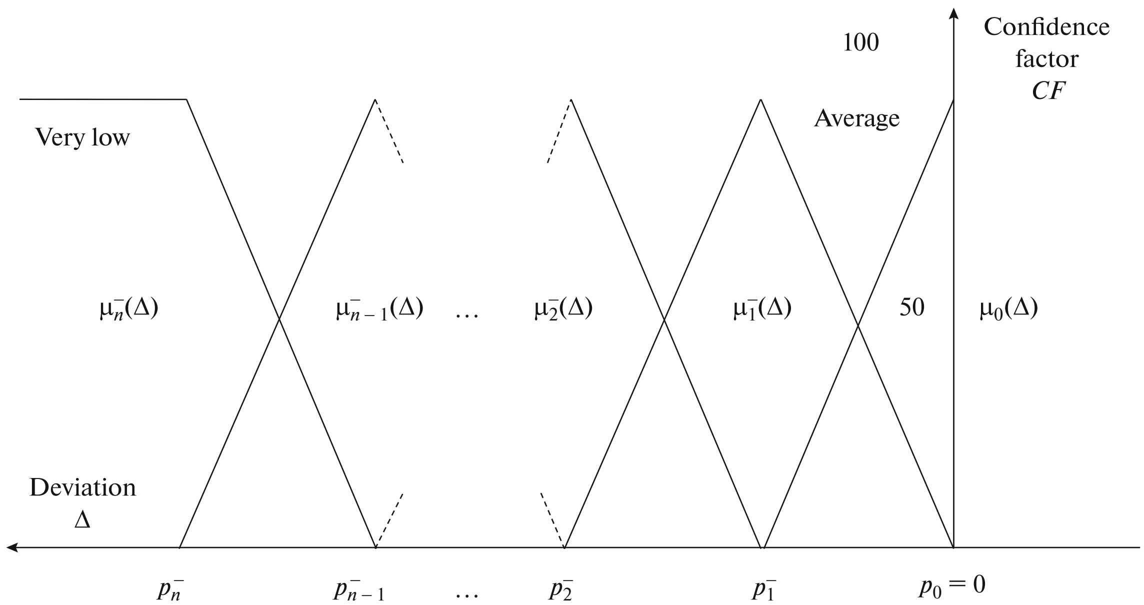

Fig. 3. A fuzzification diagram for conversion of φ_{-}.

ingly, to obtain the boundaries p_i^- of the intervals of fuzzification diagram in Fig. 3, one can use the universal formula:

$$p_i^- = -d^{i-1}, \tag{11}$$

where p_i^- is the i-th interval boundary of the fuzzification diagram in the negative direction; $d \in \mathbb{N}$; $i = 1, \ldots, n_{-}$.

At the output of the conversion of φ_{-}, we obtain the linguistic variable:

$$\varphi_{-}(\Delta_z) = \{A_{n\_}\mu_{n\_}(\Delta_z), A_{n\_-1}\mu_{n\_-1}(\Delta_z), \ldots, A_0\mu_0(\Delta_z), \ldots\}, \tag{12}$$

where φ_{-} is the conversion of fuzzification in the negative direction; Δ_z is the baseline deviation from the average value in the community z; A_i is the i-th level of the scale of linguistic variable values; μ_i is the membership function of the i-th level of the linguistic variable values; n_{-} is the number of intervals in the fuzzification transformation diagram φ_{-}; and $i = 0, \ldots, n_{-}$.

Next, the linguistic variables are merged, for which purpose the operation $\otimes$ of fuzzy multiplication is selected as the objective merge function. Thus, the linguistic variable $\hat{\Delta}$ may include the combination of linguistic variables $\hat{\Delta}_1, \hat{\Delta}_2, \ldots$ of the nested communities z_1, z_2, ... :

$$\hat{\Delta} = \otimes \hat{\Delta}_{\alpha}, \tag{13}$$

where $\hat{\Delta}$ is the integrated linguistic variable of the baseline deviation from the average values in scientific communities; $\hat{\Delta}_{\alpha}$ is the linguistic variable of the baseline deviation from the average value in the scientific community; and $\otimes$ is the multiplication operation of linguistic variables.

