The pragmatics of a diachronic journal impact factor

Peter Ingwersen

Abstract   With reference to Vanclay (Scientometrics in press, 2012) the paper argues for a pragmatic approach to the Thomson-Reuter's journal impact factor. The paper proposes and discusses to replace the current synchronous Thomson-Reuter journal impact factor by an up-to-date diachronic version (DJIF), consisting of a three-year citation window over a one year publication window. The DJIF online data collection and calculation is exemplified and compared to the present synchronous journal impact factor. The paper discusses briefly the dimensions of currency, robustness, understandability and comparability to other impact factors used in research evaluation.

Keywords   Journal impact factor · Diachronic impact factor · Synchronous JIF

Introduction

The journal impact factor, also generically named the Garfield Index and in its ISI/Thomson Reuters form as TRIF by Vanclay (2012), should *ideally* serve several purposes to the scientific community. It should signal a value of recognition of a given journal by peers, be up-to-date and be comparable to other journal scores and other impact factors assigned to other kinds of units forming part of the scientific communication process, such as institutions or scientists. It should be calculated based on rigorous data collection and analysis and signify that the journal in question is indeed peer reviewed and maintain a certain level of good editorial practice, for instance in relation to duplicate submissions, plagiarism and fraud.

In this sense I endorse most of the proposals by Vanclay (2012, p. 18) on the renewal or profound modification of the TRIF. However, I have doubt that Thomson Reuters or any

P. Ingwersen (✉)
Department of Library Science and Documentation, University Carlos III Madrid,
Calle Madrid, 126, 28903 Getafe, Madrid, Spain
e-mail: pi@iva.dk

P. Ingwersen
Royal School of Library and Information Science, Birketinget 6, 2300 Copenhagen S, Denmark

other commercial or non-commercial body could implement a rigorous and fair certification of journals, which is not based on journal editors' information on their own editorial policy. Like the scientists themselves as readers of and submitters to journals a certification agency must somehow rely on (and trust) the editorial statements of the journals. But why should the same editors who, according to Vanclay (2012), by all means strive for maximizing their own TRIF not also publicly assure rigorous quality control in their editorial practice? They could easily state one thing and in practice do something else. Only *de facto* incidents of science ethical nature might lead to some kind of measure taken against a journal, and who should then implement such measures? The academic community (as done now), a particular (non-existing) international agency or a commercial database vendor?

Consequently, I find Vanclay's proposal on the certification issue falling outside the central discussion and practice of the Garfield Index and the TRIF in particular. In my opinion, that discussion basically concerns the dimensions of conceptual nature, comparability, utility and calculation practice of the TRIF. Regardless the conceptualization and mode of calculation, the problems of author as well as citation indexing errors (by the commercial vendor) remain and should be improved. The points and illustrations made in this respect by Vanclay (2012) are just and well put. However, in my perspective the TRIF should preferably take a *diachronic* form and thus be calculated differently from its present synchronous mode (Ingwersen et al. 2001) in order to be better understood, comparable to other central indicators and useful in research evaluation.

The article is organized as follows. The diachronic form of the TRIF (DJIF) is outlined and compared to the present synchronous one, followed by a discussion of the advantage of the former when used in combination with other common research evaluation indicators, such as individual or weighted normalized institutional or country-based citation impact indicators.

