Chinese Science Bulletin, 46(6), March 2001, 524-528.
(Also in Chinese in *Chinese Science Bulletin*, 46(8), 2001, 700-704)

The publication-citation matrix and its derived quantities

Ingwersen Peter[1], Larsen Birger[1], Rousseau Ronald[2] and Russell Jane[3] *

[1] Royal School of Library and Information Science,
Department of Information Studies & Center for Informetric Studies,
Birketinget 6, DK 2300 Copenhagen S, Denmark

[2] KHBO, Dep. of Industrial Sciences and Technology, Zeedijk 101,
B-8400, Oostende, Belgium;
e-mail: ronald.rousseau@khbo.be
&
UIA, IBW, Universiteitsplein 1, B-2610 Wilrijk, Belgium

[3] Centro Universitario de Investigaciones Bibliotecológicas, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico

Abstract

In this short note we give an overview of the main data of a publication-citation matrix. We show how impact factors are defined, and, in particular, point out the difference between the synchronous and the diachronous impact factor. The advantages and disadvantages of using both as tools in research evaluation are discussed.

* Authors are in alphabetical order. No priority is implied.

Introduction

As early as 1960 Raisig [1] suggested the use of a journal impact factor. He called it the ‘index of realised research potential’. Nowadays different ‘impact factors’ are in use. First, we note that citations, and hence impact, are always calculated with respect to a certain pool of journals, e.g. all journals covered by ISI [2], or all journals included in the Chinese Science Citation Database (CSCD) [3]. We always assume that the journal under investigation belongs to a specific pool of journals.

Investigations related to journal citations and impact received a considerable impetus with the annual publication (since 1976) of the Journal Citation Reports (JCR) by the Institute of Scientific Information (Philadelphia, USA). Generally speaking, the JCR is a statistical data set providing information on how often journals are cited, how many items were published, and how often, on average, each item is cited. It also reports those source journals responsible for the references of each journal, the number of references each journal has published and the distribution of these references in time.

In this short note we give an overview of the main data of a publication-citation matrix. We show how impact factors are defined, and, in particular, point out the differences, in definition and use, between a synchronous and a diachronous impact factor.

The publication-citation matrix

Consider the following Table 1. It gives the annual numbers of published articles and citations for one – hypothetical - journal from 1995 to 2000 (citations come of course from the whole set of journals in the pool).

The first row gives the number of articles published yearly in this particular journal. We assume, for simplicity, that all articles are ‘citable’ (no editorials, obituaries, meeting abstracts etc.). The other rows are citation rows. We see that, e.g. in the year 1998 this journal received 12 citations to articles it published in the year 1997. That same year it received 30 citations to articles it published in the year 1995.

Table 1: A publication-citation matrix for a hypothetical journal

Year	1995	1996	1997	1998	1999	2000	totals
# publications	10	15	20	25	30	35	
# citations received in the year 1995	10						10
# citations received in the year 1996	20	10					30
# citations received in the year 1997	25	15	10				50
# citations received in the year 1998	30	19	12	10			71
# citations received in the year 1999	24	26	18	14	10		92
# citations received in the year 2000	23	28	29	17	13	10	120
Totals	132	98	69	41	23	10	

Impact factors, i.e. mean citedness, can be calculated using either a synchronous or a diachronous approach, and with different time windows for publication and citation data. The ISI or Garfield impact factor [4] for the year 1999 (based on Table 1) is:

$$IF_2(1999) = \frac{14+18}{25+20}$$

It is this impact factor that is published annually in the JCR. IF_2 is a synchronous impact factor involving a single citation year and two publication years. The term ’synchronous’ refers to the fact that citations used for the calculation were all received in the same year. In other words, they are obtained from reference lists published in the same year, 1999 in our example. Similarly, the 3-year synchronous impact factor for the year 1999 is:

$$IF_3(1999) = \frac{14+18+26}{25+20+15}$$

In general, the n-year synchronous impact factor of a journal J in the year Y is defined as [5] :

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

In this formula we denote the number of citations received (by a fixed journal J, from all members of the pool) in the year Y, by articles published in the year X by $CIT_J(Y,X)$. Similarly, $PUB_J(Z)$ denotes the number of articles published by this same journal in the year Z. We usually omit the index J. Citation data for a synchronous impact factor will always be found in the same row of the publication-citation matrix. Indeed the data in a certain citation row in our table corresponds to the data that can be obtained from JCR when looking at a journal in the 'Cited Journal' view.