Here, the multiplication operation $\otimes$ of linguistic variables is set in the following way:

$$\hat{X} \otimes \hat{Y} = \left\{ A_i \frac{CF_X^{A_i} CF_Y^{A_i}}{CF^{\max}} \right\}, \tag{14}$$

where $\hat{X}$ and $\hat{Y}$ are linguistic variables; $\otimes$ is their multiplication operation; A_i is the i-th level of the linguistic variable values; $CF_X^{A_i}$ and $CF_Y^{A_i}$ are the confidence factors of belonging to the i-th level of values of the linguistic variable X and Y; $CF^{\max}$ is the maximum possible value of the confidence factor, as a rule $CF^{\max} = 100$; and $i = 1, \ldots, n$.

However, a result is possible when all the output factors CF will be equal to zero. In this case, as an exception, the linguistic variable $\hat{\Delta}$ should include the union of linguistic variables $\hat{\Delta}_1, \hat{\Delta}_2$, ... nested communities z_1, z_2, using the criteria min:

$$\hat{\Delta} = \min_{\alpha}\left(\hat{\Delta}_{\alpha}\right), \tag{15}$$

In formula (15) the pessimistic criterion min is chosen as the most rational.

For the linguistic variables $\hat{A}$ and $\hat{B}$ we set the order relation:

$$\hat{A} > \hat{B} \Leftrightarrow \phi(\hat{A}) > \phi(\hat{B}), \tag{16}$$

where $\hat{A}$ and $\hat{B}$ are linguistic variables and ϕ is defuzzification conversion.

In order to improve the accuracy of calculations in the event that there is a significant difference in the values of linguistic variables of nested communities, we use the *normalization* operation, denoted as $\| \ \|$:

$$\left\|\hat{\Delta}\right\| = \|CF_1, CF_2, \ldots, CF_n\| = \left\{\frac{CF^{\max}CF_1}{\Sigma_i CF_i}, \frac{CF^{\max}CF_2}{\Sigma_i CF_i}, \ldots, \frac{CF^{\max}CF_n}{\Sigma_i CF_i}\right\}, \quad (17)$$

where $\hat{\Delta}$ is the linguistic variable of deviations of the indicator r from the average values r_{z_1}, r_{z_2}, ... in the communities z_1, z_2, ... nested in z; CF_i is the confidence factor of belonging to the ith level of the scale of the linguistic variable values, while CF_i takes a value in the interval $[0, CF^{\max}]$; n is the number of levels in the scale of linguistic variable assessment; $CF^{\max}$ is some constant, as a rule $CF^{\max} = 100$; and $\| \ \|$ is the linguistic variable rationing conversion.

Defuzzification of the linguistic variable $\hat{\Delta}$ can be carried out in various ways, for example, by the center of gravity, median, and center of maxima methods [25].

As a result, one can obtain the indicator r^* of "assessment in the eyes of the scientific community" z of some result of a study, to determine which it is necessary to carry out rationing and defuzzification:

$$r^* = \phi\left(\left\|\hat{\Delta}\right\|\right), \quad (18)$$

where r^* is the value of the baseline r normalized by scientific communities; $\hat{\Delta}$ is the linguistic variable of the deviation of the baseline r from average values in scientific communities; $\| \ \|$ is the linguistic rationing operation; and ϕ is the defuzzification conversion of the linguistic variable.

The method we proposed is quite flexible, but the accuracy of the results will be higher if we compare publications published in the same year and related to the same type of scientific community, while the baseline is identical. This will make it possible to make the indicator r^* more sustainable over time, as with the appearance of new publications in the same field, even if the value of r remains, the indicator r^* is subject to change.

In this way, it is possible to compare the assessments of various scientific results in different communities, adjusted for the scientific level of each community. Moreover, it is possible to compare local and network communities with each other. However, this is not a very transparent method, since some researchers may not be familiar with the theory of fuzzy sets. In addition, with respect to networked scientific communities, the developed methods and models lead to some technical difficulties, since the built-in functionality of existing platforms for scientific communication, unlike the *Web Of Science* scientometric base, does not necessarily provide an estimate of the average level of indicators (in this case, most likely, altmetrics) in sections of online user communities.

Table 1. Number of citations of publications

Publication	Number of citations
Publication *A*	$r_A = 5$
Publication *B*	$r_B = 3$

Table 2. The average values of the number of citations of publications on scientific communities associated with the authors of publication *A*

Scientific community	Average number of citations of publications
z_{city} (city of Moscow)	$r_{\text{city}} = 0.46$
z_{count} (country of Russia)	$r_{\text{count}} = 0.36$

AN EXAMPLE OF COMPARISON OF RESEARCH RESULTS BY SCIENTIFIC COMMUNITIES

For comparison, we selected two *English-language* articles on fuzzy decision-making, published in 2017 by teams involving mainly Russian (*A*) and Vietnamese authors (*B*). As a baseline, a bibliometric indicator of the number of citations in the *Web of Science* scientometric database was used. Assessments of the indicators r_A and r_B, given in Table 1, were taken as of May 2019.

