Scientometrics (2023) 128:1119–1132
https://doi.org/10.1007/s11192-022-04594-1

A modification of citation-based journal indexes

Hui Fang[1]

Received: 27 September 2021 / Accepted: 16 November 2022 / Published online: 28 November 2022

Abstract
The journal indexes currently used to rank journals adopt a short citation window, which has long been criticized. In particular, journals with a prolonged impact are underestimated by this method. By combining the features of synchronic and diachronic journal impact factors, Fang (Scientometrics 125:2265–2282, 2020) proposed the integral synchronic journal impact factor (IIF). The citation window of the IIF can be enlarged without reducing the index value. Therefore, the IIF can effectively and fairly reflect the impact of journals with different citation patterns on the scientific community. To overcome the defect that the IIF might be unstable if a journal publishes few papers in individual years, this work modifies the IIF by borrowing the idea of the new CiteScore: using the citation counts of the adjacent previous years to assist in evaluating the impact level of the journal in the focal year. The modified index can be regarded as a combination of the JIFs and CiteScore. The experimental results show that the modified journal index is more stable over time than the IIF while retaining the associated advantages.

Keywords Journal index · Citation · Publication-citation interval · Citation window · Stability

Introduction

The current popular journal indicators are based on the citation count of papers. Citation-based journal indicators reflect the impact of a journal on the academic community or a journal's reputation.

Garfield (1972) proposed the journal impact factor (JIF), a citation-based index. Initially, the JIF adopts a 2-year citation window, known as JIF2. Journals indexed in the Journal Citation Report (JCR) provided by the Web of Science (WoS, Clarivate Analytics) are ranked in each subject category according to JIF2 each year. As a ranking variable for journals, JIF2 is widely used because of its simplicity (Cai et al., 2019). JCR also issues a JIF with a 5-year citation window, JIF5. JIF2 and JIF5 can be regarded as the extent to which the papers published by a journal in a 2- or 5-year publication window impact the scientific

✉ Hui Fang
fanghui@nju.edu.cn

[1] State Key Laboratory of Analytical Chemistry for Life Science, School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China

community on average in the year following the publication window. This kind of JIF is denoted as synchronic JIF.

Diachronic JIFs have also been proposed and investigated (Frandsen & Rousseau, 2005; Ingwersen, 2012; Ingwersen et al., 2001). A diachronic JIF averages citations among papers published by the journal in one year. Diachronic JIFs can be regarded as measuring the extent to which the papers published by a journal in one year impact the scientific community on average in a series of successive years. A series of diachronic JIFs is needed to reflect the impact of a journal over different publication years on the scientific community.

Scopus launched CiteScore as a journal indicator (Colledge et al., 2017; James et al., 2019; Teixeira da Silva, 2020). The scheme of the first version of CiteScore that averages citations to related papers is similar to that of JIFs. Unlike JIFs, the old CiteScore adopts a 3-year citation window. In 2020, Scopus issued a new version of CiteScore. The new CiteScore of a journal for a certain year (Y) equals the average number of citations obtained during the period from $Y-3$ to Y by the eligible documents published by the journal in that period. The new CiteScore can be regarded as the extent to which the papers published by a journal in a 4-year period influence the scientific community on average in the same period.

JIF2 has long been criticized (for example, Glänzel & Schoepflin, 1995; van Leeuwen et al., 1999; Glänzel et al., 2003; Vanclay, 2009; DORA, 2012). In addition to application abuse (Amin & Mabe, 2000; Zupanc, 2014), data error (Vanclay, 2012), opaque data (Rossner et al., 2007), limited coverage of the document database (Seglen, 1997), etc., it has calculation-related shortcomings. First, the current actual JIF2 and JIF5 suffer from the problem of inconsistent citation. Specifically, the numerator of them includes citations to documents such as "News Items" or "Letters" that are not represented in the denominator, which counts only citable items (i.e., "Articles" and "Review Articles") (Seglen, 1997). Second, JIFs do not consider the skewed nature of citation distributions (Seglen, 1992), which can be rectified by geometric averaging (Thelwall & Fairclough, 2015). Third, the variation in JIF2 tends to be larger for smaller journals, and JIF2 has more variability than JIF5 (Amin & Mabe, 2000) because more samples of papers are included for the average calculation in JIF5. Fourth, JIF2, which is currently used in practice to rank journals, adopts a short citation window. Liu and Fang (2020) and Fang (2020) provide some examples of journals that have a prolonged impact that is underestimated by indexes with short citation windows. One potential reason of JIF adopting a short citation window is that JIF counts the citations received during only 1 year. Because the curve of yearly citations for all the papers published by a journal in a certain year generally decreases after its peak, two or three years after publication for many journals (Amin & Mabe, 2000; Rousseau, 2002), simply extending the citation window might reduce the JIF substantially for many journals, especially when using a long citation window (Fang, 2021). The above-mentioned flaws allow the JIF to be manipulated (Seglen, 1997).

As a comparison, CiteScore does not have the problem of inconsistent citation, and Scopus has a broader coverage of document than WoS. Besides, the value of the new CiteScore is higher than that of the previous CiteScore (Trapp, 2020) because more citations are included. Thus, the former has the potential to extend the citation window without reducing its value. However, the new CiteScore has a bias in favor of journals that have a high proportion of early citations within 4 years after publication (Fang, 2021). Therefore, the impacts of journals that have yearly citations that increase for a long time after publication are underestimated to a greater extent by the new CiteScore.

