

CiteScore-Based Quartiles for Scientometric Analysis

Mohammad R. Khosravi¹, and Varun G. Menon²

¹ Department of Electrical and Electronic Engineering, Shiraz University of Technology, Shiraz, Iran
E-mail(s): m.khosravi@sutech.ac.ir

² Department of Computer Science and Engineering, SCMS School of Engineering and Technology, Kochi, India
Email(s): varunmenon@scmsgroup.org

Corresponding author: V. G. Menon, PhD

Abstract: This study intends to review and discuss some topics around CiteScore as a new citation impact indicator of Scopus. Qualitative analysis of journals based on scientometric indicators has been usual and is a good way of rapid evaluation of them. Here, we use them to reach a conclusion. We believe that reliability of CiteScore-based quartile is more and this index is preferred because in comparison to IF-based quartile, there are some reasons of preference of CS. In addition, CiteScore can provide a more natural perception on citations comparing with SJR-based quartile. Outcome of this study can be considered as a proof that both types of SJR- and CiteScore-based quartiles should not be assumed as baseline of quality evaluations at the same time.

Key-words: CiteScore (CS), Scientometrics, Scopus, Web of Science, Statistical Quartile

1. Introduction

CiteScore is a new scientometric indicator (citation impact metric) introduced by Scopus, a product of Elsevier, in 2016 for tracking journals performance in terms of citation analysis [1]. Before introducing CiteScore (CS), Scopus has had several science quality assessment metrics like Source Normalized Impact per Paper (SNIP) and Scimago Journal Rank (SJR) indicators. However, these indicators couldn't be strongly pervasive as rivals of Impact Factor (IF) by Journal Citation Report (JCR) related to Web of Science (WoS). Maybe the reason is to use much different definitions in all the prior Scopus indicators comparing to IF, which cannot illustrate a natural perception or normal concept of yearly citation per document, to a researcher who is not expert in scientometrics. Fortunately, CS is able to give us a strong perception on the concept of citation per document because its definition is relatively

near to IF, but with some further advantages. As Fig. 1, we see a transparent definition of IF, whereas CS details are observable in Fig. 2. These figures show that there are detailed differences in definitions of CS and IF. Now, we aim to have a brief review on advantages of CS in comparison to IF. IF has two forms of 2-year and 5-year. The main form is, of course, based on 2-year evaluation frame. On the other hand, CS is calculated for a term of 3-year. It seems that 2-year period is a little short (particularly for fields that are slow in availing newer scientific findings, e.g., many areas of engineering studies) and 5-year is relatively long. However, 3-year term is much better and can provide sufficient time for most of research disciplines in order to count an effective number of citations, so 3-year value would be more accurate, adequate and fair.

The second issue is the broader and more comprehensive indexing, abstracting and citation tracking coverage of Scopus than WoS. WoS generally has less coverage on peer-reviewed scientific materials as abstracted (while it plays role of an indexing directory), and even all of its abstracted journal articles are not used for citation analysis, for example, specialized articles which are exclusively indexed and abstracted in Zoological Records database (it is not a citation database; more up-to-dated details of WoS should be reviewed at Clarivate Analytics website). Therefore, this comprehensiveness of Scopus can give a more realistic citation analysis and consequently, Scopus citation impact indicators would be more reliable currently. A third issue can be discussed around new service of Scopus entitled "CiteScore tracker" as the monthly version of CS with interesting abilities for time series prediction problems in scientometrics. Scopus is updating CS monthly and reports CS tracker whereas there is no similar facility in WoS towards IF. And as a final remark, Scopus considers some offsets of IF in calculating CS, it is towards a fairness and makes CS more accurate. In fact, denominator of IF is restricted to just citable documents but citations to a journal may be related to all published items including its citable/non-citable documents. Because CS does not consider difference between citable and non-citable documents, this indicator has no offset (or bias dedicated to a specified journal/publisher). Besides, a research [2] shows that the IF definition provides an additional bias for some types of scientific journals and also some specific fields whether in terms of having offset or short time frame (for example, Nature group's journals which are published in the field of medicine are availing this bias to reach higher IFs). According to these facts and an intellectual reasoning, we can prefer to use CS. Another issue may also be discussed is this disappointing fact that despite Scopus, WoS does not compute IF and then its Q index for many of WoS-indexed journals, out of WoS core collection or even for Emerging Sources Citation Index (ESCI) database from the core. This can be an important reason for more reliability of Scopus compared to WoS.

