

Д. Г. Газизова

ГПНТБ России

Библиометрический анализ данных по отраслевой тематике: индекс качества и производительности

В статье представлен обзор библиометрических исследований на примере геологических дисциплин. Проведена выборка данных для библиометрических исследований. Основными показателями, рассматриваемыми в исследовании, стали число публикаций, индекс Хирша, импакт-фактор журналов и индекс качества и производительности (IQp). Цели исследования – выявление взаимосвязи различных показателей авторов и библиометрических показателей научных изданий, оценка качества научной деятельности и производительности учёных на основе выбранных показателей. Высказано предположение о возможности использования IQp при предоставлении библиометрических справок экспертам в предметной области.

Статья подготовлена в рамках Государственного задания ГПНТБ России на 2017 г.

Ключевые слова: геология, библиометрия, научные исследования, импакт-фактор, IQp, оценка качества научной деятельности учёных.

Darina Gazizova

Russian National Public Library for Science and Technology, Moscow, Russia

Discipline-specific bibliometric analysis: Quality and performance indexes

Bibliometric studies in geological disciplines are reviewed. Data were selected for the bibliometric study. The main indicators analyzed were: number of publications, Hirsch index, journals impact factor, and IQp. The study task is to reveal interdependence of authors' indicators and publications bibliometric indicators, and, based on selected indicators, to assess researchers' quality and performance. The author suggests that IQp may be used to respond to bibliometric inquiries made by discipline experts.

The paper is prepared within the framework of the Government Task 2017 for the Russian National Public Library for Science and Technology.

Keywords: geology, bibliometrics, scientific research, impact factor, IQp, researchers' quality assessment.

The Ministry of Education and Science evaluates the researchers and university professors by virtue of bibliometric systems. On the ground of quantitative analysis of publications and their citations, one can compare the productivity of scientists, determine which areas are now particularly popular, and which, on the contrary, lose popularity. The study used the Web of Science database. As an interval ten full years (2007–2016) was considered. The following eight categories were chosen: Crystallography; Geochemistry & Geophysics; Geology; Geosciences, multidisciplinary; Mineralogy; Mining & Mineral processing; Paleontology. For each category there were selected 10 journals included in the first quartile and ranked by the impact factor of their area. The Hirsch index of authors with fewer publications grows less rapidly than the Hirsch index of authors with a large number of publications. The number of publications and the Hirsch index cannot be the only indicators of the contribution to science of a certain scientist. The index of quality and productivity proposed by Hirsch as an alternative to the Hirsch index takes into account the general influence of the scientist, and also makes a correction for citation rates in certain areas, scientific productivity and academic age (from the time of the author's publication activity to today's days). The index of quality and productivity, which demonstrated on the example of scientists working in the geological area, that not always quantitative indicators can serve as a measure of quality. This index is suggested to be used as a tool for assessing the productivity and usefulness of the scientist for his field, but it is also necessary to take into account the widely used Hirsch index and many other quantitative indicators. Therefore, the presented indicators can be used for a more objective assessment of scientists and organizations when making managerial decisions at the appropriate level. And on the basis of the results of a comprehensive analysis implemented with the application of various approaches and techniques, an expert of a specific subject area has the opportunity to draw up an expert opinion.

В настоящее время библиометрический и наукометрический анализ широко используется для оценки результатов научной деятельности учёных [1]. Однако следует отметить, что библиометрических показателей, с помощью которых можно абсолютно точно оценивать эффективность деятельности научных работников, не существует. Вместе с тем следует признать и важность наукометрии. С помощью количественного анализа публикаций и их цитируемости можно сравнить условную эффективность деятельности учёных, определить, какие направления науки сейчас особенно популярны, а какие, наоборот, утрачивают популярность [2].

Вначале основным библиометрическим показателем было количество печатных работ учёного [3. С. 6]. В последнее время Министерство образо-

вания и науки РФ стало оценивать эффективность деятельности научных работников и преподавателей вузов с помощью библиометрических показателей, основанных на международных системах цитирования *Web of Science*, *Scopus* и др. Такой анализ проводится с целью изучения эффективности результатов научной деятельности учёных.

Следует отметить, что определение эффективности научных исследований – непростая задача, поскольку для конкретной предметной области существуют свои показатели и индикаторы и, как следствие, свои критерии оценки. Поэтому при оценке необходимо учитывать комплекс таких факторов, как специфика научных направлений, региональные особенности проводимых исследований, определённые особенности цитирования в отдельной области знания.