To effectively and fairly reflect the impact of individual journals with different citation patterns on the scientific community, Fang (2020) proposed an integral synchronic

JIF (IIF) that incorporates the features of synchronic and diachronic JIFs. Similar to synchronic JIFs, the IIF is calculated with the citations received in the focal year. Similar to the diachronic JIF, the IIF averages citations among papers published by a journal in the same year. The length of the citation window of the IIF can be set flexibly. In particular, the IIF can adopt a long citation window; thus, it can comprehensively reflect the citation history of the journal, whether its citation peak emerges quickly or long after publication. However, this method has a defect: it may be unstable if a journal publishes few papers in individual years.

Inspired by the new CiteScore, this work proposes a modification to the IIF. The research question is, how can the stability of the IIF be improved?

Journal indexes can also be modified from other perspectives, such as self-citation exclusion (de Lusignan & Moen, 2016), normalization mechanisms to correct for discrepancies in the citation density among different fields (Batista et al., 2006; Bornmann & Marx, 2015; Radicchi et al., 2008; Tijssen et al., 2002; van Leeuwen et al., 2003; Waltman et al., 2011), weighting citations according to the importance of the citing source (Alguliyev & Aliguliyev, 2017; Ferrer-Sapena et al., 2015), and geometric averaging mechanisms (Thelwall & Fairclough, 2015). The IIF can be combined with these characteristics (Fang, 2020). This applies to the modified method proposed here as well, and we do not repeat it in this work.

The next section provides the mathematical expressions of major existing journal indexes that consider only averaging citations of individual papers, proposes a modification to the IIF based on the analysis of the new CiteScore, and discusses the issues of the modified method. Section three performs experiments to verify the effect of the modified method. The last section discusses and summarizes this work.

Methodology

Mathematical expressions of citation-based journal indexes

Papers published by a journal (J) in a certain year (Y_p) constitute a set of papers, $J(Y_p)$. For example, Scientometrics(2020) refers to the set of papers published by *Scientometrics* in 2020. $c_J(Y_c, Y_p)$ denotes the number of citations that $J(Y_p)$ received in year Y_c.

The JIF with an n-year citation window of J in year Y can be written as:

$$JIF_{n,J}(Y) = \frac{\sum_{k=1}^{n} c_J(Y, Y-k)}{\sum_{k=1}^{n} p_J(Y-k)} \tag{1}$$

where $p_J(Y_p)$ is the number of papers published by J in Y_p.

The n-year (shifted) diachronic JIF of $J(Y_p)$ is (Frandsen & Rousseau, 2005):

$$IMP_{n,J}(Y_p) = \frac{\sum_{i=s}^{s+n-1} c_J(Y_p + i, Y_p)}{p_J(Y_p)}, \tag{2}$$

where $s = 0, 1, 2\ldots$ denotes a possible shift with respect to Y_p.

The new CiteScore of journal J in year Y can be written as:

$$CiteScore_J(Y) = \frac{\sum_{i=0}^{3}\sum_{j=i}^{3} c_J(Y-i, Y-j)}{\sum_{k=0}^{3} p_J(Y-k)} \tag{3}$$

The IIF (Fang, 2020) combines the features of synchronic and diachronic JIFs as:

$$IIF_J(Y) = \sum_{i=0}^{n} \frac{c_J(Y, Y-i)}{p_J(Y-i)} \tag{4}$$

The scheme of all the above indexes is to average the citations of related papers. To describe their differences, we here denote the cited window as the successive period in which papers whose citations are counted in the journal index were published. The citing window is denoted as the successive period in which the citations that were counted in the journal index occurred. The citing window of $JIF_{n,J}(Y)$ is Y, and its cited window is from $Y-n$ to $Y-1$. The citing window of $IMP_{n,J}(Y_p)$ is the n successive years beginning with Y_p+s, and its cited window is Y_p. The citing and cited windows of $CiteScore_J(Y)$ both range from $Y-3$ to Y. In $IIF_J(Y)$, the citing window of every summation term is Y, and the cited window of every summation term is a unique year between $Y-n$ and Y.

A potential modification to the new Scopus

Here, we define the publication-citation interval as the interval in years between publication and citation. An irrationality of JIF2, JIF5 and CiteScore is to average citations among papers with different publication-citation intervals, which causes different expected average citations of papers. The IIF and diachronic JIF do not have this defect. Each summation term of the IIF and diachronic JIF corresponds to a unique publication-citation interval.

The following modification can overcome this problem for CiteScore:

$$\begin{aligned} CiteScore_{1J}(Y) &= \sum_{i=0}^{3} \frac{\sum_{k=0}^{3-i} c_J(Y-k, Y-k-i)}{\sum_{k=0}^{3-i} p_J(Y-k-i)} \\ &= \frac{\sum_{k=0}^{3} c_J(Y-k, Y-k)}{\sum_{k=0}^{3} p_J(Y-k)} + \frac{\sum_{k=0}^{2} c_J(Y-k, Y-k-1)}{\sum_{k=0}^{2} p_J(Y-k-1)} \\ &\quad + \frac{\sum_{k=0}^{1} c_J(Y-k, Y-k-2)}{\sum_{k=0}^{1} p_J(Y-k-2)} + \frac{c_J(Y, Y-3)}{p_J(Y-3)} \end{aligned} \tag{5}$$

In Eq. 5, the four summation terms each correspond to a unique publication-citation interval. The bias in CiteScore (Fang, 2021) is also addressed because the four terms have the same weight.

A modification to the IIF

In Eq. 5, the first item estimates the average number of citations of a journal's papers received in the publication year with a 4-year cited window, and the remaining 3 items estimate the average number of citations of a journal's papers received in the i-th ($i = 1, 2, 3$) year following publication with a $4-i$-year cited window.