Data on scientific communities were taken from the information about the authors. A publication belongs to the scientific community if at least one of the authors of the publication belongs to this scientific community. Thus, collective articles may relate to several communities at once. However, to simplify the calculations, we will consider only those scientific communities to which the author who is mentioned first belongs. Table 2 shows the average number of citations in Moscow and the Russian Federation of articles of domestic authors in Russian, published in 2017.

Table 3 contains data on the average number of citations of foreign publications published in 2017, in the city of Hanoi and the country of the Socialist Republic of Vietnam.

Table 4 shows the parameters of the fuzzification conversion of baseline deviations for scientific communities, obtained by substitution $d = 5$ in formulas (9) and (11).

In Table 5 we present the results of fuzzification of citation number deviations for the publication *A*, using fuzzy logic production rules such as

$$\text{IF}\,\Delta_z \geq p_1^+ \wedge \Delta_z < p_2^+, \text{THEN}\,\hat{\Delta}_z = \{\text{“Above the average”}\,\mu_1(\Delta_z), \text{“Almost high”}\,\mu_2(\Delta_z)\}.$$

Substituting the data from Table 5 in formula (13) for the publication *A*, we obtain the integrated linguistic variable of the citation number deviation of the

Table 3. The average values of the number of citations of publications on scientific communities associated with the authors of publication *B*

Scientific community	Average number of citations of publications
z_{city} (city of Hanoi)	$r_{city} = 4.36$
z_{count} (country of Vietnam)	$r_{count} = 4$

publication *A* from the average values in the scientific communities z_{city} and z_{count}:

$$\hat{\Delta}_A = \hat{\Delta}_{city} \otimes \hat{\Delta}_{count} = \{\text{“Above the average”}\ CF = 1.035, \text{“Almost high”}\ CF = 80.535\}. \quad (19)$$

According to formula (17) for the publication *A*, the citation index r_A^* normalized by the scientific communities is calculated:

$$\begin{aligned} r_A^* &= \phi\left(\left\|\{\text{“Above the average”}\ CF = 1.035, \text{“Almost high”}\ CF = 80.535\}\right\|\right) \\ &= \phi\left(\{\text{“Above the average”}\ CF = 1.269, \text{“Almost high”}\ CF = 98.731\}\right) = 0.748. \end{aligned} \quad (20)$$

Guided by the same fuzzification rules, in Table 6 we obtain linguistic variables of citation number deviations for publication *B*.

Thus, we have an integrated linguistic variable of the citation number deviation of the publication *B* from the average values in the scientific communities z_{city} and z_{count}:

$$\begin{aligned} \hat{\Delta}_B &= \hat{\Delta}_{city} \otimes \hat{\Delta}_{count} \\ &= \{\text{“Below the average”}\ CF = 91\}. \end{aligned} \quad (21)$$

Accordingly, the citation index normalized by the scientific community for publication *B* is equal to:

$$\begin{aligned} r_B^* &= \phi\left(\left\|\{\text{“Below the average”}\ CF = 91\}\right\|\right) \\ &= \phi\left(\{\text{“Below the average”}\ CF = 100\}\right) = 0.375. \end{aligned} \quad (22)$$

From this it follows that the publication by domestic researchers has a greater number of citations and should be valued more highly than the publication of foreign colleagues, given the appropriate adjustment for the scientific communities to which the authors of the publications belong.

Table 4. The boundaries of the fuzzification conversion intervals for deviations from the average number of citations of publications by scientific communities

Boundary point	Fuzzification parameters in the positive direction	Fuzzification parameters in the negative direction
$i = 1$	$p_1^+ = 1$ (Above average)	$p_1^+ = -1$ (Below average)
$i = 2$	$p_2^+ = 5$ (Almost high)	$p_2^- = -5$ (Almost low)
$i = 3$	$p_3^+ = 25$ (High)	$p_3^- = -25$ (Low)
$i = 4$	$p_4^+ = 125$ (Very high)	$p_4^- = -125$ (Very low)

Table 5. The fuzzification of deviations of the number of citations of publications from the average values in scientific communities associated with the authors of the publication *A*