Изучив иностранные источники [4], можно говорить о том, что библиометрическими исследованиями охвачены практически все области знаний. Однако в отечественной практике подобные исследования в отдельных предметных областях значительно уступают зарубежным, а некоторые области и вовсе не освещены.

В этой статье предполагается шире раскрыть вопросы наукометрического анализа в таких областях знаний, как геология, геохимия, геофизика, кристаллография.

Количественные данные о цитировании публикаций отражают влияние результата исследования на научное сообщество, а также насколько то или иное исследование полезно для других учёных [5, 6]. Сами по себе эти данные не оценивают качество публикации, поэтому их стоит рассматривать как индикаторы, показывающие, что данная работа с некоторой долей вероятности может оказаться довольно значимой для исследователей, занимающихся схожими проблемами. Анализ цитирования конкретной публикации может помочь в выявлении популярных работ среди авторов одной области.

Цель моего исследования – проанализировать научные журналы нескольких смежных областей геологических дисциплин, выделить группы авторов по количеству публикаций, проследить их распределение по различным журналам и областям и выявить взаимосвязь различных показателей авторов и библиометрических показателей научных изданий, а также оценить качество научной деятельности и производительность учёных, основываясь на индексе *I_{Qp}* (ИКП).

Методика исследования

В ходе исследования использовалась база данных *Web of Science*. В качестве временного промежутка рассматривался интервал в десять полных лет (2007–2016 гг.). Выбор такого интервала обусловлен тем, что в геологии

данные устаревают не так быстро, и показатели *cited half-life* и *citing half-life* для журналов этой области равны в среднем десяти годам. При этом необходимо, чтобы все учёные были в равных условиях, не было существенного превосходства учёных со стажем над молодыми.

В качестве основы для проведения исследования выбрана классификация *Web of Science*: были отобраны направления, соответствующие принятым в этой базе данных категориям, по которым производилась выборка научных изданий и авторов. Для геологии это восемь следующих категорий: *Crystallography; Engineering, geological; Geochemistry & Geophysics; Geology; Geosciences, multidisciplinary; Mineralogy; Mining & Mineral processing; Paleontology*.

Для каждой категории было отобрано по 10 журналов, входящих в первый квартиль и ранжированных по величине импакт-фактора своей области. Для выявления закономерностей изменения показателей авторов и журналов во времени по каждому журналу отобрано по три автора, имеющих разные наукометрические показатели. Все авторы были поделены на три группы, ранжированные по количеству публикаций за исследуемый период. Первая группа – А – включала в себя авторов с числом публикаций более 150, вторая – В – от 51 до 150, третья – С – от 1 до 50 публикаций. Всего в выборке было 240 авторов.

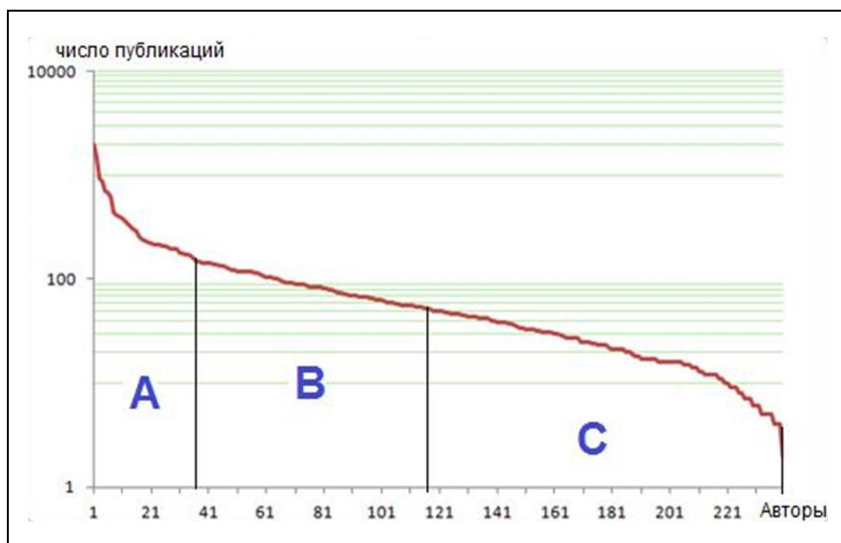


Рис. 1. Выборка авторов и количество их публикаций за 10 лет

График (рис. 1) показывает: большую часть (122) авторов составляют группы С и В (82 автора), а в группе А всего лишь 36 авторов. Это также наглядно отображает ABC-анализ как частный случай принципа Паретто [7]: в данном случае по количеству публикаций 51% авторов вносят 11,86% вклада, 35% – 29,57%, 15% – 58,57%. Эти данные отображены в таблице.