We can further modify Eq. 5 so that the four items have the same length of the cited window, as follows:

$$CiteScore_{2J}(Y) = \sum_{i=0}^{3} \frac{\sum_{k=0}^{3} c_J(Y-k, Y-k-i)}{\sum_{k=0}^{3} p_J(Y-k-i)} \tag{6}$$

This inspires us to modify the IIF to overcome its defect with the idea of using the citation counts that occurred in the adjacent previous years to assist in evaluating the impact level of the journal in the focal year. We denote the new index as JIc for journal J in year Y, defined as:

$$JIc_J(Y) = \sum_{i=0}^{n} \frac{\sum_{k=0}^{L} c_J(Y-k, Y-k-i)}{\sum_{k=0}^{L} p_J(Y-k-i)} \tag{7}$$

Equation 7 has $n+1$ summation terms, and each estimates the average number of citations received by the journal's papers with a unique publication-citation interval. Each publication-citation interval (i) uses the same length of the cited window (from $Y-L-i$ to $Y-i$), and the citations are received by the journal in $L+1$ successive year(s) (from $Y-L$ to Y) correspondingly. If $L=0$, $JIc_J(Y_p)$ becomes $IIF_J(Y_p)$, as in Eq. 4.

Figure 1 shows the citing relationship in the above journal indexes. In Fig. 1(a), the citations used in calculating JIF2 are generated in Y (citing window) for the papers published in the preceding 2 years ($Y-2$ and $Y-1$, as in the cited window). In Fig. 1(b), the citations used in calculating the n-year (shifted) diachronic JIF occurred in the years between Y_p+s and $Y_p+s+n-1$ (citing window) for the papers published in Y_p (cited window). In Fig. 1(c), the citations used in calculating the new CiteScore occurred in the years between $Y-3$ and Y (citing window) for the papers published in the same period (cited window). In Fig. 1(d), in each summation term of the IIF, the citations occurred in Y (citing window) for the papers published in a unique previous year (cited window). In Fig. 1(e), each summation term of JIc uses the citations that occurred in the $L+1$ successive years (ending at the focal year) and corresponds to a unique publication-citation interval. The citing window of each summation term consists of $L+1$ sub-citing-window, as does the cited window. Each sub-citing-window or sub-cited-window represents a calendar year. They form $L+1$ pairs of a sub-citing-window and a sub-cited-window. The time interval between the sub-citing-window and sub-cited-window in each pair is the same in each summation term. The time interval between the sub-citing-window and sub-cited-window in each pair is 0 for the first summation term (top of Fig. 1(e)), 1 year for the second term, …, and n years for the last term (bottom of Fig. 1(e)).

Overcoming the defect of the IIF

Although the IIF can fairly reflect the impact level of individual journals with different citation patterns because it can use a long citation window without decreasing its value, it may be unstable if a journal publishes few papers in individual years (Fang, 2020). To address this problem, Fang (2020) suggests combining adjacent publication years into a segment if the journal published fewer papers in these years. However, in such a segment, paper-citation pairs may have different publication-citation intervals. This also results in the absence of the relevant summation term in the calculation formula of the IIF. In contrast, in JIc, papers and the corresponding citations have the same publication-citation interval in each summation term, and no required summation term is missed. A requirement of setting the length of the cited window of JIc is that the number of papers in every

Fig. 1 Citation averaging scheme for JIF2 (**a**), the n-year (shifted) diachronic JIF (**b**), the new CiteScore (**c**), the integral synchronic JIF (**d**) and JIc proposed in this work (**e**). The dashed lines from the citing window to the cited window delimit a publication-citation period pair: the period in which citations occur and the period in which the papers that received the citations were published. Parameters are described in the main text

cited window should reach a threshold (such as 20) to ensure the stability of the index. This threshold is denoted as the JIc paper number threshold (JIc-PNT) here.

Setting JIc parameters

Setting the number of publication-citation intervals and the length of the cited window of JIc are subject to several restrictions. Suppose Y_{ini} is the first year in which the database records citations of journal J; then, $n+L \le Y-Y_{\text{ini}}$ for $JIc_J(Y)$ in Eq. 7.

As discussed above, the main difference between JIc and the IIF is that JIc has a longer cited window to avoid the instability of the index induced by a small number of papers published by some journals in individual years. However, if the cited window length $(L+1)$ is fixed, it may be short for journals that publish few papers per year and may be long for journals that publish many papers annually. In addition, many journals publish different numbers of papers in different years. To ensure that JIc is appropriate for different types of journals, JIc can have an adjustable cited window:

$$JIc_J(Y) = \sum_{i=0}^{n} \frac{\sum_{k=0}^{L_i} c_J(Y-k, Y-k-i)}{\sum_{k=0}^{L_i} p_J(Y-k-i)}. \tag{8}$$