Scientific community	Deviation of the number of citations of publications from the average value	Linguistic variable of the deviation of the number of citations of publications from the average value
z_{city}	$\Delta_{city} = 4.54$	$\hat{\Delta}_{city} = \{$“Above average” $CF = 11.5$, “Almost high” $CF = 88.5\}$
z_{count}	$\Delta_{count} = 4.64$	$\hat{\Delta}_{count} = \{$“Above average” $CF = 9$, “Almost high” $CF = 91\}$

Table 6. The fuzzification of deviations of the number of citations of publications from the average values in scientific communities associated with the authors of the publication *B*

Scientific community	Deviation of the number of citations of publications from the average value	Linguistic variable of the deviation of the number of citations of publications from the average value
z_{city}	$\Delta_{city} = -1.36$	$\hat{\Delta}_{city} = \{$“Below average” $CF = 91$, “Almost high” $CF = 9\}$
z_{count}	$\Delta_{count} = -1$	$\hat{\Delta}_{count} = \{$“Below average” $CF = 100\}$

CONCLUSIONS

The formalization of approaches to the assessment of scientific activity allows us to explain the choice of indicators that are convenient in assessing the results of research work. The opinion of the professional scientific community is also important. At the same time, it is necessary to understand what each individual scientific community represents, since all of them differ significantly in both number and competence and may not be equivalent to each other, despite the respect that the scholarly profession enjoys virtually universally.

Comparison of the performance indicators of research regarding the level of scientific communities would be useful when evaluating publications, including dissertation papers written in different places or at different periods of time. In addition, the need for such a comparison appears when deciding on the composition of participants in conferences and other scientific events when the presence of representatives of approximately the same level from different local communities is desirable.

Professional scientific communities can be guided by the formulated criteria for evaluating research results, scientific publications, and research organizations. This will make it possible to move away from the total dominance of representatives of the most reputable scientific communities in achieving the highest scientific contributions.

The method of comparative analysis of research results in scientific communities opens up new opportunities for determining the degree of compliance of Russia's scientific priorities with the world's scientific mainstream, as well as for monitoring our country's international scientific cooperation in priority scientific areas, thereby contributing to the formation of an effective system of scientific organizations, an increase in their role in the socio-economic development of the country, and an increase in the prestige of national science.

FUNDING

This work was carried out as part of a study on the topic 0003-2019-0001 of the state task of the VINITI RAS.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Dutta, B., The journey from librametry to altmetry: A look into the past, *Mezhdunar. Forum Inf.,* 2015. vol. 40, no. 2, pp. 3–12.
2. Zinov'eva, I.V., Inozemtseva, S.N., Kosarenkova, M.I., Markusova, V.A., Mindeli, L.E., Solomentseva, O.A., and Solopova, M.A., *Nauka, tekhnologii i innovatsii Rossii: Krat. stat. sb.* (Science, Technology, and Innovation of Russia: Small Statistical Book), Mindeli, L.E., Ed., Moscow: Inst. Probl. Razvit. Nauki Ross. Akad. Nauk, 2017.
3. Varganova, T.N., Zinov'eva, I.V., Inozemtseva, S.N., Markusova, V.A., Solomentseva, O.A., Solopova, M.A., and Kosarenkova, M.I., *Nauchnyi potentsial Rossii za 2006–2014 gody: Analit.-stat. sb.* (The Scientific Potential of Russia for 2006–2014: Analytical-Statistical Book), Mindeli, L.E., Ed., Moscow: Inst. Probl. Razvit. Nauki Ross. Akad. Nauk, 2017.
4. Gorodnikova, N.V., Gokhberg, L.M., Ditkovskii, K.A., Kotsemir, M.N., Kuznetsova, I.A., Lukinova, E.I., Martynova, S.V., Ratai, T.V., Rosovetskaya, L.A., Sagieva, G.S., Strel'tsova, E.A., Suslov, A.B., Tarasenko, I.I., Fridlyanova, S.Yu., and Fursov, K.S., *Indikatory nauki: 2018: Stat. sb.* (Indicators of Science: 2018: Statistical Book), Gokhberg, L.M., Kuz'minov, Ya.I., Laikam, K.E., and Matveev, S.Yu., Eds., Moscow: Nauchno-Issled. Univ. Vyssh. Shk. Ekon., 2018.
5. Science & Engineering Indicators 2018. National Science Board. https://nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf.
6. Bornmann, L. and Marx, W., Critical rationalism and the search for standard (field-normalized) indicators in bibliometrics, *J. Inf.,* 2018, vol. 12, no. 3, pp. 598–604.
7. Mart'yanova, N.A., Construction of professional associations: From professional groups to expert communities, *Izv. Ross. Gos. Pedagog. Univ. im. A. I. Gertsena,* 2013, no. 162, pp. 136–140.
8. Maslova, Yu.V., Social professional communities: Typology, structure, and purpose (on the example of the library community), *Materialy elektronnoi Vserossiiskoi nauchnoi konferentsii "Professional'noe bibliotechnoe soobshchestvo XXI veka: professionalizm, grazhdanstvennost', tolerantnost'"* (Proceedings of the Electronic Library Conference Professional Library Community of the 21st Century: Professionalism, Citizenship, and Tolerance), Kalegin, O.A. and Kormishin, G.M., Eds., Kazan: Kul'tura, 2011, pp. 74–85.
9. Loiko, V.I., Loiko, V.I., Romanov, D.A., Kushnir, N.V., and Kushnir, A.V., Qualimetric assessment of integration of scientific staff into the scientific community based on citation analysis, *Politemat. Setevoi Elektron. Nauchn. Zh. Kuban. Gos. Agrarn. Univ.,* 2016, no. 118, pp. 1168–1186.
10. Titov, V.N., Institutional and ideological aspects of the functioning of science, *Sotsiol. Issled.,* 1999, no. 8, pp. 62–70.
11. Feigel'man, M.V. and Tsirlina, G.A., Bibliometric passion as a result of the lack of scientific expertise, *Upr. Bol'shimi Sist.: Sb. Tr.,* 2013, no. 44, pp. 332–345.
12. Garcia, N.R., Salinas, D.T., Zahedi, Z., and Costas, R., New data, new possibilities: Exploring the insides of Altmetric.com, *Prof. Inf.,* 2014, vol. 23, no. 4, pp. 359–366.
13. Yurevich, M.A. and Tsapenko, I.P., Prospects for the use of altmetrics in the social sciences and humanities, *Inf. O-vo.,* 2015, no. 4, pp. 9–16.
14. Gureev, V.N. and Mazov, N.A., The place of altmetrics in quantitative methods for assessing scientific activity, *Inf. Innovatsii,* 2018, no. 1, pp. 18–21.
15. Ramos-Vielba, I., D'Este, P., Woolley, R., and Amara, N., Introduction to a special section: Balancing scientific and societal impact—a challenging agenda