As the paper organization, in the second part, our methodology is presented and in the last part, discussions and conclusion will be given.

2. Methodology

Nowadays, qualitative analysis of journals based on scientometric indicators has been usual and is a good way of rapid evaluation of them. WoS has provided a classification based on statistical quartile concept using computations of median on journals with IF descending order, for each subject category. Scopus has similarly classified all of the covered journals for assigning a quartile rank (Q index) through

ScimagoJR database on which all journals are classified into four classes Q1, Q2, Q3 and Q4 (Q1 consists of top journals of the subject category), based on values of their SJR indicator.

Also, other periodical/serial publications including conference proceedings and book series may be considered for Q index, but some details exist which are addressed to [1], [3]-[5] on which all types of serial titles (journals, magazines, book series, and conference proceedings having ISSN) can be allocated to a CS-based quartile, but continuously published conference proceedings are not currently evaluated for SJR-based quartile computations (even though SJR is computed for them) and just scientific and trade journals along with book series can be classified by SJR-based quartiles. Using the qualitative ranking of Q index has two benefits; at first, perfect perception and interpretation of quantitative indicators (e.g., SJR) is difficult, therefore, the preference is to convert quantitative information (IF, SJR and CS) to qualitative results like all types of quartiles (IF-based quartiles in WoS, and SJR-based/CS-based quartiles in Scopus) in order to compare journals with each other, for example, is a journal with SJR=0.45 high-quality? Surely response to this question is not easy. The second benefit is to attend to effect of different scientific areas, for example, an engineering journal with IF=3 may be Q1, but a medical journal with IF=4 may be Q2; these observations are due to various subject categories in Q index classification (amongst quantitative indicators, SNIP is also trying to create such benefit, but when we are hearing SNIP value of a journal, we don't interpret it like IF or CS towards the concept of citation per document).

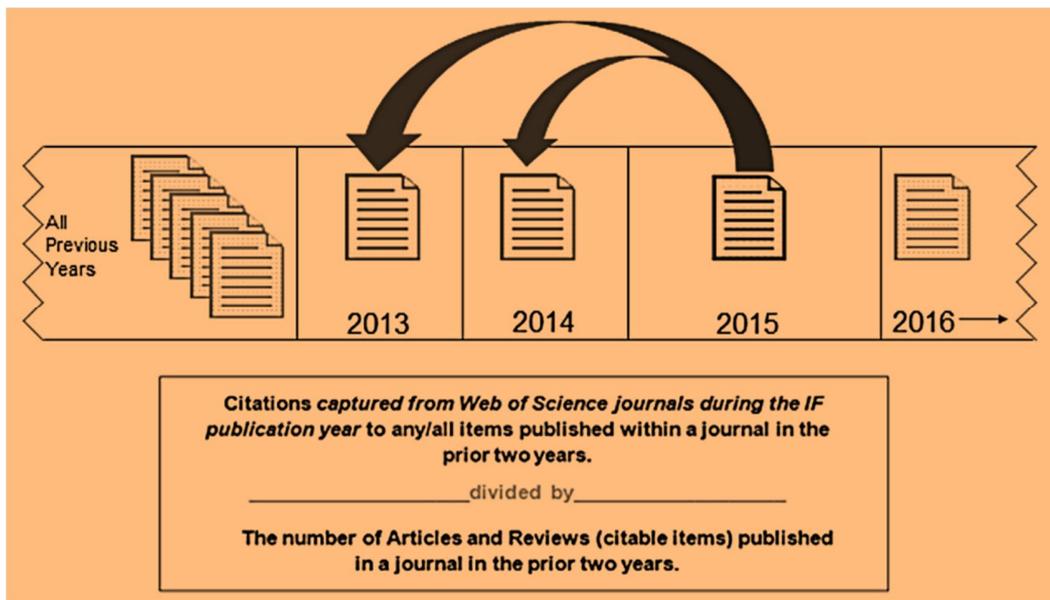
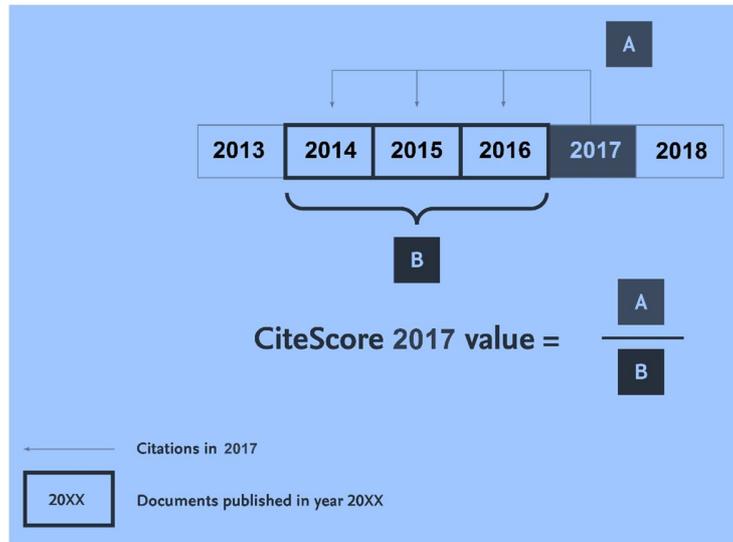