The diachronic journal impact factor

Conceptually the current TRIF, as all the Garfield Indexes, is made in a *synchronous* form—Fig. 1. It starts at a given point in time, say year $y = 2011$, and calculates the number of citations given to the publications of a journal *back in time* in the preceding two years, Y-1 and y-2 (2010 and 2009). Already Christensen and Ingwersen (1996) and Ingwersen et al. (2000) pointed out the advantage of a diachronic journal impact
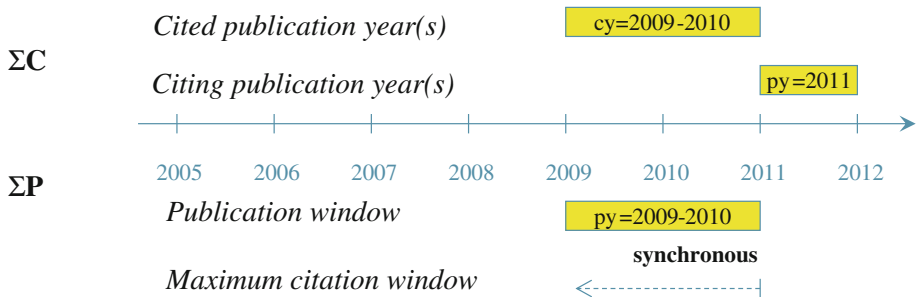
Fig. 1 Synchronous TRIF 2011 illustration for 2009–2010 publications; TRIF publishable April 2012. *C* citations, *P* publications (revised version from Ingwersen et al. 2000)
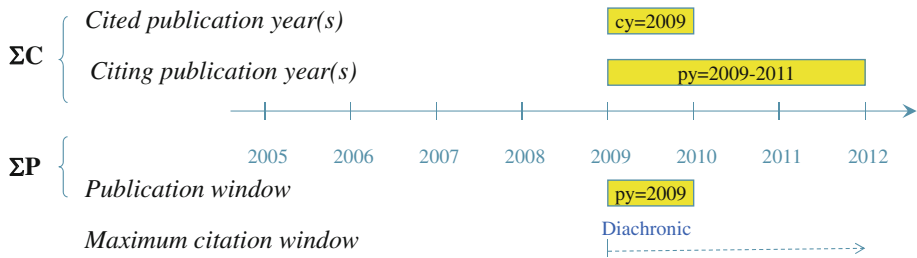
Fig. 2 Diachronic JIF illustration, 3-year citation window for 2009 publications; DJIF publishable April 2012. *C* citations, *P* publications (revised version from Ingwersen et al. 2000)

factor—Fig. 2, in particular for the purpose of online calculation and comparison to other citation impact based indicators, such as impact of institutions or scientists. Diachronic citation analysis is commonly used when calculating the cited half-life of a journal or a specific article (Egghe and Rousseau 1990). Such analyses are historical in the sense that one calculates the total amount of citations given annually to an article *forward* in time. As stated in Ingwersen et al. (2000, p. 373), in the diachronic analysis it is possible to compare all citations or the cited half-life of a journal article with the corresponding amount of citations for the journal itself. This comparison is not possible in a synchronous citation analysis, and a comparison between a diachronic article citation analysis and a synchronous journal citation analysis is likewise meaningless.

For both the synchronous and diachronic JIF the data collection can be done manually online via Web of Science (WoS) or the Thomson Dialog versions of SCI and SSCI. The factor is consequently reproducible. However, the analyst has more control over the data collection if the WoS *Search* mode is applied than if the *Cited Reference* search mode is used. The difference is that in the *Search* mode one initially selects and isolates the citable items (articles, review articles, conference papers, letters and notes) for the given journal in the given publication window—Figs. 1 and 2. It is vital to enter the correct publication years in the query box in Search mode *and* to click on the relevant citation years in the 'from' and 'to' box to form the citation window further down the interface page.[1] Thus, the denominator of the impact factor calculation is restricted to the relevant citable types of publications. Prior to performing a *Citation Report* analysis of the set it can be sorted by citations. It is now possible to detect the *citedness* ratio of the set of citable publications, i.e. how many items that actually contribute to the current diachronic TRIF.

A Citation report of the set returns the number of citations given in the citation window to the journal's citable publications. However, the citation volume (in the enumerator) is *unrestricted* in the sense that the citing documents contain all kinds of document types, including non-scientific book reviews and editorial materials. One can manually detect such types by selecting and viewing the set of *citing* documents. By displaying the document types of the set through WoS' rank-display facility (left-hand side of webpage) one may refine the set to isolate the smaller number of non-scientific material and thus be able to calculate and subtract the number of citations that derives from such materials. Journal self-citations can easily be retrieved and its ratio calculated by additional searching.

---

[1] In case of synchronous TRIF for logical reasons in WoS the citation window must cover the publication years in Search mode; then the citation number from the given citation year, say 2011, can be extracted from the result table of the Citation Report.