for academic research, *Sci. Publ. Policy,* 2018, vol. 45, no. 6, pp. 749–751.

16. Vlasov, V.V., The value of scientific publications in specialized journals, *Profil. Med.,* 2010, vol. 13, no. 4, pp. 44–47.
17. How is the Altmetric Attention Score calculated. https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-score-calculated-.
18. Galyavieva, M.S., Altmetrics and libraries: Trends, opportunities, and problems, *Vestn. Kazan. Gos. Univ. Kul't. Iskusstv,* 2013, nos. 4-1, pp. 27–32.
19. Schreiber, M., Malesios, C.C., and Psarakis, S., Exploratory factor analysis for the Hirsch index, 17 h-type variants, and some traditional bibliometric indicators, *J. Inf.,* 2012, vol. 6, no. 3, pp. 347–358.
20. Waldrop, M.M., Science 2.0: Great new tool, or great risk?, *Sci. Am.,* 2008, vol. 298, no. 5, pp. 69–73.
21. Efimova, G.Z., Analysis of the effectiveness of scientometric indicators in the evaluation of scientific activity, *Vestn. Tyumen. Gos. Univ., Sots.-Ekon. Pravovye Issled.,* 2012, no. 8, pp. 101–108.
22. Batygin, G.S. and Gradosel'skaya, G.V., Network communications in the professional community of sociologists: Methodology for content-analytical research of biographies, *Sotsiol. Zh.,* 2001, no. 1, pp. 88–109.
23. Hood, W. and Wilson, C., The literature of bibliometrics, scientometrics, and informetrics, *Scientometrics,* 2001, vol. 52, no. 2, pp. 291–314.
24. Medynskaya, M.K., Theory of fuzzy sets. The concept of a fuzzy set, *Mezhdunarodnaya konferentsiya po myagkim vychisleniyam i izmereniyam* (International Conference on Soft Computing and Measurement), 2015, vol. 1, pp. 46–48.
25. Rotshtein, A.P. and Shtovba, S.D., The influence of defuzzification methods on the tuning speed of a fuzzy model, *Kibern. Sist. Anal.,* 2002, no. 5, pp. 169–176.

Translated by S. Avodkova