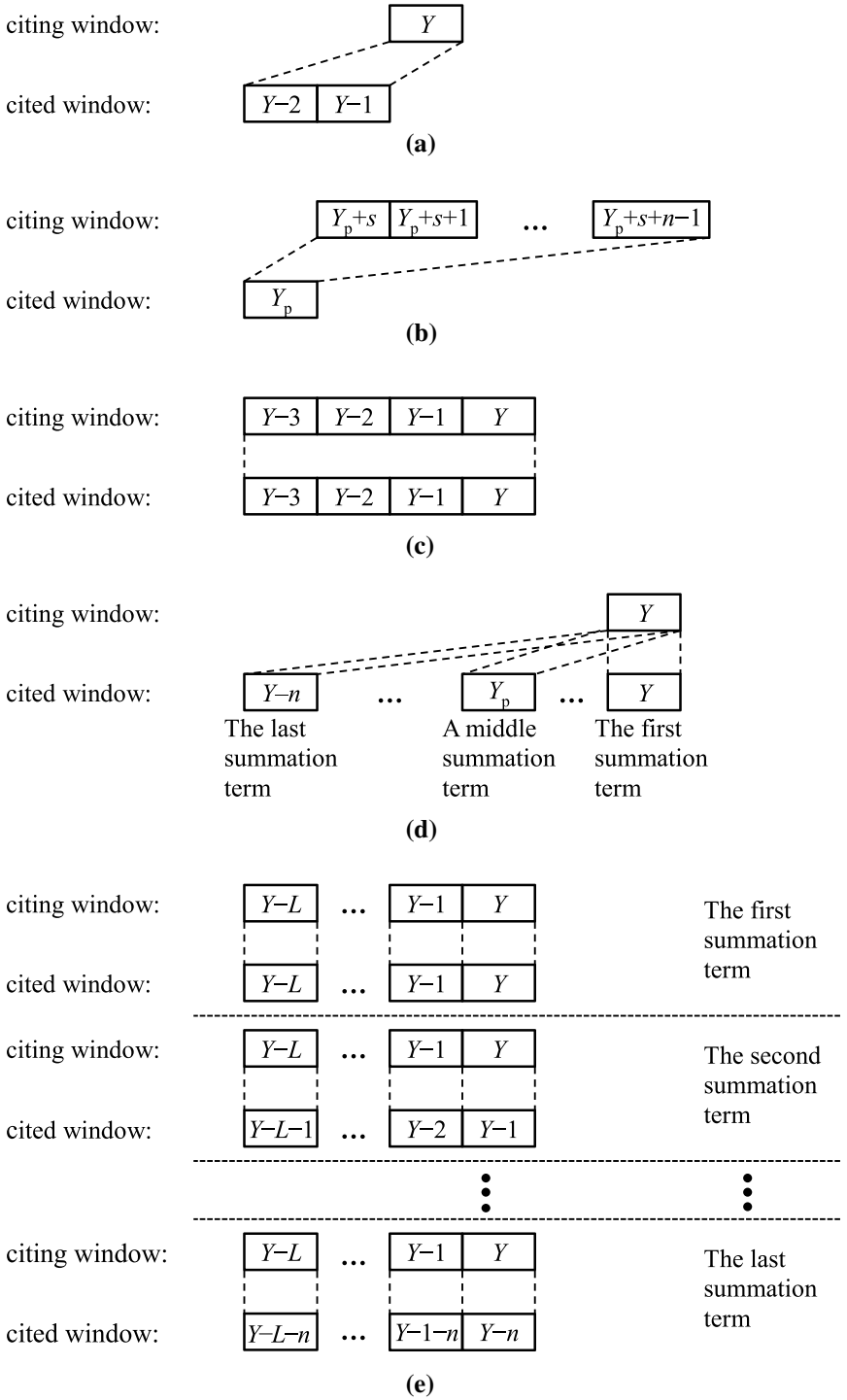
In Eq. 8, for each i, L_i+1 is the length of the shortest cited window (ending at $Y-i$) in which the number of papers published by J reaches JIc-PNT. Therefore, for different publication-citation intervals, the cited window length may be different. If a journal publishes many papers every year, enough to ensure the stability of the IIF, then there is no necessity to enlarge the cited window, and $L_i=0$ for each publication-citation interval.

The number of summation terms in Eq. 8 satisfies $n \le Y-Y_{\text{ini}}-L_n$. n can be set to a large value to make use of all the citation histories of a journal in the database, such as the WoS and Scopus. This method may incur the criticism that it would be beneficial to journals with long histories. Fang (2020) argues that this method is reasonable if journal indicators are regarded as a measure of journal reputation (Jacso, 2009) because the long history of a journal itself is prestigious.

Fang (2020) also provides solutions to the number of publication-citation intervals of the IIF if a journal indicator must have the same measurement conditions for all journals. These methods apply to JIc. For example, we can adopt a moderate upper limit of the publication-citation interval, such as 15. Up to this period, some delayed recognition papers with short dormancy periods will have been well recognized (Glänzel et al., 2003).

Experiment and results

To show the effect of the modification of the IIF (as JIc), we present some examples. The citation data of journal papers were collected from the "Citation Report" of the WoS Core Collection. To calculate the indexes of a journal, only citations of papers with a document


citing window:
cited window:
Y
Y−2
Y−1
(a)
citing window:
cited window:
Yp+s
Yp+s+1
…
Yp+s+n−1
Yp
(b)
citing window:
cited window:
Y−3
Y−2
Y−1
Y
(c)
citing window:
cited window:
Y
Y–n
…
Yp
…
The last summation term
A middle summation term
The first summation term
(d)
citing window:
cited window:
Y−L
…
Y−1
The first summation term
citing window:
cited window:
Y−L−1
Y−2
The second summation term
citing window:
cited window:
Y–L–n
Y−1−n
Y−n
The last summation term
(e)

type of "Articles" and "Review Articles" were counted. The data were collected on July 19, 2021. The papers and citations after 2020 are ignored. When calculating the IIF and JIc, n in Eqs. 4 and 8 is empirically set as 15. JIc-PNT for calculating JIc is empirically set as 20.

Figure 2 shows that the number of papers published by *Malacologia* in each year varies considerably. This number was zero in 1987, 1994 and 1997 and was only one in 2005.