Next, we introduce the diachronous impact factor, denoted (here) by IMP. The 1997 two-year diachronous impact factor for the journal represented in Table 1 is:

$$IMP_2(1997) = \frac{12+18}{20}$$

Or, if you include the year of publication (which is often done):

$$IMP_2^0(1997) = \frac{10+12}{20}$$

Similarly, the 4-year 1995 diachronous impact factor (including the year of publication) is:

$$IMP_4^0(1995) = \frac{10+20+25+30}{10}$$

In general, the n-year diachronous impact factor of a journal for the year Y is:

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

where k = 0 or 1, depending on whether you include the year of publication or not. Citation data for the diachronous impact factor are always found in the same column of the publication-citation matrix. Therefore in order to collect data for calculations of the diachronous impact factors several volumes of JCR are needed. Alternatively, data may be collected using an online methodology (see below). The term 'diachronous' refers to the fact that the data used to calculate it derive from a number of different years with a starting point somewhere in the past looking towards later years. Thus the diachronous impact factor can be said to reflect the development over time.

We note that, in our notation Raisig [1] actually determined $IF_{55}(1926)$ and $IF_{10}(1926)$ for a number of chemical journals, using a pool of one journal, namely the *Journal of the American Chemical Society*.

Why a diachronous impact factor should be used for research evaluations?

Impact factors are mainly applied in two areas: for the measurement of research performance, and as a tool for journal selection and journal de-selection in libraries and other information services. The emphasis in this article is on measurement of research performance.

Selection of appropriate citation and publication windows is crucial in these practical applications of impact factors [6]. Synchronous impact factors "mix" different publication years rendering them more robust when the object under study is the journal itself. They better represent the permanent impact of journals whereas diachronous indicators characterise their actual impact. Consequently when the article (or the scientist who wrote it) is being evaluated diachronous impact factors are more commonly employed. Indeed, calculating diachronous indicators over different windows is the preferred method for evaluation studies by CWTS, the Leiden (the Netherlands) based Centre for Science and Technology Studies [7],[8]. Following their lead, Rousseau [9] used IMP_4^0 for the LUC (Limburgs Universitair Centrum) evaluation studies.

We would like to elaborate on the reasons why it is better to use a diachronous impact for evaluation purposes. Naturally, it is of the utmost importance for the persons, groups and institutes under evaluation that this is performed fairly. One issue here is that like should be compared with like. When using the ISI impact factor the pool of articles consists of all those published during the previous two years in a particular journal (for journal impact factors) or in a particular subfield (when determining subfield impact). One of these two years is not the publication year of the target article, and some circumstances may have changed considerably. A case in point is when one year the journal publishes articles presented at a conference and the next it does not, but, instead, has several special theme issues. In contrast, the diachronous impact factor always makes comparisons with articles published in the same year as the target article. When changing citation windows, the difference between the two methods is more marked. The set of articles used for comparison changes when one moves from a 2- to a 3- or 4-year synchronous impact factor. On the other hand, no change occurs with the diachronous one. The same group of articles are compared, only over a different period.

Another point related to the fair treatment of those being evaluated, is the fact that only journal article evaluations can be performed using a synchronous impact factor. Diachronous impact factors, on the other hand, can also be calculated for one-off publications, such as books containing contributions of different authors, or conference proceedings [10]. Scientists often complain that contributions to such publications are not taken into account during research evaluation exercises [9]. In some fields the most important results are made public during conferences. The fields of computer sciences and of information technology are cases in point. If conferences are held on a regular, e.g. yearly, basis then it is possible to extend the pool of source items (often ISI's database in practice) to cover this series of conference proceedings, and use it as a basis for diachronous impact factors. This has been done for the first and second international conference on bibliometrics, scientometrics and informetrics [10]. Note that the number of citations, and hence the impact factors of journals must be recalculated accordingly (incorporating citations in these proceedings).

Some further comments on research evaluation and the use of impact factors in general

Obviously, for a librarian the long-term impact (perhaps 10 year) is of considerable more importance than the short-term (2 year) impact of a journal. Using different generalised impact factors, or different windows, allows a comparison of the long-term versus the short-term journal impact. Garfield [11] performed such an investigation. He found that journals such as *Cell*, *The New England Journal of Medicine*, *Proceedings of the National Academy of Sciences*, *Nature,* and *Science* always had a high impact, whatever the period chosen (two, seven or fifteen years). Other journals moved up or down significantly. Letter journals in particular suffered considerable downward changes in ranking.

