

Paper presented at the 4th International Conference on University Evaluation and Research Evaluation, Wuhan University (China), September 27-October 1, 2004

Impact factors and databases as instruments for research evaluation

Ronald Rousseau

KHBO, Industrial Sciences and Technology (IWT),
Zeedijk 101, B-8400 Oostende, Belgium
Email: ronald.rousseau@khbo.be

&

University of Antwerp (UA), Information and Library Science (IBW)
Universiteitsplein 1, B-2610 Wilrijk, Belgium

Abstract

Starting from the question “What should be the real purpose of research evaluation?”, I focus on the use of instruments for science policy purposes, presenting an overview of several types of impact factors as used in the scientometric literature. The recently introduced median impact factor receives special attention. I promote the use of citation-publication tables as a visual aid for thinking about impact. Mathematical properties, such as the general pool decomposition theorem are highlighted. Better instruments, i.e. better artefacts and mental schemes, make for better evaluations. Consequently, much more should be done in order to make publication and citation databases optimal instruments for research evaluation.

Key words: research evaluation, impact factors, instrument, artefact, mental scheme, publication-citation table, median impact factor, mathematical properties

1. Introduction

All over the world and in all domains of human behaviour we experience a tendency for short-term “value for money”. Regrettably this is also true in the fields of education and research. Scientists must publish, the more the better, and this – almost exclusively – in journals covered by the Institute for Scientific Information (ISI). Note that I think that it is only natural that scientists obtain

results, publish, and strive for a wide audience. The point is that not all benefits are short-term, and not all interesting journals are covered by ISI. We often see that in evaluation exercises these other journals, congress proceedings, even books are not taken into account. In Europe, money for scientific research is often distributed through international projects which must involve industrial partners and collaborators of different countries. It is as if an individual scientist can do no valuable research anymore.

This leads me to the question: What should be the real purpose of research evaluation? The answer, as given in (Russell & Rousseau, 2002) is that the real goal of any form of research evaluation is providing those people and institutions with the talent and motivations to carry out scientific research, with the best conditions possible under which to do so. Of course, budgetary and other kinds of constraints make evaluations necessary for the equitable distribution of resources (Jin & Rousseau, 2004). The evaluation of short-term strategic research as well as long-term curiosity-driven research for new knowledge demands the same accountability and rigorous standards as scientific research requires of itself. For this reason the challenge is not only for the application of bibliometric and scientometric techniques in research evaluation to keep up with the rapid changes occurring in scientific communication patterns and practices but also to constantly improve the theoretical foundation for the construction of all kinds of indicators needed for peer review (Russell-Rousseau, 2002).

In order to perform evaluations instruments are needed. But what is an instrument? An instrument can be defined as an artefact plus a mental scheme (Vygotski, 1978; Drijvers & Gravemeijer, 2004). An artefact (in Chinese: Ren2 Zao4 Wu4 Pin3 人造物品) is just a thing. A typical example of an artefact is a previously unknown object found on an archaeological site. No one knows its use: it is just an artefact, a thing. An artefact becomes an instrument when knowledge and understanding about the proper use of the artefact develops in a person's mind. This knowledge is called a mental scheme. For most people a computer once was an artefact; knowledge about how to use it makes it an instrument. Clearly with such a complex artefact, knowledge is often partial, so it often happens that the instrument is not used to its full capacity.

Similarly, the SCI, Journal Citation Reports (JCR), CSCD, CSTPC and the CSSCI are just artefacts. Only a mental scheme, i.e. thorough understanding of their functionalities, makes these databases into instruments fit for evaluation of scientific research.

In this article I will try to augment readers' mental schemes about such databases by focusing on the mathematical properties of the impact factor, and related indicators that can be derived from artefacts such as the Web of Science, the CSTPC, and the CSCD (Jin & Wang, 1999; Wu et al., 2004).

2. Citation impact and the publication-citation table

Citations over time can best be studied using the so-called publication-citation table (Ingwersen et al., 2001; Liang, 1992). Consider the following hypothetical publication-citation table (Table 1).

