EDICIONES
COMPLUTENSE

Mediterranean Botany
ISSN-e: 2603-9109

EDITORIAL

Assessing scientific publications by their impact. A possibility for a more accurate evaluation of researchers

Javier Loidi
University of the Basque Country (UPV/EHU). 48080 Bilbao, Spain

https://doi.org/10.5209/mbot.99402 Received: 3 April 2024 / Accepted: 18 December 2024 / Published online: 24 March 2025

Abstract. The use of metrics in the evaluation of a researcher's output is now a common practice and has a decisive influence on their career (metricracy). It is widespread to consider the Journal Impact Factor (JIF) as a means of assessing the quality and significance of a paper. It has been discussed that it is appropriate to use a metric that evaluates each paper based on its citations rather than the journal in which it was published. Other parameters such as the number of years since publication, the number of authors and their position in the author list are also discussed and a formula for scoring each published article is proposed.
Keywords: citations, Journal Impact Factor, scientific value, metricracy, Google Scholar, Scopus, Web of Science.

How to cite: Loidi, J. 2024. Assessing scientific publications by their impact. A possibility for a more accurate evaluation of researchers. Mediterr. Bot. 46(1), e99402. https://doi.org/10.5209/mbot.99402

Introduction

The "metricracy"

Recently, competition for resources of all kinds, especially for permanent positions, has increased among young scientists and has developed in line with the current globalization in our world. Scientific and technological research is an activity in which more and more professionals are involved worldwide, currently up to 20 million (Prichina *et al*., 2017). A large proportion of them publish in approximately 126,000 journals (Kurmis, 2003) and produce more than 5 million publications per year (KNOEMA; WORDSRATED). It is assumed that the peer review process is widespread in many of these journals.

With such numbers, a quantitative evaluation system is necessary to allocate positions, projects and other resources, especially by public bodies such as governments, national and international funding organizations, universities, institutes, etc. (Simons 2008). For the last three decades, various metrics have been used to assess the output of countries, institutions and individuals (Marchant, 2009; Wilsdon *et al*., 2015). This has led to a kind of "tyranny of academic accountability" (Campbell, 2008) or "metricracy". The whole current evaluation system is supported by the information provided by the Web of Science (WoS), SCOPUS and Google Scholar (https://scholar.google.com/).

The way to evaluate a researcher's output is essentially to assess the number and quality of papers published in indexed journals. Assessing quantity is straightforward as it can simply be calculated based on the number of papers published, but assessing quality is a more difficult question. For this, the "quality" of the journal in which the individual articles were published is used. This quality is estimated using the Journal Impact Factor (hereafter JIF) of the year in which the evaluated paper was published. This method, which has been widely used for many decades, is calculated for a given year by dividing the number of citations of papers published in a journal in the last two years by the number of citable papers published in the same period (Garfield, 2003) and has become one of the most important tools for evaluating journals, research programs and even the scientific production of institutes (Hecht *et al*., 1998). Although this metric is almost universally recognized, it is sometimes debated whether it is appropriate for assessing the scientific value or quality of a journal (Hecht *et al*., 1998; Kavic & Satava, 2021; McKiernan *et al*., 2019; DORA). In any case, it is an evaluation of the journal, but not of the individual articles published by it (Alberts, 2013). The highly cited works compensate for the bad ones (lead in the wings), creating an average that masks individual impact or acceptability. A small percentage of highly cited published articles are responsible for a journal's high JIF (Kurmis, 2003).

When we rely solely on journals for scientific publications, we give a significant amount of power to those who control them, including private companies, universities, and scientific societies, because the use of these metrics has become critical to a researcher's career advancement. Every position you aspire to, every project you submit to a scientific authority for approval, and every step in your scientific and professional career depends on the value you achieve with them. Today, these metrics are simply critical (Alberts, 2013), hiring, faculty promotion, and research funding allocation all depend on them (Simons, 2008). Some of them, such as the h-index (Hirsch, 2005; Bornmann & Daniel, 2007), have an individual character as they are based on the citations of the articles published by an author, while others, such as the IF of the journals in which it is published, are mere proxies of lower reliability.