Before 2005, the IIF is lower than JIc. One reason is that there are no papers in 1987, 1994 or 1997 indexed in the WoS; therefore, the summation terms in the IIF that use citations of papers published in these years are zero. However, the corresponding summation terms in JIc are not zero. For example, in calculating $IIF_{malacologia}(1993)$, the summation term corresponding to a publication-citation interval of 6 is 0 because no papers were published in 1987. In calculating $JIc_{malacologia}(1993)$, citations that occurred in 1992 for papers published in 1986 were used to estimate the impact level corresponding to a publication-citation interval of 6 for this journal. Therefore, the corresponding summation term is larger than 0. However, the IIF began to be larger than JIc in 2005 and then rose sharply, while JIc rose gradually, as shown in Fig. 2. After 2007, the IIF is more than two times higher than JIc. The reason is that the only paper published by *Malacologia* in 2005 happens to be highly cited. For the IIF after 2005, the citation of this paper obtained in the focal year is the value of one summation term, which is high. For JIc, this paper's citations obtained in the focal year are averaged together with 30 papers published in 2004, and none of these 30 papers is highly cited.

Figure 3 shows the citation trajectory of the only paper published in 2005 by *Malacologia* and the difference between the IIF and JIc. Between 2005 and 2017, the difference between the IIF and JIc is slightly less than the number of citations of the highly cited paper published in 2005, which mainly determines the change in the IIF curve between 2005 and 2020. After 2017, the difference between the IIF and JIc is higher than the number of citations of the paper published in 2005. The reason is that *Malacologia* published only three papers in 2017, and one of them is highly cited. The number of citations of this highly cited paper in 2017 is averaged among only three papers in the IIF, and thus, the corresponding summation term in IIF is further increased.

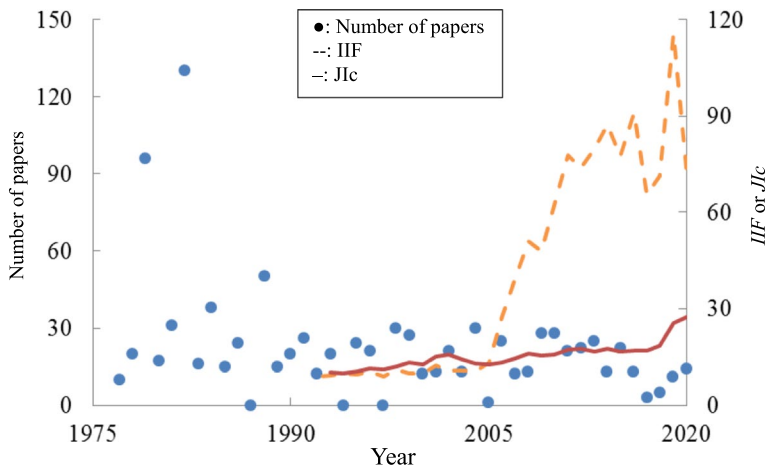


Fig. 2 Number of papers published by *Malacologia* in each year (●) and the integral synchronic JIF (IIF, dashed line) and JIc proposed in this work (solid line) of *Malacologia* for each year

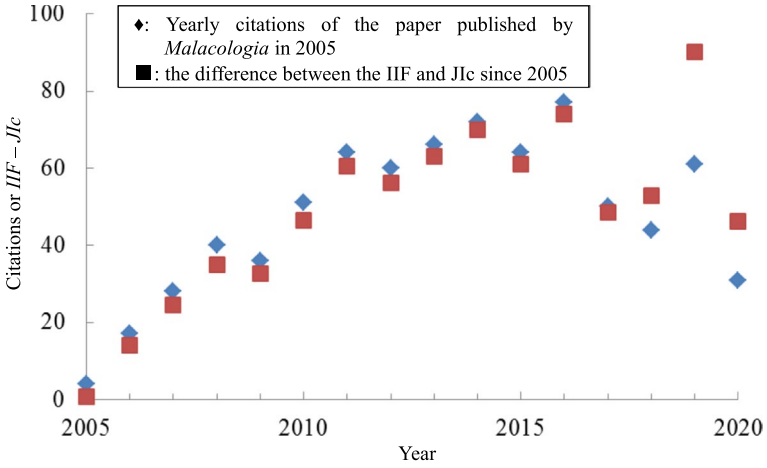


Fig. 3 Yearly citations of the paper published by *Malacologia* in 2005 (♦) and the difference between the IIF and JIc of *Malacologia* since 2005 (■). The IIF and JIc are defined in the main text

Figure 4 shows that the annual number of papers published by the *Journal of Classification* remained between 10 and 20 in most years. Before 2016, the difference between the IIF and JIc is small. After 2016, the IIF is obviously larger than JIc. The reason is that this journal published a highly cited paper in 2014, together with 15 other papers. The difference between the IIF and JIc for the *Journal of Classification* since 2016 is much less significant than that for *Malacologia* since 2006. The main reason is that the citation number of the highly cited paper published by the *Journal of Classification* in 2014 is averaged together with 15 "normal" papers in calculating the IIF. In contrast,