Table 1 Comparison of DJIF and TRIF for *Journal of Documentation* publications 2009, cited 2009–2011 (DJIF) and publications 2009–2010, cited 2011 (TRIF)

	Total items (cited)	Citable items (cited)	All citations	Citations restricted	All citing items	Citing items restricted	IF all citations	IF restricted
DJIF[a]	74 (34)	42 (32)	85	82	73	70	2.02	1.95
DJIF[b]	–	40 (32)	86	82	73	69	2.15	2.05
TRIF[a]	133 (54)	86 (52)	69	67	60	58	0.80	0.78
TRIF[b]	–	51 (43)	121	–	110	97	2.37	–

Data collected on February 3, 2012

[a] WoS search mode, [b] WoS cited reference search mode

This way of data collection for DJIF is reproducible and maintains some fault-tolerance, since it includes all citations matching the journal title in question and publication/citation year(s), even if citation errors exist. By this diachronic approach citing as well as cited-side controls are introduced to a degree but probably more relaxed than ideally proposed and described in Vanclay (2012). The synchronous TRIF mode is slightly more cumbersome to produce in *Search* mode. If the Cited Reference search facility in WoS is used, the analysis become really cumbersome to perform and the analyst has less transparency and control of the outcome—see comparative example, Table 1.

Table 1 demonstrates that aside from potential errors associated with the author's typing and Thomson-Reuter's (or other database providers') indexing procedure the way of searching in the same database may also produce very different IF scores, depending on the manual retrieval procedure and underlying vendor algorithms. The most conservative (and most transparent and two-side restricted) estimate of the DJIF is the (a) version applying the *Search* facility in WoS (1.95). In the (b) version it is not possible directly to restrict the citing side (the enumerator). This has to be done in more cumbersome ways, and the vendor algorithm does not retrieve all citable items (40 vs. 42 units). In the TRIF (b) version applying the *Cited Reference* search tool the WoS algorithms retrieves *all* citations given the two publication years (2009–2010) up including 2011, i.e. also the citations obtained in 2009 and 2010. Hence the large amount of unrestricted citations (121). This unfortunate algorithmic phenomenon is explicitly described by Thomson-Reuters on the WoS interface; but analysts should take care not to apply the displayed number of citations. The number of citable publications in TRIF (b) is almost identical to the number found cited in the retrieval version TRIF (a) (51 vs. 52), but 8 items cited are in fact not of the citable kind (no DOI are assigned to the cited items). We observe here the same kind of problematic data twisting as displayed by Vanclay (2012, p. 13, Fig. 10). Problems may thus arise in the calculations if automated data collection takes place.

Comparability of the diachronic JIF

The most important dimensions of the DJIF, aside from the technical data collection and calculation issues, associates with the following issues:

1. Currency and sufficient data availability for calculation and validity.
2. Easy to understand for scientists and policy makers.
3. Easy to compare to other citation-based indicators.

The proposed DJIF operates with a three year citation window, i.e. one year more than the traditional TRIF annually published in JCR, but is just as much up-to-date for the published items cited, see Figs. 1 and 2. This facet of *currency* of the impact factor is vital for a commercial vendor of the product. The 5-year synchronous TRIF also given by JCR does not serve any real purpose because the published citable items are too remote in time, but was probably published because it is easy to calculate for Thomson-Reuters. The proposed three-year citation window assures an adequate citation volume for the calculation of the DJIF. From Table 1 it is clear that even for a smaller-size journal like *Journal of Documentation* published in a citation-low field, three years of citations to a one-year journal volume forms a more robust data foundation than the one-to-two year citation window in the present TRIF (a) (82 citations vs. 67). We include the document type restrictions introduced above in the reproducible online calculation mode.

This issue of satisfactory volume is important since for many research fields in science the cited half-life is 2–4 years. Thus, over a three-year period a major bulk of citations may hence be expected to be available. In some social science and humanity fields this may take longer time. It is an advantage from a utility point of view that the same citation window is used for all the fields. From a diachronic perspective it is quite easy to apply differentiated window length through, for instance the cited half-life of each discipline, as suggested by Rousseau (1998) for mathematics. This may probably produce a more balanced and just impact factor, but makes its comprehension among academics more difficult and comparative analyses even more cumbersome than today.

The uninterrupted three year citation window that provides an up-to-date citation history in the DJIF may in addition make it easier to understand what is done during calculation and how to interpret the results. If a scientist has published a citable item in 2009 it is easy to understand that one calculates the number of citations from 2009 up to include last year's citations. Further, it is rather easy to make scientists and policy makers understand the simplistic comparison between the given article's citation amount from 2009 to present and the corresponding DJIF for the journal publishing the article. Both are diachronic, covering the same citation window and the latter signifying the *expected* number of citations an *average* article published in that journal ought to receive.