Some drawbacks of the current accounting system

The comparison between different scientific fields

Fields belonging to disciplines with different citation rates per paper introduce an inevitable bias when we try to make numerical comparisons (Campbell, 2008; Kumar, 2018). These differences can be marked: the geographical and conceptual breadth of many papers, the smaller number of active researchers in the least cited fields, etc. To avoid this error, it has been recommended that these metrics should not be used to compare the relative importance of the output of authors from different disciplines (Kelly & Gennions, 2006). For this reason, I believe it is essential that comparisons between journals be limited to the scope within a discipline (Linardi *et al*., 1996).

The explosion in the number of authors. Certainly, it is now common to undertake projects involving a larger number of researchers, including data providers. This inevitably leads to an increase in the number of authors who need to be categorized in different ways to distinguish the degree of their involvement in each paper. Generally, it is tacitly assumed that the 1st author is the main producer of the work, its designer and the main editor. The last author is usually the senior colleague who provides funding and integrates the research into larger programs; he or she also intervenes in a relevant way in the execution of the work.

Then there is a group of two to five authors, the core authors, who were instrumental in the writing of the article, in the analyses carried out, and in its drafting, at least in some of its sections. The other authors are only involved to a limited extent, often only by providing unpublished data and editing drafts prepared by others. In this context, and to the extent that effective control is not possible, it is not uncommon to see “honorary” authors (Campbell, 2008) who, for personal, family or practical reasons, or even on the basis of inter-group agreements, decide to involve colleagues from other groups in the authorship of their publications in return for the latter doing the same with theirs (Reciprocal Citation Agreements, RCAs).

Opportunistic practices. Other well-known practices used by unscrupulous authors and editors to artificially increase the number of citations in articles, journals and CVs include “self-citations”, “salami publishing” and editorial recommendations to authors asking to include self-citations in submitted articles. They are not uncommon and were perfectly documented a long time ago (Kurmis, 2003).

Predatory journals. Another undesirable effect of the current system combined with the development of the digital era is the recent emergence of journals organized around a different financing model. Instead of subscription fees, they require authors to pay Article Processing Charges (hereafter APCs) and offer fully open access to published articles. This has the advantage that the articles are read and cited by a larger audience, of course only if you have enough money to pay the APCs. This is also done by some reputable journals, but it helps to create a new environment in which for-profit companies are entering the scientific editorial field and imposing new logic based on commercial criteria, leading to difficulties for researchers from developing countries who have less access to financial resources (Dengler, 2023).

We are witnessing increasing corruption favored by the commodification of the world of science, especially in the field of dissemination, which is related to the gaps that remain open in the computerized and globalized system that we have developed in recent decades. I believe that an individualization of estimates can help to control these phenomena to a certain extent.

The special case of biodiversity sciences

In the case of the sciences that deal with biodiversity, such as botany or vegetation science (phytosociology), the problem has been dramatic. These fields have a long tradition of scientific publications, dating back to the XVIII century in the case of classical botany or zoology. As a result, a huge amount of literature has accumulated over the centuries where modern procedures for publishing science have not been applied. In the last decades of the 20th century, the use of the above-mentioned metrics was demanded by funding agencies and governments in all fields of science, including the sciences dealing with biodiversity. At that time, there were a large number of journals publishing a large number of papers for which metrics had not previously been used. Up to this point, researchers were not evaluated using metrics and peer review was not a common practice. This was in line with the old tradition of the “feudal” university that prevailed on the old continent, where favoritism, nepotism and privileges of social rank dominated much of the academic world before the Second World War (Hecht *et al*., 1998). Nonetheless, the acceptance of papers for publication was relatively rigorous and effective enough to ensure good quality of published material, and many highly relevant papers were published in this pre-metrics period. The European environment was multilingual, and scientists had to invest a lot of time and effort in learning languages. The transition to the post-metric phase was difficult, slow and often traumatic (Saiz-Salinas, 1996), but it eventually took place. The current scenario is virtually monolingual (English only) and all relevant journals now apply procedures to achieve metrics.