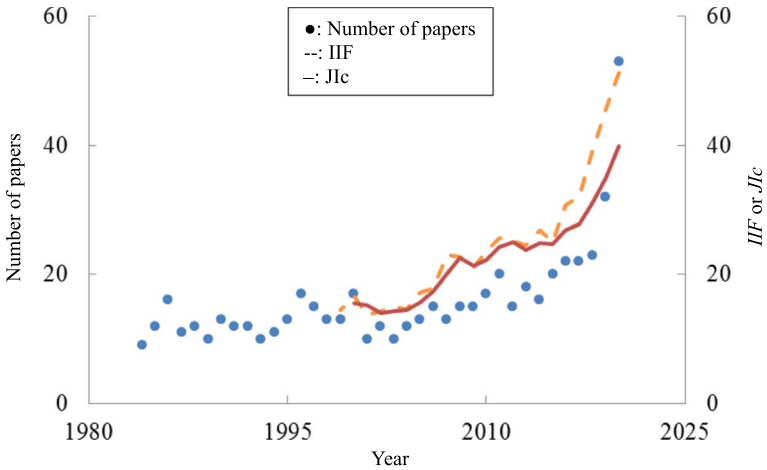


Fig. 4 Number of papers published by *Journal of Classification* in each year (●) and the integral synchronic JIF (IIF, dashed line) and JIc proposed in this work (solid line) of *Journal of Classification* for each year

the citation number of the highly cited paper published by *Malacologia* in 2005 is not "shared" by other papers in calculating the IIF.

In 1985 and 1996, the *Journal of Classification* published fewer than 20 papers, including two papers cited more highly than the highly cited paper published in 2014. These two highly cited papers do not obviously cause an increase in the IIF. The reason is that they were highly cited more than 15 years after publication, and in the experiment, we only considered citations with publication-citation intervals of no more than 15 years. Thus, the impact of these two highly cited papers is not reflected when the publication-citation interval is not sufficiently long.

Next, we inspect the effect of the modification of JIc on the IIF using a set of journals that satisfy the following conditions:

(1) The journal has at least one paper collected in the WoS every year since the first year it was collected by the WoS.
(2) There is at least one year during the period in which the journal was collected by the WoS in which the number of papers collected in the WoS is less than 20.
(3) In the first year the journal was collected by the WoS, its number of papers collected in the WoS is at least 20.
(4) The period in which the journal was collected by the WoS covers the years from 1996 to 2019.

Condition (1) is used to avoid missing summation terms when calculating the IIF. Condition (2) is used to avoid having the same value of the IIF and JIc in every year. Condition (3) is used to ensure that the journals each have the same length of the IIF and JIc. Condition (4) ensures that the IIF curve and JIc curve of the journal each have at least 10 points. With these four conditions, we screened 165 journals that had the same length but different values of eligible IIF and JIc. Of the 165 journals, if one journal was not collected by the WoS in 2020, it was first collected before 1996. Therefore, the IIF curve and JIc curve of the journal each have at least 10 points.

In this experiment, we compare the stability between the IIF and JIc. If an index is less volatile from year to year, it is more stable. We use the gradient to reflect the degree of fluctuation of the index curve:

$$\nabla z_J(Y) = \frac{z_J(Y) - z_J(Y-1)}{Y-(Y-1)} = z_J(Y) - z_J(Y-1), \tag{9}$$

where z represents the IIF or JIc. Using a gradient directly cannot indicate the relative change in the index. For example, suppose Δz equals 1; z changes by 10% if z equals 10, whereas z changes by only 1% if z equals 100. We use the relative gradient (Bettens & Collins, 1999; Le & Raff, 2010) to compare the changes in indexes with different values equally:

$$\nabla_{\mathrm{r}} z_J(Y) = \frac{z_J(Y) - z_J(Y-1)}{[z_J(Y) + z_J(Y-1)]/2}. \tag{10}$$

In Eq. 10, the denominator is not $z_J(Y)$ or $z_J(Y-1)$ but is the average of $z_J(Y)$ and $z_J(Y-1)$ to prevent the denominator from being zero if $z_J(Y)$ or $z_J(Y-1)$ is zero. If both $z_J(Y)$ and $z_J(Y-1)$ are zero, the relative gradient is zero because there is no change in value from $z_J(Y-1)$ to $z_J(Y)$.

Here, we define the sum of squares of the relative gradient of the index in different years as the overall fluctuation level of the index:

$$SRG\_z_J = \sum_{Y=Y_{\text{first}}+1}^{Y_{\text{last}}} \left(\nabla_{\text{r}} z_J(Y)\right)^2, \tag{11}$$

where Y_{first} and Y_{last} are the first and last years in which journal J has the index z. The larger the value of SRG_z is, the greater the fluctuation in the curve of z and the lower the stability of z.

Of the 165 journals screened, 137 journals had lower *SRG_JIc* than *SRG_IIF*, as shown in Fig. 5. One journal has same *SRG_JIc* and *SRG_IIF*. Figure 5(a) shows that the 27 journals that have higher *SRG_JIc* than *SRG_IIF* all have lengths of the IIF and JIc that are