Авторы и их доля в общем объёме публикаций

Группа авторов	Количество авторов в выборке	Количество авторов, %	Количество публикаций	Доля публикаций от общего числа, %
А	36	15	14 592	58,57
В	82	34	7 368	29,57
С	122	51	2 954	11,86
Всего:	240	100	24 914	100

Однако не стоит судить о вкладе учёного в определённой области только по количеству его публикаций. В нашем исследовании проведён анализ и других показателей, рассмотрены факторы, оказывающие не меньшее влияние на оценку вклада учёного в своей предметной области.

Наукометрический анализ авторов

На основе собранных данных были построены графики, которые отразили изменение величины индекса Хирша (*h*-индекса) в зависимости от роста числа публикаций за 10 лет (рис. 2) и количества цитирований (рис. 3) для учёных выделенных групп.

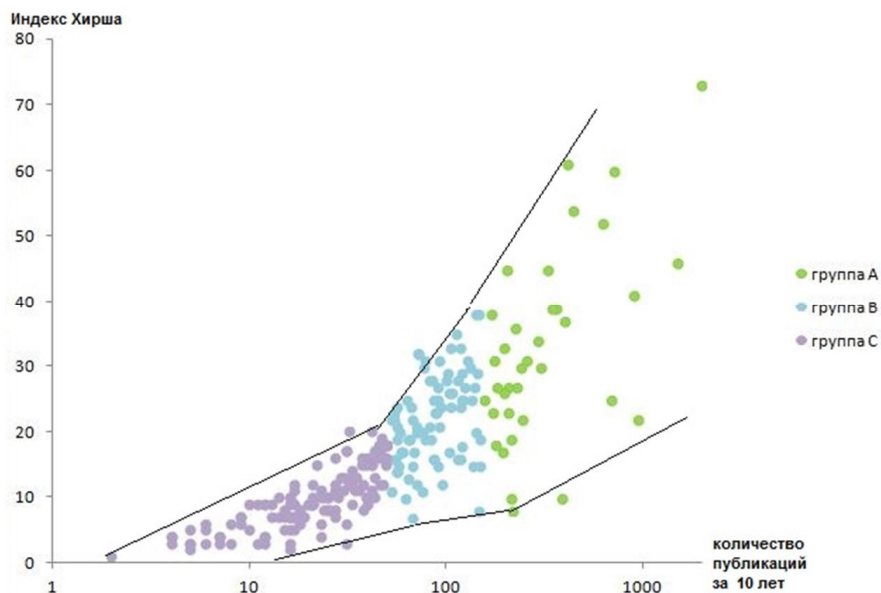


Рис. 2. Изменение индекса Хирша в зависимости от роста числа публикаций за 10 лет

График, представленный на рис. 2, показывает, что индекс Хирша авторов, имеющих меньшее число публикаций и относящихся к группе С, растёт медленнее, чем индекс Хирша авторов с большим числом публикаций. Но при этом учёные группы А характеризуются большим разбросом по значению индекса Хирша, чем учёные группы С.

На другом графике (рис. 3) прослеживается чёткая зависимость индекса Хирша от количества цитирований автора. При этом для группы С не характерно существенное увеличение этого показателя, его резкий рост начинается примерно после значения 1 тыс. цитирований для каждого автора (группы В и С).