Table 1: A publication-citation table for a hypothetical set of articles: cited dimension

Publication year	1999	2000	2001	2002	2003	2004
# publications	20	25	30	30	40	45
# citations received in the year 1999	8					
# citations received in the year 2000	14	5				
# citations received in the year 2001	25	14	8			
# citations received in the year 2002	35	20	12	8		
# citations received in the year 2003	25	26	18	11	6	
# citations received in the year 2004	17	20	22	14	6	3

The first row of Table 1 gives the publication year, the second row the number of articles published yearly by this particular article set. The other rows are citation rows. We see that, e.g. in the year 2003 this set of articles received 18 citations

for articles it published in the year 2001. In 2004 it received 3 citations for articles published that same year. Note that I refer to a 'set of articles'. Such a set can be a journal, but can also be any other meaningful collection of articles.

Based on such a publication-citation matrix, it is now easy to explain the mathematical form of the ISI impact factor. Table 2 illustrates the elements needed in order to calculate the ISI impact factor for this set of articles. Recall that citations always originate from a pool of articles. In the case of the JCR this pool is the set of all journals covered by ISI.

Table 2 Elements for the calculation of an ISI impact factor for the year 2003.

Publication year	1999	2000	2001	2002	2003	2004
# publications	20	25	30	30	40	45
# citations received in the year 1999	8					
# citations received in the year 2000	14	5				IF(2003) =
# citations received in the year 2001	25	14	8			(18+11)/(30+30)
# citations received in the year 2002	35	20	12	8		= 0.483
# citations received in the year 2003	25	26	18	11	6	
# citations received in the year 2004	17	20	22	14	6	3

The ISI or Garfield-Sher impact factor is a synchronous impact factor, which means that citations used for its calculation are given in the same year. In this example the year 2003 has been considered. Hence citation data can all be

found in the same row. The other citation rows play no role at all for this particular calculation. The ISI impact factor uses a 2-year window, but it is obvious that other windows are possible. This leads then to the following general definition:

Definition (Rousseau, 1988). The synchronous journal impact factor for the year Y is defined as

$$IF_n^s(Y) = \frac{\sum_{i=s}^{s+n-1} CIT(Y-i, Y)}{\sum_{i=s}^{s+n-1} PUB(Y-i)}$$

where $PUB(Y-i)$ denotes the number of sources included in the set of articles under study, published in the year $Y-i$; and $CIT(Y-i, Y)$ denotes the number of citations received by this set of articles in the year Y , to articles published in the year $Y-i$, s is called the offset period. If necessary we also show the journal J we are considering:

$$IF_n^s(J, Y) = \frac{\sum_{i=s}^{s+n-1} CIT(J, Y-i, Y)}{\sum_{i=s}^{s+n-1} PUB(J, Y-i)}$$

The classic Garfield-Sher, or ISI, impact factor is a synchronous impact factor with offset $s = 1$ and publication window equal to the citation window, equal to two ($n = 2$ in the formula). In the new notation it becomes $IF_2^1(Y)$.

In research evaluation exercises, however, the use of a diachronous impact factor is often preferred (Ingwersen et al., 2001; Russell & Rousseau, 2002). Table 3 illustrates the definition of this other type of impact factor.

Table 3. Elements for the calculation of a 3-year diachronous impact factor for the year 2002

Publication year	1999	2000	2001	2002	2003	2004
# publications	20	25	30	30	40	45
# citations received in the year 1999	8					
# citations received in the year 2000	14	5				DIF₃(2002) =
# citations received in the year 2001	25	14	8			(8+11+14)/30
# citations received in the year 2002	35	20	12	8		= 1.10
# citations received in the year 2003	25	26	18	11	6	
# citations received in the year 2004	17	20	22	14	6	3

The general formula for the calculation of a diachronous impact factor is given in the following definition.