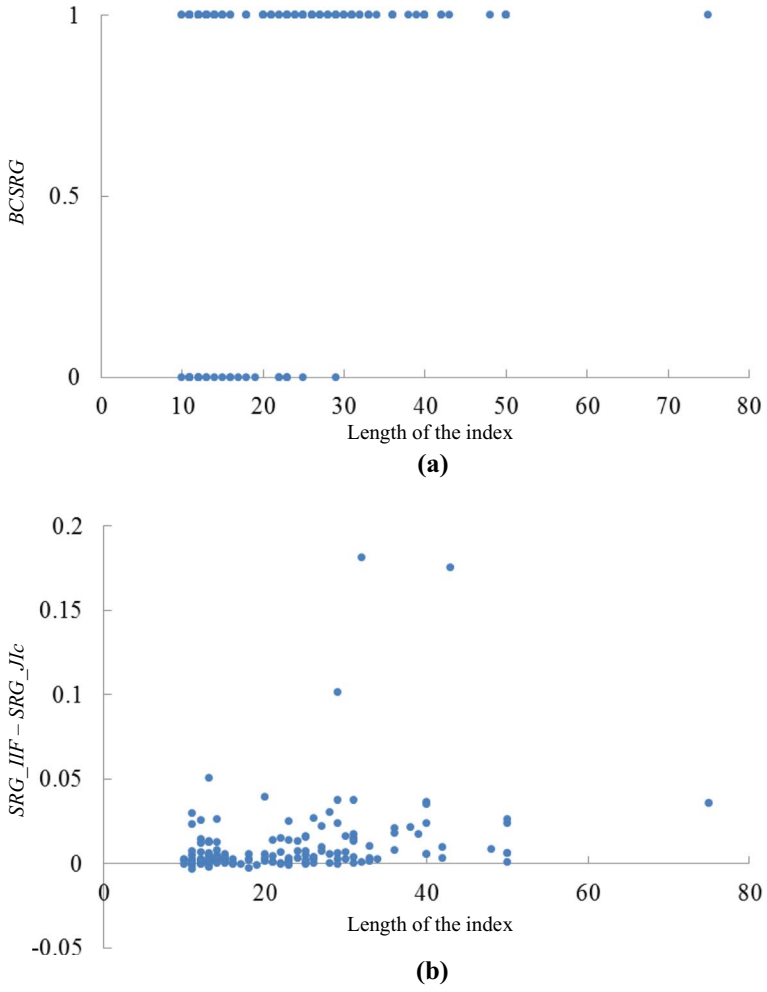


Fig. 5 Overall comparison of stability between the IIF and JIc. **a** Binarization comparison: if *SRG_IIF* > *SRG_JIc*, then *BCSRG* = 1; otherwise, *BCSRG* = 0; **b** The difference between *SRG_IIF* and *SRG_JIc*. *SRG_IIF* and *SRG_JIc* are defined in the main text

less than 30. Figure 5(b) shows that the difference between *SRG_JIc* and *SRG_IIF* for all the journals with higher *SRG_JIc* is much closer to 0 than that of journals with higher *SRG_IIF*.

The fact that some journals with short lengths of the IIF and JIc have higher *SRG_JIc* than *SRG_IIF* can be attributed to certain coincidental factors because there are more causes associated with the stability of the indexes, such as different distributions of the number and quality of papers published in different years. When the length of the index increases, the index curve can experience more potential coincidental cases that individually increase or decrease the fluctuation level of the IIF and JIc. Therefore, the performance of the IIF and JIc for journals with a long series of the two indexes, as shown in Fig. 5, verifies that the modification of JIc with respect to the IIF improves the stability of the index. Similarly, the combination of the performance of the IIF and JIc for the 165 screened journals also shows that JIc has higher stability than the IIF overall.

Discussion and conclusions

This work draws on the idea of the new CiteScore that citation counts from adjacent previous years can be used in estimating the journal index of the focal year to modify the IIF (Fang, 2020). The modified IIF proposed here, named JIc, incorporates the features of the JIFs and CiteScore. The key difference of the IIF and JIc from JIFs and CiteScore, which are used in practice to rank journals, is that the IIF and JIc average the citations among papers with the same publication-citation interval, which was first used in diachronic JIFs (Frandsen & Rousseau, 2005; Ingwersen, 2012; Ingwersen et al., 2001).

JIc inherits the advantages of the IIF, which can set the length of the citation window flexibly. Citation-based journal indexes with long citation windows can reflect the impact level more comprehensively than those with short citation windows. The citation window of the IIF and JIc can be enlarged without reducing the index value. JIc with a long cited window is fair for journals with a prolonged impact that is underestimated by a journal index with a short citation window. In addition, JIc is more stable than the IIF. The reason is that if a journal publishes few papers in the years involved in the calculation of JIc, JIc can be estimated with the help of the citation counts of the adjacent previous years.

The experiment results confirm that JIc is more stable than the IIF. In the experiment, n in Eqs. 4 and 8 is empirically set as 15. The reason for this setup is explained in the end of the methodology section. If this parameter is increased, the IIF and JIc reflect more citation information of a journal if the journal has a sufficiently enough citation history. The meaning of this parameter is the same for the IIF and JIc, and thus the change in its value does not affect the comparison between the IIF and JIc.

The IIF can be regarded as JIc with JIc-PNT being 0. The results show that a larger JIc-PNT makes JIc more stable. However, the degree of increase in the stability of JIc decreases with increasing JIc-PNT because the average result for singular value of citation numbers is inversely proportional to JIc-PNT for journals that publish only a few papers a year. Therefore, in the experiment, JIc-PNT is empirically set to a moderate value of 20.

Determining a more reasonable value of JIc-PNT needs further investigations. Because the distribution of citation counts for individual papers is highly skewed (Seglen, 1994), the percentage of highly cited papers is small (Thelwall & Fairclough, 2015). If a journal publishes many papers per year, then the papers it publishes in a year have a high probability of covering papers with a wide range of impacts, including highly cited papers. Then,

the skewedness of the citation distribution among papers causes little variation in averaging the citations among papers published in one year for such journals. However, for journals publishing few papers, the skewedness of the citation distribution among papers will cause large fluctuations if the index adopts a short citation window because papers published by such journals in every year may not cover papers with a wide range of impacts. The problem of determining JIc-PNT then becomes that of determining how many papers have a high probability of covering papers with a wide range of impacts. This requires comprehensively investigating the distribution of citations among papers. This issue awaits future work for perfecting citation-based journal indexes.