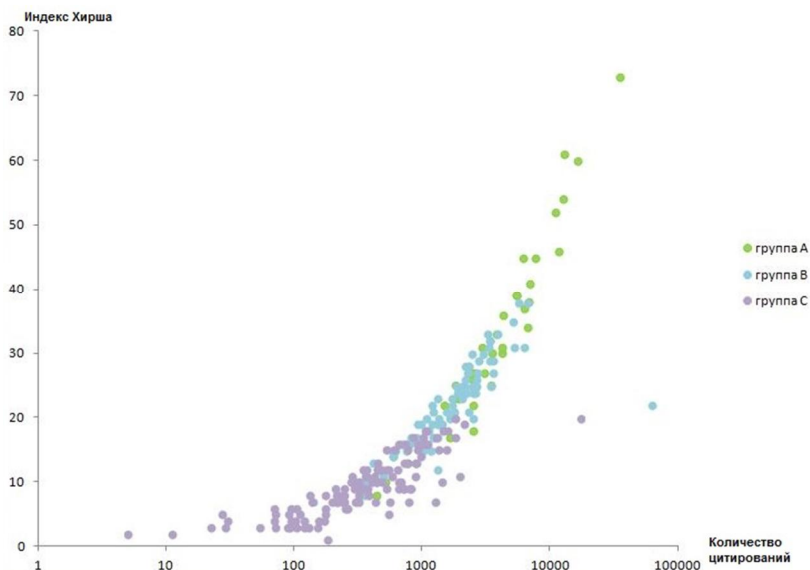


Рис. 3. Зависимость индекса Хирша от количества цитирований

Индекс качества и производительности

Однако, как было сказано выше, по отдельности количество публикаций и индекс Хирша не могут являться единственными показателями вклада учёного в науку. Предложенный в качестве альтернативы индексу Хирша ИКП, или IQp [8], принимает во внимание общее влияние учёного, а также делает поправку на показатели цитирования в определённых областях, научную продуктивность и академический возраст (время с начала публикационной активности автора и до сегодняшнего дня). Для расчёта этого показателя используется следующая формула [9. С. 198]:

$$IQp = \frac{C_T}{P + \frac{C_p}{P}} = \frac{P \times C_T}{P^2 + C_p},$$

где C_T и P – общее число цитирований и публикаций автора соответственно, а C_p – расчётное значение цитирований, которые автор мог бы получить, если бы достиг среднестатистических успехов в своей области.

Этот показатель вычисляется по следующей формуле:

$$C_p = \frac{Y_a \times c \times (P+1)}{2}.$$

Значение Y_a – это так называемое время работы в науке, годы, в которые автор публиковался, а показатель c является поправочным коэффициентом, отражающим уровень цитирования публикации среднего качества в конкретной области исследования. Это выражение применимо при условии, что автор публикует определённое количество работ (P) за равные промежутки времени, в течение которого он занимается наукой, получая среднее в его области исследований число цитирований на его публикации в год.

Показатель c вычислялся на основе совокупного и средневзвешенного импакт-фактора в трёх основных категориях, в которых цитировался автор. Импакт-фактор журнала показывает, сколько раз в среднем цитируется каждая опубликованная в журнале статья в течение двух последующих лет после её выхода. Импакт-факторы журналов были взяты из опубликованных данных с сайта *Journal Citations Reports* [10]. Более подробный расчёт показателя c приводится в [8].

Таким образом, для каждого учёного будет характерен свой коэффициент в зависимости от того, в каких предметных областях и каких журналах размещены его публикации.

В упрощённом виде формула расчётных цитирований автора будет такой:

$$C_p \text{ (расчетные цитирования)} = \frac{C \times \text{годы} \times (\text{количество публикаций} + 1)}{2}.$$

Подставив все значения в начальную формулу и проведя все преобразования, получаем следующую формулу расчёта ИКП:

$$IQ_p = \frac{\text{Цитирование автора}}{\text{публикации} + \frac{\text{годы} \times c \times (\text{количество публикаций} + 1)}{2 \times \text{публикации}}}.$$

В ходе исследования для имеющейся выборки учёных были произведены расчёты ИКП.

Для наглядного отображения полученных данных построены графики и диаграммы распределения ИКП в зависимости от количества публикаций (рис. 4) и h -индекса автора (рис. 5).

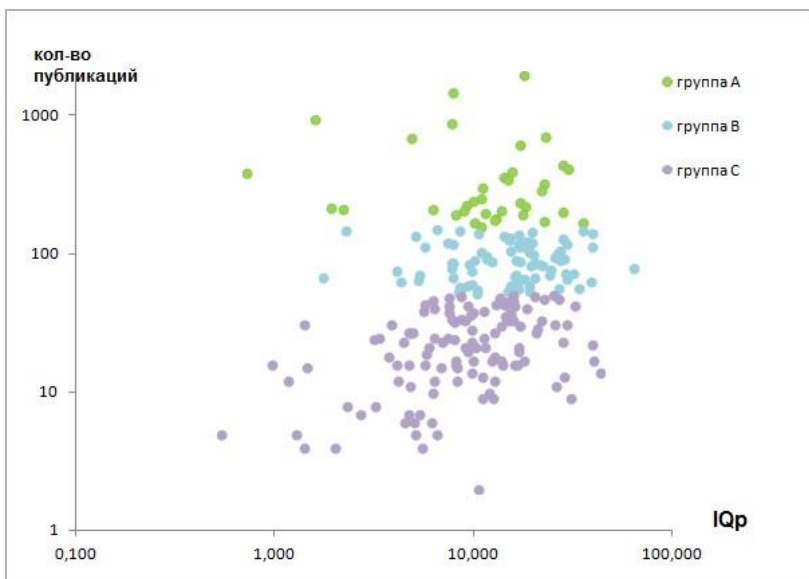


Рис. 4. Индекс качества и производительности (IQ_p) и количество публикаций авторов