Definition. The diachronous n-year journal impact factor for the year Y is defined as

$$DIF_n^s(Y) = \frac{\sum_{k=s}^{s+n-1} CIT(Y, Y+k)}{PUB(Y)}$$

where $PUB(Y)$ denotes the number of sources included in the set of articles under study, published in the year Y ; and $CIT(Y, Y+k)$ denotes the number of citations received by this set of articles in the year $Y+k$, to articles published in the year Y , s is again an offset period. For diachronous impact factors the offset period is usually 0. If necessary we also show the journal J we are considering:

$$DIF_n^s(J, Y) = \frac{\sum_{k=s}^{s+n-1} CIT(J, Y, Y+k)}{PUB(J, Y)}$$

The so-called immediacy index as published by ISI in the Journal Citation Reports is a synchronous, as well as a diachronous impact factor. In our notation it is either IF_1^0 or DIF_1^0 .

We have already stated that impact factors may not only be calculated for journals, but also for other sets of articles. We note, in particular, that diachronous impact can easily be calculated for conference proceedings. This has been done in (Rousseau, 1997). Clearly, the concept of 'impact' can be applied in most source-item environments. It is, moreover, not necessary that the whole table is known. One item row suffices for the calculation of a synchronous impact, while one item column suffices for the calculation of a diachronous impact.

Once the publication-citation table is accepted as the framework in which citations are considered, it becomes natural to consider more than one row or one column. This leads to the generalized impact factor as described by Frandsen & Rousseau (2005).

Let us consider some examples of general source-item relations.

a) Scientific journal \rightarrow (publishes) articles \rightarrow (receives) citations

A scientific journal considered as a set of articles is the source. The items produced by this source are the citations received by the articles. All this happens, of course, within a certain timeframe and a certain pool of citing documents. The ratio (# citations)/ (# articles) is then known as an impact factor, more precisely, a journal article impact factor.

Similarly we may consider the following variations:

b) Scientific journal \rightarrow (publishes) issues \rightarrow (receives) citations

In this case the journal is considered as a set of issues (the source). Now issues receive citations, not individual articles, and one may define, similar to the journal

article impact factor, the ratio (# citations)/ (# issues) as a journal issue impact factor. As far as I know this has not been done yet, but it might be interesting to study the difference between a journal issue impact factor and a journal article impact factor. Clearly, the journal issue impact factor will always be smaller, but it is the comparison of different journals that might lead to new insights. Recall that in case a) one article cites another article at most once. Here, in case b), it is understood that one article cites one issue at most once.

c) Scientific journal → (publishes) one particular issue → (receives) citations

Here, one special issue is the source. This situation leads to the definition of an issue article impact factor, defined as (# citations)/ (# articles in this particular issue). This notion is interesting for a study of special issues, dedicated to one particular theme. Note that for this case only a diachronous impact factor can be studied, a synchronous impact factor makes little sense.

d) Considering one particular special issue of a journal, or a book published as a congress proceedings, such as the *Proceedings of the 9th International Conference on Scientometrics and Informetrics* (Jiang et al., 2003) is conceptually the same. This leads to the following situation:

Congress proceedings → (consists of) articles → (receives) citations

The source is the set of articles making up a proceedings volume, items are citations received by these articles. This leads to the proceedings (diachronous) impact factor.

3. A new approach: the median impact factor (MIF)

Recently, a group of Thai colleagues proposed a new kind of impact factor, the cited half-life impact factor, or better median impact factor, in short MIF (Sombatsompop et al., 2004; Rousseau, 2004). I think the idea introduced by Sombatsompop, Markpin and Premkamolnetr has a great theoretical value as these colleagues are – to the best of my knowledge – the first to consider impact factors not in terms of years as a basic ingredient, but in terms of the actual form of the citation curve. Taking the citation curve into account makes this impact factor better suited for a comparison of impact among fields and subfields. It is defined as follows:

$$MIF_J(Y) = \frac{TOT_J(Y)/2}{CPUB_J(Y-X, Y)}$$

In this formula, $TOT_J(Y)$ denotes the total number of citations received by journal J in the year Y; $CPUB(Y-X, Y)$ denotes the cumulative number of publications in the journal J, during the period $[Y-X, Y]$, where X denotes the median cited age. As citations are collected in the same year (Y) this is a synchronous impact factor.