Fig. 1. JCR's IF definition for sample year of 2015.
(Source: Minnick et al. [7])



(a)

CiteScore 2017 methodology

CiteScore 2017 counts the citations received in 2017 to documents published in 2014, 2015 or 2016, and divides this by the number of documents published in 2014, 2015 and 2016.

Documents from 3 years

Citations in 2017

CiteScoreTracker 2018 uses the same methodology with citations based on the latest 2018 data.

3-year publication window

Document types

All types of documents (research articles, review articles, conference proceedings, editorials, errata, letters, notes, and short surveys) are included in the CiteScore calculation. Although articles in press are included in Scopus they are not included in the calculation.

(b)

Fig. 2. CS definition and computing details for sample year of 2017, computation has been done in 2018; (a) Definition: Its general form is similar to IF, but without bias and with a longer time frame, (b) Details: Despite SJR, CS considers all documents, not just citable items. (Source: Scopus [4], and [1])

In addition to quartile-based classification, there are other kinds of qualitative rankings, Scopus provides percentile-based classification for all indexed journals based on CS values. This classification can give us the second benefit, but might not be well-interpreted for the first one. This is because percentiles are not as sound as quartiles on the viewpoint of qualitative inference. For instance,

suppose there are 120 journals in a subject category which should be compared with each other. A journal is directly ranked as 106th from 120 (descending order; lower direct ranks have higher values of numerical indicators and indicate better quality). Its percentile would therefore be 11th.

It is explicit that descriptions of 106/120 (fully quantitative view) and Q4 (qualitative view) are more descriptive than the 11th percentile. So, we believe that quartiles are better than percentiles for quick description of journals' quality. Scopus has determined a formula for computing CS-based quartile [1], [3]-[4], [6], but its results are currently not given directly. If we wish to compute CS-based quartile for each journal in order to have all advantages of CS in comparison to other numerical indicators and also, all benefits of quartile classification (especially comparing with SJR-based quartile because we know both are using Scopus scientific coverage and are similar in this view), a formula can be derived as follows. The formula is another description of a formula based on statistical computations provided by Scopus and makes a connection among direct rank, percentile rank and quartile rank with applying ceiling function. Assume that a journal based on CS information is ranked m^{th} among all n journals existing in that category (normally $n \gg 1$). Thus, CS-based percentile and quartile are specified as below. 0th percentile class (or rank) includes the lowest quality journals, with a CS less than the first statistical percentile value. On the other hand, 99th percentile and Q1 classes include the best journals (definitely accuracy of percentile is higher). Scopus already provides the fraction of m/n and the percentiles directly based on CS. In ScimagoJR as a part of Scopus, the m/n and quartiles are provided directly based on SJR.

$$\text{Percentile Rank } (P^{\text{th}} \text{ class}) = 100 - \left\lceil \frac{m}{n} \times 100 \right\rceil \in \{0, 1, \dots, 99\}$$

$$\text{Quartile Rank } (Q^{\text{th}} \text{ class}) = \begin{cases} 4 & 0 \leq p \leq 24 \\ 3 & 25 \leq p \leq 49 \\ 2 & 50 \leq p \leq 74 \\ 1 & 75 \leq p \leq 99 \end{cases}$$