In the evaluation studies mentioned above data for impact factor calculations often have been extracted from databases that are constructed especially for research evaluation purposes. At the Centre for Informetric Studies (CIS) in Copenhagen an alternative data collection strategy has been developed in which data is collected directly from the online versions of the ISI (Institute for Scientific Information) databases [12],[13]. This makes it possible to use diachronous impact factors for research evaluation purposes without having to construct a database especially for this purpose. This method is also accessible to a broader group of people, e.g. research librarians and other information professionals working in a range of institutions. All that is needed is online access to a database host that provides access to the citation databases produced by ISI, e.g. Dialog as used by CIS in an evaluation of nine Danish research centres carrying out environmental research. In this study a range of different indicators were calculated including baseline values that make it possible to compare the performance of the nine centres with the domain impact for the research carried out in Denmark and in the world. These domain impact baselines can be tailored to match the research profile of each individual centre. A combination of data collection methods was used: citations to the individual articles and data for calculation of impact factors were collected online, whereas data from National Science Indicators (NSI) were used to establish the baselines for domain impact. As all indicators are calculated using a diachronous methodology, they are directly comparable. (See [14] for an account of the methodological aspects of the study).

Although the methodology for calculating diachronous impact factors has been known for some time and applied in a number of evaluation studies as indicated above, diachronous impact factors have not yet gained widespread recognition and use outside the bibliometric community. The main reason for this is probably that impact factors from JCR are easy and cheap to get hold of. Thus it is more common that impact factors from JCR are used in measurements of research performance. Yet, ill-considered use of these may lead to wrong conclusions.

One typically assigns impact factors of journals, as obtained from the JCR, to each article published by the scientists (or scientist) under evaluation. These impact factors are then cumulated and maybe averaged. The resulting number is seen as a measure of the total or average impact. In this simple way the synchronous impact factors from JCR are used directly as a measure of research performance. The demand for resources are quite low compared to actually collecting citations to each individual article, because the same impact factor from JCR will be assigned to all articles from the same journal. Although it is a cheap and easy way to evaluate research there are a number of problems associated with this method that affects the validity of the results obtained.

Journal impact is not equal to article impact

Firstly, it is important to note that the impact factor of a journal is a measure based on the averages number of citations of all the articles in the journal, and that the impact of an individual article may be considerably different from the overall impact factor of the journal. Indeed it has been shown that in a typical journal a few articles receive most of the citations, and consequently contribute heavily to the impact factor [15]. Differences between the few articles receiving many citations and the large numbers of normal articles that receive few or no citations will be masked [16]. Assigning JCR's journal impact factor to each article could well result in an unfair assessment of research impact.

Dependence on the citation window

Secondly, the type of journal affects the size of impact factors. As noted above letter journals benefit from the short citation windows used to calculate impact factors in JCR, mainly because letters receive citations only for a short period after publication. Review journals typically have high impact factors as well, but as reviews are often cited for longer periods after publication their impact factors are affected less by short citation windows.

Field dependence

A third problem is that it is very difficult to compare impact factors directly between different research fields without some form of normalisation because the size and nature of different research fields affect the numerical values of impact factors. The top journals in large fields, with a high number of journals and many published articles, often have large impact factors because the pool of citers is large. The impact factors of journals in fields that traditionally have many references per article are higher on average, and fields with rapid publication patterns benefit from short citation windows (like the ones on JCR) in calculations of impact factors. Based on earlier work Seglen [15] points out that certain fields like clinical medicine cite basic medicine, but not vice versa, resulting in higher impact factors for basic medicine on average. Thus even comparisons between impact factors from closely related research fields become unreliable.

These issues need to be taken into account when applying impact factors (synchronous as well as diachronous) for research evaluation and other purposes. Generally, it is not a good idea to use aggregations of two-year synchronous impact factors from JCR to compare research performance between different research groups because of the problems mentioned above. Instead of using impact factors directly as a measure of research performance, they should be used as a measure of the *expected impact* of a group of articles. In this way the differences between the research profiles of the individual groups are normalised and the performance of the groups can be compared directly in a fair manner. One problem remains with this journal-based diachronous approach: a group that produces low quality articles and publishes in low impact journals might do as well as a group that produces high quality articles published in high impact journals [7]. In order to avoid this the impact of a group should be compared to the average impact of the domains in which the group conducts research (a so-called Field Citation Score [8]). This increases the demand for further data collection and processing, but it also results in detailed and accurate results.

As stated above diachronous indicators best characterise the actual impact of articles and journals. Although some problems related to using impact factors from JCR are avoided, the diachronous approach to research evaluation is not without problems, albeit of a less serious nature. First, the diachronous method demands more resources than simply using impact factors from JCR, because it has to be based on manual collection of data, either online or offline. This reduces the total number of articles that can be evaluated. Secondly, because of the time lag in scientific communication, a number of years (ideally more than three) have to pass after publication before an article can be evaluated on the basis of the citations it receives. This means that the latest research published in the last 2-3 years cannot be evaluated using diachronous methods.