However, this change has also had some negative effects. The provision of baseline data is an essential task in these biodiversity sciences, as it forms the fundamental matrix for the creation of distribution models, phenology, community classification, taxonomy and much more. The problem is that reports providing basic data on flora or vegetation are rarely accepted in international journals with a high JIF. In addition, there are local or regional surveys of taxonomy or syntaxonomy describing new taxa or syntaxa that receive few citations because their potential interest is not very broad. However, their results are incorporated into reviews of general scope that collect large amounts of local information, and it is expected that these general, comprehensive reviews will be highly cited (Hecht et al., 1998; Simons, 2008; Kumar, 2018). This dynamic between reports and reviews illustrates the conflict in which the work of the many benefits a few. This phenomenon leads to a general discouragement of researchers, preventing them from producing this type of fundamental report. Moreover, it has led to the decline and even the disappearance of many journals in non-English-speaking countries, as researchers from these countries have been forced to publish in journals with a high impact factor. Had they survived, they would have played an important role in scientific dissemination and as support for local or regional research (Kurmis, 2003).

Another issue is the validity of a publication over time. Publications in the field of biodiversity sciences often have a long period of time in which they are citable (Kurmis, 2003). The JIF does not do justice to this characteristic at all, as it favors research areas that produce predominantly short-term studies, since it is only applied to the two years following publication (Hecht *et al*., 1998). Luckily, in the last years it is also used a five-years JIF, better adapted to the long-term validity ecological studies.

The collection and publication of basic data has thus been discouraged and has declined drastically in recent decades. The situation now is that very little new data is being made available. Fortunately, there is a large amount of data from old publications, mostly in digital repositories, which were collected in the pre-metrics era and are used for many synthesis studies. In the past, a scientist could build his/her career by publishing papers based on such raw data, but today this is still difficult as he/she will have to find a high-ranking JIF journal to publish them. In the future, unless new raw data comes along, the results of these synthesis surveys will gradually become obsolete.

A formula to evaluate the value of each scientific publication

Calculating the value of a publication using a proxy such as the JIF of the journal in which it was published is an outdated and old-fashioned approach, given the availability of new information technologies today. The JIF cannot be considered an effective method for assessing the value of an individual article or author (Kurmis, 2003). This could be corrected by introducing a method that assesses individual articles rather than the journal in which they were published. The basic idea is that the scientific value of a paper should be judged by its content and impact and not by its packaging (Hecht *et al*., 1998). All other integrative assessments can be carried out using this type of individual assessment: Institutions, programs, projects, etc. For this reason, the use of the following parameters is suggested:

Number of citations. The best estimate of a paper's influence is the number of citations it receives, either from WoS, Scopus or Google Scholar. They can include all citations or exclude the self-citations. Of course, citations increase over the years and so does the value of the paper, but we can balance this effect by also considering the number of years since publication.

Authorship. Another problem with the method currently used is that it rarely takes into account the number of authors. Recently, a large number of publications with multiple authors have appeared, due to the contribution of many data providers, but also partly due to invited authors. It may be appropriate to apportion the value of the paper among them, taking into account their position in the author list. It is to be expected that papers with many authors will be cited more frequently as many of them tend to cite them. The position of the author in question in the author list should also be taken into account.

Duration of validity. As far as the period of validity of a paper is concerned, it can be assumed that it decreases over the years. As a rule, papers are frequently cited in the first years after their publication, after which their citation rate decreases. However, as already mentioned, in some scientific fields they continue to be cited regularly for a long time. In the case of biodiversity sciences, the only two years that predominantly JIF takes into account are completely insufficient. If we only took into account the number of citations, the value of the work would constantly increase over time. To achieve a balance, we suggest taking into account the number of years since publication.