An example: the MIF(2003) for the journal *Scientometrics*.

The total number of citations received by the journal *Scientometrics* in the year 2003 was 1012. Hence $TOT = 1012$ and thus $TOT/2 = 506$. Further data can be found in Table 4.

Table 4 Data for the calculation of *Scientometrics'* MIF(2003)

Year	1997	1998	1999	2000	2001	2002	2003
# publications	75	84	128	82	91	84	83
Cumulative number of publications (going back to the past)	627	552	468	340	258	167	83
Number of citations in the year 2003	64	78	68	85	123	96	22
Cumulative % of citations (going back to the past)	52.96	46.64	38.93	32.21	23.81	11.66	2.17

The median is attained somewhere between the 6th and the 7th year. The cumulative number of citations received by articles published during the first six most recent years is 472. As $(506-472) = 34$, and 34 is 0.53 of 64, the median citation age is 6.53 (rounded in the JCR to 6.5). The number of articles published during the most recent six years in *Scientometrics* is 552 (see row three of Table 4). Finally $0.53 \times 75 = 39.75$ is added to this number, yielding 591.75. This is the denominator for the calculation of the MIF. It is concluded that the 2003 MIF of *Scientometrics* is $506/591.75 = 0.855$, a value which is smaller than the corresponding ISI impact factor (1.251).

We may say that recently not only the standard ISI impact factor has received attention, but instead a whole battery of impact factors is used for science evaluation purposes.

4. Mathematical properties of the diachronous impact factor

Recall that the diachronous impact factor is defined as:

$$DIF_n^s(Y) = \frac{\sum_{k=s}^{s+n-1} CIT(Y, Y+k)}{PUB(Y)}$$

Clearly, when at least one the citation data increases, and all other data stay the same, then the DIF increases. The same happens when the number of

publications decreases. If all citation data are multiplied by the same factor, say a , then the DIF also increases by this same factor: $DIF \rightarrow a \cdot DIF$. Similarly, when the number of publications increases by a factor b then the DIF becomes $(1/b) \cdot DIF$. If $a = b$ then the DIF stays invariant. Note though that thinking in percentages is more difficult. If each citation data increases by 10% then $a = 1.1$ and DIF also increase by 10%. If however PUB decreases by 10% then $b = 1.1$ and $(1/b) = 0.91$. Hence an increase of PUB by 10% leads to a decrease of the DIF by 9% (and not 10 %!).

What happens if citations come from different pools of citing articles? We first assume that the total pool of citing articles is subdivided into two disjoint pools. We denote by CIT_j ($j = 1, 2$) the number of citations given by journals belonging to pool j ($j=1, 2$), and use a similar notation for the DIF. Hence DIF_j ($j=1, 2$) denotes the diachronous impact factor calculated exclusively with respect to pool j . Then we have the following simple decomposition:

$$DIF_n^s(Y) = \frac{\sum_{k=s}^{s+n-1} CIT_1(Y, Y+k) + \sum_{k=s}^{s+n-1} CIT_2(Y, Y+k)}{PUB(Y)} = DIF_{n,1}^s(Y) + DIF_{n,2}^s(Y)$$

because the pools are disjoint.

Clearly, if the global pool is subdivided into m disjoint subpools then the decomposition takes the following form:

$$DIF_n^s(Y) = \sum_{j=1}^m DIF_{n,j}^s(Y).$$

What happens if we have m possibly overlapping pools, as in the case of JCR's journal categories?

A citing journal may belong to 1, 2, ..., m pools, making the situation difficult to oversee. For two overlapping pools the answer is relatively easy:

$$DIF(Y) = DIF_1(Y) + DIF_2(Y) - DIF_{1,2}(Y)$$

where, to simplify the notation, we have omitted the length of the citation window (n) and the offset period (s). The index $1, 2$ refers to the journals in the intersection of pool 1 and pool 2.

In the general case we need the so-called inclusion-exclusion theorem. For a proof of this theorem the reader is referred to (Feller, 1968, p. 99). The inclusion-exclusion theorem then immediately leads to the following general pool decomposition theorem.