References

Alguliyev, R. M., & Aliguliyev, R. M. (2017). Modifications to the journal impact factor. *COLLNET Journal of Scientometrics and Information Management, 11*(1), 25–43.

Amin, M., & Mabe, M. (2000). Impact factors: Use and abuse. *Perspectives in Publishing, 1*, 1–6.

Batista, P. D., Campiteli, M. G., & Kinouchi, O. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics, 68*(1), 179–189.

Bettens, R. P. A., & Collins, M. A. (1999). Learning to interpolate molecular potential energy surfaces with confidence: A Bayesian approach. *Journal of Chemical Physics, 111*(3), 816–826.

Bornmann, L., & Marx, W. (2015). Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? *Journal of Informetrics, 9*(2), 408–418.

Cai, L., Tian, J., Liu, J., Bai, X., Lee, I., Kong, X., & Xia, F. (2019). Scholarly impact assessment: A survey of citation weighting solutions. *Scientometrics, 118*(2), 453–478.

Colledge, L., James, C., Azoulay, N., Meester, W., & Plume, A. (2017). CiteScore metrics are suitable to address different situations - A case study. *European Science Editing, 43*(2), 27–31.

de Lusignan, S., & Moen, A. (2016). Extracting oneself from the citation-stacking bear trap. *Methods of Information in Medicine, 55*(4), 301–302.

DORA. (2012). San Francisco declaration of research assessment. Retrieved October 19, 2022 from http://www.ascb.org/files/SFDeclarationFINAL.pdf.

Fang, H. (2020). Investigating the journal impact along the columns and rows of the publication-citation matrix. *Scientometrics, 125*(3), 2265–2282.

Fang, H. (2021). Analysis of the new Scopus CiteScore. *Scientometrics*. https://doi.org/10.1007/s11192-021-03964-5

Ferrer-Sapena, A., Sánchez-Pérez, E. A., González, L. M., Peset, F., & Aleixandre-Benavent, R. (2015). Mathematical properties of weighted impact factors based on measures of prestige of the citing journals. *Scientometrics, 105*(3), 2089–2108.

Frandsen, T. F., & Rousseau, R. (2005). Article impact calculated over arbitrary periods. *Journal of the American Society for Information Science and Technology, 56*(1), 58–62.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science, 178*(4060), 471–479.

Glänzel, W., Schlemmer, B., & Thijs, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics, 58*(3), 571–586.

Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science, 21*(1), 37–53.

Ingwersen, P. (2012). The pragmatics of a diachronic journal impact factor. *Scientometrics, 92*(2), 319–324.

Ingwersen, P., Larsen, B., Rousseau, R., & Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin, 46*(6), 524–528.

Jacso, P. (2009). Five-year impact factor data in the Journal Citation Reports. *Online Information Review, 33*(3), 603–614.

James, C., Colledge, L., Meester, W., Azoulay, N., & Plume, A. (2019). CiteScore metrics: Creating journal metrics from the scopus citation index. *Learned Publishing, 32*(4), 367–374.

Le, H. M., & Raff, L. M. (2010). Molecular dynamics investigation of the bimolecular reaction $BeH + H_2 \rightarrow BeH_2 + H$ on an ab initio potential-energy surface obtained using neural network methods with both potential and gradient accuracy determination. *Journal of Physical Chemistry A, 114*(1), 45–53.

Liu, X. Z., & Fang, H. (2020). A comparison among citation-based journal indicators and their relative changes with time. *Journal of Informetrics, 14*(1), 101007.

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America, 105*(45), 17268–17272.

Rossner, M., Van Epps, H., & Hill, E. (2007). Show me the data. *Journal of Cell Biology, 179*(6), 1091–1092.

Rousseau, R. (2002). Journal evaluation: Technical and practical issues. *Library Trends, 50*(3), 418–439.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*(9), 628–638.

Seglen, P. O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science, 45*(1), 1–11.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal, 314*(7079), 498–502.

Teixeira da Silva, J. A. (2020). CiteScore: Advances, evolution, applications, and limitations. *Publishing Research Quarterly, 36*(3), 459–468.

Thelwall, M., & Fairclough, R. (2015). Geometric journal impact factors correcting for individual highly cited articles. *Journal of Informetrics, 9*(2), 263–272.

Tijssen, R. J., Visser, M. S., & Van Leeuwen, T. N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics, 54*(3), 381–397.

Trapp, J. V. (2020). The new Scopus CiteScore formula and the Journal Impact Factor: A look at top ranking journals and middle ranking journals in the Scopus categories of General Physics and Astronomy, Materials Science, General Medicine and Social Sciences. *Physical and Engineering Sciences in Medicine, 43*(3), 739–748.

van Leeuwen, T. N., Moed, H. F., & Reedijk, J. (1999). Critical comments on Institute for Scientific Information impact factors: A sample of inorganic molecular chemistry journals. *Journal of Information Science, 25*(6), 489–498.

van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics, 57*(2), 257–280.

Vanclay, J. K. (2009). Bias in the journal impact factor. *Scientometrics, 78*(1), 3–12.

Vanclay, J. K. (2012). Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics, 92*(2), 211–238.

Waltman, L., Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. (2011). Towards a new crown indicator: An empirical analysis. *Scientometrics, 87*(3), 467–481.

Zupanc, G. K. H. (2014). Impact beyond the impact factor. *Journal of Comparative Physiology A-Neuroethology Sensory Neural and Behavioral Physiology, 200*(2), 113–116.