For $n < 100$, computation of percentile is not suitable and can be without importance, so in these cases, just a quartile computing is sufficient. However, the equation is able to compute P , and based on this computed P , we can calculate Q . Normally, while having $n < 100$, some high-rank classes of P like 99th, 98th, etc., include no journal. Also, in general form for any n , it is possible that the worst direct rank of journals in a category becomes $n-1$, $n-2$, $n-3$, etc. (less than n), this happens because some journals have a similar value of CS or SJR, and therefore, they achieve the same direct rank; in these cases, some lower percentile ranks (0th, 1st, etc.) may not contain any journal. As a final point, we should say that this definition may be a little different from the Scopus definition for percentiles (because of there is no unique definition in statistics), however our definition is simply computed. Just in less number of journals in a specific subject category (much less than 100), we might accidentally observe a difference of 1 or 2 for percentile value between our definition and Scopus' formal definition on which in rare cases it may affect Q ranking, however in practice number of journals are not often less than 100, so we can ignore this difference.

3. Analysis and Conclusions

As a consequence, we believe that reliability of CS-based quartile is more and this index preferred (read a similar work about journal reliability in [8], because:

- In comparison to IF-based quartile, we discussed some reasons of preference of CS comparing with IF.
- In comparison to SJR-based quartile, CS can provide a more natural perception on citations.

Moreover, investigation in case studies shows that there are differences in results of CS-based and SJR-based quartiles for many journals. In Table 1, details of CS and SJR are observed for *Library Hi Tech News* (LHTN). The table shows that in this journal for a specific year and a selected subject category, quartiles extracted based on SJR and CS may not approve results of each other, although both are communicating citation statistics of Scopus. In addition to difference in supporting conference proceedings, some of other concurrent resources of SJR have not been covered by CS (see denominators in direct CS and SJR ranks). This can be considered as a proof that both types of quartiles should not be assumed as baseline of quality evaluations at the same time.

Table 1. LHTN results, assessed using Scopus information in 2017 publication year (CS=0.33, SJR=0.23). Best quartile (category-independent) in terms of both of CS and SJR is Q3.

Subject Category	CS Rank	SJR Rank	CS-based Percentile	SJR-based Percentile	CS-based Quartile	SJR-based Quartile
Library and Information Sciences	126/202	118/213	37th	44th	Q3	Q3
Information Systems	221/251	182/259	11th	29th	Q4	Q3

References

- [1] Zijlstra, H., McCullough, R. (2016), "CiteScore: a new metric to help you track journal performance and make decisions", available at: <https://www.elsevier.com/editors-update/story/journal-metrics/citescore-a-new-metric-to-help-you-choose-the-right-journal>
- [2] Eigenfactor (2019), "Comparing Impact Factor and Scopus CiteScore", available at: <http://eigenfactor.org/projects/posts/citescore.php>
- [3] Zijlstra, H., McCullough, R. (2016), "CiteScore: a new metric to help you choose the right journal", available at: <https://www.elsevier.com/authors-update/story/impact-metrics/citescore-a-new-metric-to-help-you-choose-the-right-journal>
- [4] Scopus (2016), "Journal Metrics Frequently Asked Questions (FAQs)", available at: <https://journalmetrics.scopus.com/index.php/Faqs>
- [5] Borrego, A. (2018), "Are mega-journals a publication outlet for lower quality research? A bibliometric analysis of Spanish authors in PLOS ONE", Online Information Review, doi: 10.1108/OIR-04-2018-0136
- [6] Waltman, L. (2016), "Q&A on Elsevier's CiteScore metric", available at: <https://www.cwts.nl/blog?article=n-q2y254>
- [7] Minnick, J. (2016), "CiteScore: A Non-Rival for the Journal Impact Factor", available at: <https://clarivate.com/blog/science-research-connect/research-management/citescore-a-non-rival-for-the-journal-impact-factor/>
- [8] Khosravi, M.R. (2018), "Reliability of scholarly journal acceptance rates", *Library Hi Tech News*, Vol. 35, No. 10, pp. 7-8, doi: 10.1108/LHTN-07-2018-0044