The general pool decomposition theorem for m possibly overlapping pools of citing documents

$$DIF(Y) = \sum_{i=1}^m DIF_i(Y) - \sum_{i_1, i_2, i_1 \neq i_2} DIF_{i_1, i_2}(Y) + \sum_{\substack{i_1, i_2, i_3 \\ \text{no 2 equal}}} DIF_{i_1, i_2, i_3}(Y) - \dots (-1)^{m-1} DIF_{i_1, \dots, i_m}(Y)$$

Clearly, similar mathematical results can be obtained for other types of impact factors.

5. Conclusion

We have given an overview of a number of impact factors as used in the scientometric literature, including the recently introduced median impact factor. A study of journal issue impact factors combined with a comparison with journal article impact factors is suggested. Some mathematical properties of impact factors are highlighted. Better instruments, i.e. better artefacts and mental schemes, make for better evaluations. Consequently, much more should be done in order to make publication and citation databases optimal instruments for research evaluation.

References

Drijvers, P. and Gravemeijer, K. (2004). Computer algebra as an instrument: examples of algebraic schemes. In: *The didactical challenge of symbolic calculators: turning a computational device into a mathematical instrument*, (D. Guin, K. Ruthven & L. Trouche, eds.). Kluwer: Dordrecht, pp. 171-206.

Feller, W. (1968). *An introduction to probability theory and its applications. Vol. 1 (third ed.)*. Wiley: New York.

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

Ingwersen, P., Larsen, B., Rousseau, R. and Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin*, 46(6), 524-528. Also (in Chinese): *Kexue Tongbao*, 2001, 46(8), 700-704.

Jiang, G., Rousseau, R., and Wu, Y. (2003). *Proceedings of the 9th International Conference on Scientometrics and Informetrics*. Dalian University of Technology Press, Dalian (P.R. China).

Jin, B. and Rousseau, R. (2004). Evaluation of research performance and scientometric indicators in China. In: *The Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies on S&T Systems* (Henk F. Moed, Wolfgang Glänzel and Ulrich Schmoch, eds.). Kluwer Academic Publishers, p. 497-514.

Jin, B. and Wang, B. (1999). Chinese Science Citation Database: its construction and application. *Scientometrics*, 45, 325-332.

Liang, L. (1992). A bibliometric study of the journal *Studies in Dialectics of Nature*, *Studies in Dialectics of Nature*, 8(8), 18-29 (in Chinese).

Rousseau, R. (1988). Citation distribution of pure mathematics journals. In: *Informetrics 87/88* (Egghe L. & Rousseau R., eds). Elsevier, Amsterdam; 249-262.

Rousseau, R. (1997). The proceedings of the first and second international conferences on bibliometrics, scientometrics and informetrics: a data analysis. In: *Proceedings of the Sixth Conference of the International Society for Scientometrics and Informetrics*, (Peritz B. & Egghe L., eds.). Hebrew University of Jerusalem, Jerusalem, 371-380.

Rousseau, R. (2005). The median and percentile impact factors: a set of new indicators. *Scientometrics* (to appear).

Russell, J., and Rousseau, R. (2002). Bibliometrics and institutional evaluation. In: *Encyclopedia of Life Support Systems (EOLSS). Part 19.3 Science and Technology Policy*, edited by R. Arvantis, 2002. Developed under the auspices of the UNESCO, EOLLS Publishers, Oxford (UK).

Sombatsompop, N., Markpin, T., Premkamolnetr, N. (2004). A modified method for calculating the impact factors of journals in ISI Journal Citation Reports: polymer science category in 1997-2001. *Scientometrics*, 60, 217-235.

Vygotski. L.S. (1978). *Mind in society: the development of higher psychological processes*. Harvard University press: Cambridge (MA).

Wu Y. , Pan, Y., Zhang, Y., Ma, Z., Pang, J., Guo, H., Xu, B. and Yang, Z. (2004). China Scientific and Technical papers and Citations (CSTCP): history, impact and outlook. *Scientometrics*, 60, 385-397.