The proposed formula is as follows:

$$V = \frac{c+1}{\sqrt{y+\ a}} + C$$

Where:
V = value of the published element (paper).
c = total number of citations. The more citations, the higher the value; it is a measure of the paper's impact.
y = number of years since publication. The value of the paper decreases with its age.
a = number of authors. A high number of them reduces the value of the paper.
C is a constant that depends on the position of the researcher in the list of authors: = 1 if he is the first author; = 0.75 if he is the last author and = 0.5 if he is between the 2nd and 5th author. For the rest = 0.

The impact factor of the journal in which the article (paper) was published has no influence on this formula. Only the performance and conditions of the paper are taken into account: Its impact (evaluated based on citations), its age and its authorship.

To test the validity of this approach, a selection of six researchers of the same scientific field was made on the Google Scholar website and their value for 2019 was selected. The results are shown in Table1 and Figure 1

Table 1. Results on the validation of the new approach. Abbreviations are: Papers, n. papers*2; Citations, n. citations/10; Authors, a/10; Index (V), new formula (see text above); c/p and a/p are the quotient of c or a (see V) with Papers/2.

Researchers	Papers	Citations	Authors	Value Index (V)	c/p	a/p
Res. 1	14	17,4	19,4	**34,13**	24,86	27,71
Res. 2	46	49,9	102,4	**83,47**	21,70	44,52
Res. 3	30	29,8	36,8	**84,47**	19,87	24,53
Res. 4	24	61,8	45,9	**93,69**	51,50	38,25
Res. 5	24	63,6	56	**99,14**	53,00	46,67
Res. 6	32	91	35,1	**187,16**	56,88	21,94

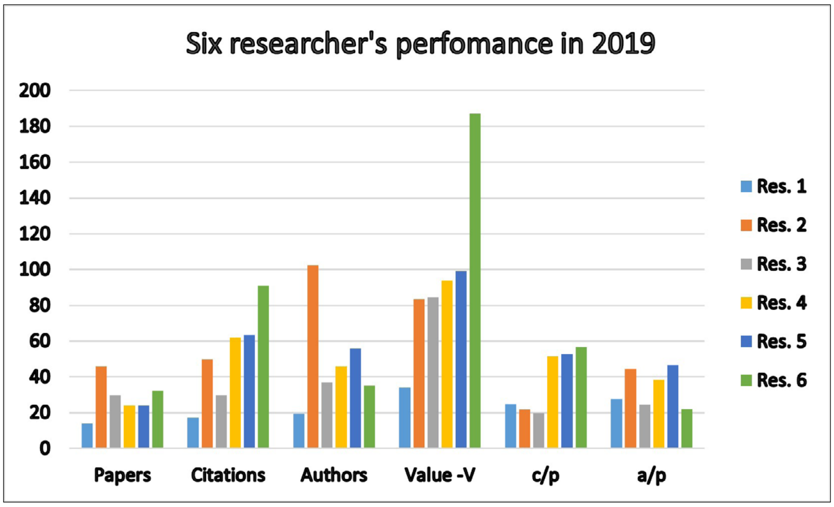

Figure 1. Bar-graph of the six hypothetic researchers evaluated with the new approach. See Table 1 for explanations.

Researcher 1 is a novel researcher with a lower number of published papers this year. His V value is therefore the lowest, but this is somewhat compensated for by the regular values for citations per paper (c/p) and authors per paper (a/p). Researcher 2 has a high number of published articles, that accumulate a large number of citations, but a high ratio of a/p, which lowers the V-value could be expected. Researcher 3 has a low number of citations but achieves a high V-value as it has a low a/p ratio. Researchers 4 and 5 show a regular pattern in terms of number of citations and authors per paper, and Researcher 6 scores the highest value for V, which is enhanced by a low a/p ratio.