Although research evaluation should clearly be based on the diachronous approach to ensure a fair appraisal, complementary synchronous methods could be used to estimate the expected impact of articles published in the last 2-3 years. The field effects may be eliminated to a certain extent by normalising the impact factor from JCR of the used journal set with the average impact factor of all the journals in the field. A number of methods have been suggested for this [17],[18]. Even though the result will only be a rough estimate it will still be valuable in combination with the diachronous approach. It can be seen as a measure of the ability of the research groups under consideration to get articles accepted in high impact journals. Finally, we refer the reader to [19] for a review on research evaluation methods.

What more can be derived from Table 1?

If all data in one row are divided by the total of that row, we obtain a journal's synchronous citation distribution for that particular year. Of course, as we present no data older than 1995 we have here a truncated distribution. Similarly, if all citation data in one column are divided by the total of the column we obtain the diachronous citation distribution for that year. Diachronous citation distributions are always truncated but grow longer each year. Finally, using relative (i.e. after division by the row totals) citation data situated on the same downward-sloping line, we can calculate the average contribution of the first, second, third, and so on, years to the total (synchronous) citation distribution, leading to an 'average' synchronous citation distribution for that journal.

Conclusion

The differences in definition between a synchronous and a diachronous impact factor have been demonstrated based on a publication-citation matrix and formulas for both are given. It is shown that the diachronous impact factor compares like with like, and hence is the preferred one in evaluation studies. The diachronous impact factor can, moreover, be applied on a larger group of publications, including congress proceedings. Further, problems associated with the use of impact factors, synchronous as well as diachronous, for research evaluation purposes are discussed. It is recommended that diachronous impact factors be used as a measure of the expected impact of the articles under consideration, i.e. by comparing with the Field Citation Score.

Acknowledgement. The authors thank the journal's anonymous referees for their helpful comments.

References

[1] Raisig L.M. Mathematical evaluation of the scientific serial. *Science*, 1960, 131:1417-1419.

[2] Garfield E. *Citation indexing - its theory and application in science, technology and humanities.* New York: Wiley, 1979.

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

[4] Garfield E., Sher I.H. New factors in the evaluation of scientific literature through citation indexing. *American Documentation,* 1963, 14:195-201.

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

[6] Rousseau R. Bibliometric and econometric indicators for the evaluation of scientific institutes (in Chinese). In: *R&D Evaluation and Indicators*, (Jiang Guo-Hua, ed.), Red Flag Publishing House, 2000: 16-37.

[7] Moed H.F., Burger W.J.M., Frankfort J.G. & Van Raan A.F.J. The use of bibliometric data for the measurement of university research performance. *Research Policy*, 1985, 14: 131-149.

[8] De Bruin R.E., Kint A. , Luwel M. & Moed H.F. A study of research evaluation and planning:the University of Ghent. *Research Evaluation* 1993, 3: 25-41.

[9] Rousseau R. *A scientometric study of the scientific publications of LUC. Period 1984-1996*. LUC report, 1998.

[10] Rousseau R. 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.) Jerusalem, 1997: 371-380.

[11] Garfield E. Long-term vs. short-term journal impact: does it matter? *The Scientist,* 1998, 12(3): 10-12.

[12] Hjortgaard Christensen F., Ingwersen P. Online citation analysis: a methodological approach. *Scientometrics*, 1996, 40: 529-540.

[13] Hjortgaard Christensen F., Ingwersen P. & Wormell I. Online determination of the journal impact factor and its international properties. *Scientometrics*, 1997, 37: 39-62

[14] Ingwersen P. , Larsen B. &. Wormell I. Applying diachronic citation analysis to ongoing research program evaluations In: *The web of knowledge: A Festschrift in honor of Eugene Garfield* (Cronin, B. & Atkins, H.B., eds.) Medford, N.J.: Information Today, Inc. & American Society for Information Science, 2000, 373-387.

[15] Seglen P.O. The skewness of science. *Journal of the American Society for Information Science*, 1992, 43: 628-638.

[16] Seglen P.O. Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science*, 1994, 45: 1-11.

[17] Braun T., Glänzel W. United Germany: The new scientific Superpower? *Scientometrics*, 1990, 19: 513-521.

[18] Ramirez A.M., Garcia E.O. & Del Rio A. Renormalized impact factor. *Scientometrics*, 2000, 47: 3-9.

[19] Russell J. , Rousseau,R. Bibliometrics and institutional evaluation. In: *Encyclopedia of Life Support Systems (EOLSS)* Part 1.30 Science and Technology Policy, edited by Rigas Arvantis (to appear).