In contrast the synchronous nature of the Garfield indexes makes them incomparable to other unit's impact for a given period—and also quite uncertain. Which TRIF year should one compare to as the correct one if a citable item has been published in a journal in 2009? TRIF 2011 would provide a higher score than TRIF 2010 since the former is less immediate than the latter, but both covers citation-wise 2009. And there is the empty time slot of two years of citations (2009–2010) for which no citations are counted in the traditional TRIF 2011[2]— Fig. 1. That is the reason for its incomparability to other kinds of citation impact factors.

Also Vanclay (2012) is aware of the misuse of applying the TRIF (or DJIF) value as a research evaluation tool for individual researchers or institutions, replacing the *real* amount of citations received by the TRIF/DJIF values of the journals used by the scientist or institution. The method involves assigning the relevant TRIF/DJIF score to each item published by a person or the institution. This procedure is unethical and unfortunate to all involved since (1) the impact factor is a short-term impact measure and, by replacing the real score by an expected one, (2) it is assumed that *all* articles in each journal actually obtain citations and that (3) the scientist is an *average scientist*. Already Seglen (1995, 1997) demonstrated that only few articles per volume obtain many citations and thus contribute positively to the journal impact while many receives very few or none. He also

---

[2] In TRIF 2010 one year of citations are lacking (from 2009).

argued against this unethical (and unwise) misuse of such scores in research performance evaluations (Seglen 1997).

By applying the real number of citations received diachronically by the citable items involved in a research evaluation process *and* comparing to the corresponding diachronic DJIFs of the journals used for publishing the items the analyst obtain *relative* research performance indicators that are normalized and robust. Via adequate grouping of journals (and aggregation of their DJIF scores) one obtains diachronic research field impact scores. The comparison between the latter field scores and a given DJIF (over the same time period) provides information about the impact-defined journal quality (high/low) in context of the expected field citation impact. This has been common knowledge in the sciento-metric community for many decades (Moed 2005). Thus, one can only urge the citation database vendors to produce diachronic-based impact factors and to introduce more rig-orous input control than done at present.

Concluding remarks

It is evident that a diachronic journal impact factor can be misused in line with the present synchronous one. Also the DJIF may suffer from author and vendor errors. The proposed DJIF has the pragmatic advantage of being citing and cited side restricted to proper document types and being reproducible in the WoS database by means of the Search facility tool. However, it does not fulfill all the ideal requirements put forward by Vanclay (2012). It does not claim to hold a one-to-one relationship between citing and cited items.

The disadvantage is that the DJIF (like the TRIF and other arithmetic average-based impact factors) displays properties that make it vulnerable in a mathematical sense as discussed among others recently in (Rousseau and Leydesdorff 2011). Also, in line with the present TRIF version, it informs about historic events going three years back when published. This is unavoidable but the DJIF may indeed be differentiated across disci-plines, if required. The proposed DJIF is much more straight-forward in its form than the TRIF, thus much more understandable for scientists and policy makers.

References

Christensen, F. H., & Ingwersen, P. (1996). Online citation analysis—A methodological approach. *Scientometrics, 37*(1), 39–62.

Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics*. Amsterdam: Elsevier.

Ingwersen, P., Larsen, B., & Wormell, I. (2000). Applying diachronic citation analysis to research program evaluations. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: Festschrift for Eugene Garfield* (pp. 373–388). Medford: Information Today.

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

Moed, H. (2005). Citation analysis in research evaluation. Dordrecht: Springer.

Rousseau, R. (1998). Citation distribution in pure mathematics journals. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88* (pp. 249–260). Amsterdam: Elsevier Science.

Rousseau, R., & Leydesdorff, L. (2011). Simple arithmetic versus intuitive understanding: The case of the impact factor. *ISSI Newsletter, 7*(1), 10–14.

Seglen, P. O. (1995). 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 evaluation research. *British Medical Journal, 314*, 498–502.

Vanclay, J. K. (2012). Impact factor: Outdated artefact or stepping-stone to journal certification? Sciento-metrics (in press).