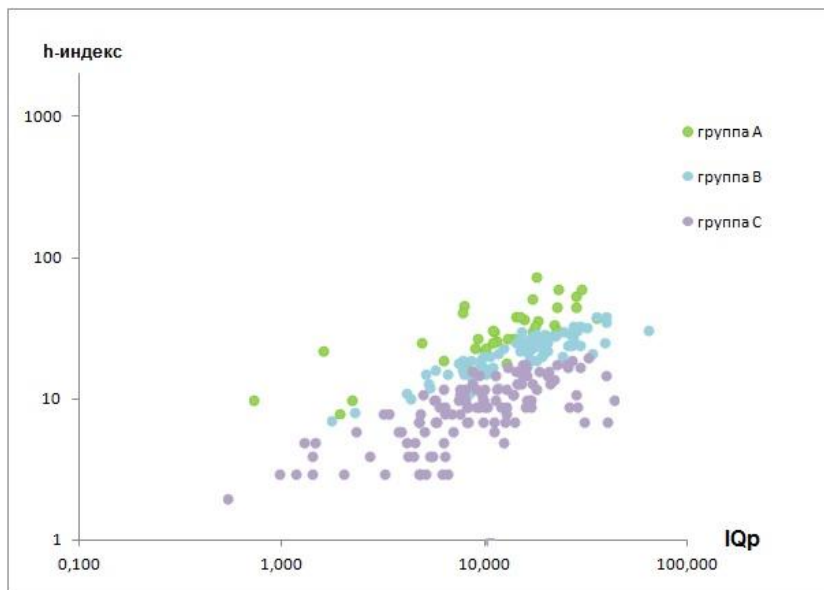


Рис. 5. Индекс качества и производительности (IQ_p) и индекс Хирша

Из представленных графиков видно, что учёные, принадлежащие к одной группе (выделенные на основании количества публикаций), могут иметь совершенно разные ИКП. Например, значение ИКП учёных группы С имеет разброс от 0,5 до 100, группы В – от 1 до 100, группы А – от 0,5 до 100. Внутри групп не прослеживается никакой зависимости параметра ИКП от числа публикаций. Это свидетельствует о том, что независимо от числа публикаций учёный может иметь как низкие, так и высокие значения ИКП.

Если посмотреть на распределение ИКП относительно индекса Хирша, то также можно заметить, что этот показатель имеет различные значения в каждой из групп. В то же время внутри самих групп прослеживаются линейные тренды прямой зависимости ИКП и h -индекса. Нижнюю границу тренда образует группа С, верхнюю – группа А, и наиболее чётко эта зависимость прослеживается в группе В. Однако при этом наблюдается закономерность: учёные с высокими значениями индекса Хирша будут иметь, скорее, средние и высокие значения ИКП, тогда как для учёных с низкими значениями индекса Хирша характерен широкий разброс по значениям ИКП.

Соответственно, возможны случаи, когда учёный с меньшим числом публикаций либо невысоким значением h -индекса может иметь высокое значение ИКП, а учёный с большим числом публикаций и/или хорошим значением h -индекса – невысокое значение ИКП. Например, китайский учёный *Zhang Q. Q.* имеет большое количество публикаций (1 480) и цитирований (11 542) и высокий индекс Хирша (46). Но при этом у него очень небольшой ИКП – 7,766. Другой пример: учёный *Desiraju G. R.*, у которого публикаций в 10 раз меньше (140) и в два раза меньше цитирований, чем у первого автора, при примерно равном индексе Хирша (39), имеет ИКП значительно выше – 39,127.

Таким образом, показатель ИКП можно использовать для оценки научного вклада учёного, но при этом выделенные группы учёных, ранжированные по числу публикаций, перестанут иметь значение.

Выводы

Наукометрические показатели индексов цитирования играют важную роль, они необходимы при оценке научной деятельности авторов. Но следует иметь в виду, что количественные показатели не могут являться единственным критерием качества и значимости учёного. При проведении оценки научной деятельности учёного либо научного издания следует проводить комплексный анализ данных, применимых для *конкретной предметной области*.