In conclusion, the more publications produced by a researcher and more citations obtained, the higher the V values. This formula modulates the quantitative assessment of each researcher's results, as a high number of authors per paper significantly reduces their final value, as well as the number of years since it was published.

References

Alberts, B. 2013. Impact factor distortions. Science 340: 787. doi: 10.1126/science.1240319

Bornmann, L. & Daniel, H.-D. 2007. What do we know about the h-index? J. Am. Soc. Inf. Sci. Tec. 58(9): 1381–1385. doi:10.1002/asi.20609

Dengler, J. 2023. Priorities in journal selection for authors, reviewers, editors, librarians and science funders. Vegetation Classification and Survey 4: 219–229. doi: 10.3897/VCS.110296

Campbell, P. 2008. Escape from the Impact Factor. Ethics in Science and Environmental Politics 8: 5–7. doi: 10.3354/esep00078

Garfield, E. 2003. The meaning of the Impact Factor. Int. J. Clin. Hlth. Psyc. 3(2): 363–369.

Hecht, F., Hecht, B.K. & Sandberg, A.A. 1998. The journal Impact Factor: a misnamed, misleading, misused measure. Cancer Genet. Cytogenet. 104: 77–81.

Hirsch, J.E. 2005. An index to quantify an individual's scientific research output. PNAS 102 (46): 16569–16572. doi: 10.1073/pnas.0507655102

Kavic, M.S. & Satava, R.M. 2021. Scientific Literature and Evaluation Metrics: Impact Factor, Usage Metrics, and Altmetrics. Journal of the Society of Laparoscopic and Robotic Surgeons 25(3): e2021.00010. doi: 10.4293/JSLS.2021.00010

Kelly, C.D. & Gennions, M.D. 2006. The h index and career assessment by numbers. Trends Ecol. Evol. 21(4): 167–170. doi:10.1016/j.tree.2006.01.005

Kumar, A. 2018. Is "Impact" the "Factor" that matters ...? (Part 1). Journal of Indian Society of Periodontology 22: 95–96. doi: 10.4103/jisp.jisp_195_18

Kurmis, A.P. 2003. Understanding limitations of the Journal Impact factor. J. Bone Joint Surg. 85-A(12): 2449–2454.

Linardi, P.M., Coelho, P.M. & Costa, H.M. 1996. The "impact factor" as a criterion for the quality of scientific production is a relative, not absolute, measure. Braz. J. Med. Biol. Res. 29: 555–61.

Marchant, T. 2009. An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. Scientometrics 80 (2): 327–344. doi:10.1007/s11192-008-2075-y

McKiernan, E., Muñoz Nieves, C., Matthias, L., Niles, M.T., Alperin, J.P. & Schimansky, L. 2019. Use of

the Journal Impact Factor in academic review, promotion, and tenure evaluations. eLife 8: e47338. doi: 10.7554/eLife.47338
Prichina, O., Orekhov, V. & Schennikova, E. 2017. World number of scientists in dynamic simulation for the past and future. 25th International Scientific Conference on Economic and Social Development. XVII International Social Congress. Pp. 69–81. Moscow.
Saiz-Salinas, J.I. 1997. Failed professor. Nature 381: 186.
Simons, K. 2008. The misused Impact Factor. Nature 322: 165. doi: 10.1126/science.1165316
Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwell, M., Tinkler, J., Viney, I. & Vouters, P. 2015. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. Sage Publications, London. doi: 10.13140/RG.2.1.4929.1363

Websites

DORA (Declaration on Research Assessment) https://sfdora.org/
GOOGLE SCHOLAR. https://scholar.google.com/
KNOEMA. https://knoema.com/atlas/ranks/Number-of-scientific-journal-articles
SCOPUS. https://www.scopus.com/home.uri
Web of Science. https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/
WORDSRATED. https://wordsrated.com/number-of-academic-papers-published-per-year/

Supplementary material

Appendix S1. Evaluation data and graphs.