В ходе исследования сделан вывод: ни по числу публикаций, ни по индексу Хирша по отдельности невозможно объективно оценить учёного и его вклад в конкретную предметную область. Для этого предложено использо-

вать новый показатель – *IQR*. Его использование применительно к учёным, работающим в геологической области, продемонстрировало, что не всегда количественные показатели могут служить мерой качества. ИКП предложено применять как инструмент оценки продуктивности и пользы учёного для своей области, но при этом стоит учитывать и широко используемые индекс Хирша и многие другие количественные индикаторы.

Следовательно, представленные показатели можно применять для более объективной оценки учёных и организаций при принятии управленческих решений. На основе результатов комплексного анализа, проведённого с привлечением различных подходов и методик, эксперт конкретной предметной области имеет возможность составить экспертное заключение.

СПИСОК ИСТОЧНИКОВ

1. **Шрайберг Я. Л.** Библиотеки, музеи, вузы и книжный рынок в едином информационном цифровом пространстве: общее и особенное : ежегод. докл. конф. «Крым», 2017 г., Судак / Я. Л. Шрайберг. – Москва : ГПНТБ России, 2017. – 79 с.

Shrayberg Ya. L. Biblioteki, muzei, vuzy i knizhnyy rynek v edinom informatsionnom tsifrovom prostranstve: obshchee i osobennoe : ezhegod. dokl. conf. «Crimea», 2017 g., Sudak / Ya. L. Shrayberg. – Moskva : GPNTB Rossii, 2017. – 79 s.

2. **Полянин А. Д.** Недостатки индексов цитируемости и Хирша и использование других наукометрических показателей // Математ. моделирование и числ. методы. – 2014. – № 1 (1). – С. 131–144.

Polyanin A. D. Nedostatki indeksov tsitiruемости i Hirsha i ispolzovanie drugih naukometricheskikh pokazateley // Matemat. modelirovanie i chisl. metody. – 2014. – № 1 (1). – S. 131–144.

3. **Штовба С. Д., Штовба Е. В.** SH-Индекс – новая дробная модификация индекса Хирша // Науч. тр. Винниц. нац. техн. ун-та, 2011. – № 3. – С. 6.

Shtovba S. D., Shtovba E. V. SH-Indeks – novaya drobnaya modifikatsiya indeksa Hirsha // Nauch. tr. Vinnits. nats. tehn. un-ta, 2011. – № 3. – S. 6

4. **Geisler E.** The measurement of scientific activity: Research directions in linking philosophy of science and metrics of science and technology output // Scientometrics. – 2005. – Vol. 62. No. 1. – P. 269–284.

5. **Гарфилд Ю.** Можно ли выявлять и оценивать научные достижения и научную продуктивность? // Вестн. АН СССР. – 1982. – № 6. – С. 42–50.

Garfield Yu. Mozhno li vyyavlyat i otsenivat nauchnye dostizheniya i nauchmuyu produktivnost? // Vestn. AN SSSR. – 1982. – № 6. – S. 42–50.

6. **Прайс Д.** Малая наука, большая наука // Наука о науке : сб. ст. / под ред. В. Н. Столетова. – Москва : Прогресс, 1966.

Prays D. Malaya nauka, bolshaya nauka // Nauka o nauke : sb. st. / pod red. V. N. Stoletova. – Moskva : Progress, 1966

7. **Репин В.** Процессный подход к управлению. Моделирование бизнес-процессов / Владимир Репин, Виталий Елиферов. – Москва : Манн, Иванов и Фербер, 2013. – 544 с.

Repin V. *Protsessnyy podhod k upravleniyu. Modelirovanie biznes-protsessov* / Vladimir Repin, Vitaliy Eliferov. – Moskva : Mann, Ivanov i Ferber, 2013. – 544 s.

8. **Antonakis J., Lalive R.** Quantifying scholarly impact: IQp versus the Hirsch h. // Journal of the American Society for Information Science and Technology. – 59 (6). – 2008. – P. 956–969.

9. **Todeschini R., Baccini A.** Handbook of Bibliometric Indicators. Quantitative Tools for Studying and Evaluating Research. – Milan : 2016, Wiley-VCH Verlag GmbH.

10. **Journal Citation Reports** [Электронный ресурс] // InCites by Clarivate Analytics. – Режим доступа: <https://jcr.incites.thomsonreuters.com/JCRJournal>.

Darina Gazizova, *Researcher, Russian National Public Library for Science and Technology;*

darina.gpntb@gmail.com

17, 3rd Khoroshevskaya st., 123298 Moscow, Russia