Impact Factor

B K Sen
80 Shivalik Apartments, Alaknanda, New Delhi – 110019
Email: bksen@airtelmail.in

The paper highlights role of impact factor as Garfield’s major contribution in the field of bibliometrics. Defies impact factor and discusses the criteria fixed for calculating impact factor. Informs the manner in which impact factor can be represented mathematically. Highlights features, uses and role of impact factor in evaluation of research output of CSIR scientists. Lastly, informs about Normalized Impact Factor, a variant of the impact factor.

Introduction

Dr Eugene Garfield has conceived a number of bibliometric concepts of which impact factor, I feel, has made the greatest impact. How the idea of impact factor dawned on his mind is not known to me. In all probability, the following circumstances led to the birth of the idea.

The job of compilation of Science Citation Index (SCI) started in early 1960s. For this particular publication, Garfield had to subscribe to more than 600 periodicals from all over the world. The problem was the criteria of selection. For this purpose he exploited his experience of the past years during which he brought out current awareness periodicals. The criteria that was fixed were as follows:

(i) The serial should be a research or review periodical.
(ii) It should be peer reviewed.
(iii) It should be international in scope.
(iv) It should be very regular as to its publication schedule.

The criteria he chose worked more or less satisfactorily. As SCI went on coming Garfield acquired a weapon in the form of citation to evaluate a periodical more scientifically and objectively.

Even before the advent of impact factor, journals have been ranked on the basis of citations, productivity, opinion survey and so on. Ranking by all these methods were not free from flaws. Take, for example, citations. Suppose, the journals J_a, J_b, J_c, J_d, and J_e pertain to the same subject and are considered to be the standard journals on the subject by the experts. The data pertaining to their various parameters are as in Table 1.

From Table 1, it is clear that though the journal J_a received least number of citations per paper per year, still it is having the first rank because of largest number of citations it received. On the other hand, the journal J_e received three times more citations per paper per year, still it has occupied the lowest position. This is the basic flaw of ranking by citation count. Impact factor was conceived to overcome this particular flaw.

Impact Factor

Definition – Garfield defined impact factor as the ratio of number of citations received by source items in a particular year to the number of source items published over a fixed period of time in a particular periodic publication, say a journal1.

‘Source items’ mentioned in the definition include original articles, editorials, letters, short communications, reports of meetings, correction, notes, review articles, etc2. The impact factors provided in Journal Citation Reports are computed considering the ‘fixed period of time’ as two years. This particular impact factor, popularly known
as JCR impact factor or SCI impact factor, is usually used all over the world.

**Symbol**

While defining the term ‘impact factor’, Garfield did not suggest any symbol for the term, as a result different authors have used different symbols. Alvarej and Pulgarin, Osareh and Wilson, Pestana et al, Royle and Davis, Suriya and Manimekalai, Garg and Dutt, have all used the symbol IF. Nisonger, Marshakova-Shaikevich, and Bhattacharyya have respectively used the symbol I.F., Ip, and IMPH respectively. Most authors till date use the IF. IF, I.F. and IMPH are basically abbreviations and not symbols. Ip is a symbol, but it lacks the mnemonic value. Considering these Sen proposed the symbol $I_f$.

**Theoretical Foundation**

Not much work has come to my notice on the theoretical foundation of impact factor. The present author did some work on this some years ago. A glimpse of the same with a little more addition is being presented here.

The impact factor is based on certain axioms as enumerated below.

(i) Most source items will be cited.

(ii) The citation generation potential (CGP) will vary from source item to source item.

(iii) The citations received by a source item will vary from year to year with some exceptions. A source item may receive 10 citations in the first year. In the second year it may receive 20 or many more citations. With the passage of time the CGP of a source item generally decreases. It has also been seen that some source items do not receive any citation at all.

Basing these axioms it may be stated that

(i) The impact factor of a journal ($I_f(J)$) will be dependent on the database providing the impact factor.

(ii) If ($I_f(J)$) will vary from year to year as the source items and their respective CGPs will vary. Every year a journal will have a new set of source items and all these source items will have their own CGPs and that will alter the $I_f(J)$.

(iii) If ($I_f(J)$) is likely to vary if the character of the source items of a journal varies. It has been seen that on average a review article receives more citations. If a purely research journal starts including a few review articles in a year, the $I_f(J)$ is likely to increase. Again, if a journal of mixed content (both review and research articles) is converted into a purely research journal or a review journal, its impact factor is likely to decrease and vice versa.

(iv) The variation in $I_f(J)$ is likely to be significant if the database adds to or deletes from its stock a significant number of journals.

$I_f(J)$ being provided by a database will never be **absolute**, since the database will cover only those citations received by a source item that figure in the journals covered by the database. Suppose a database covers 10,000 journals pertaining to all fields. This database will be able to capture only those citations that are figuring in these 10,000 journals. The citations appearing in the remaining journals numbering more than 100,000 will not be taken into account. As a result the impact factor a database
will provide will always be a working impact factor and not an absolute impact factor.

**Mathematical Formula**

The impact factor can be represented as

\[
I_f(J) = \frac{C_1 + C_2}{S_1 + S_2}
\]

where \( I_f(J) \) denotes impact factor of the journal J for the year \( Y \);

\( C_1 \) denotes the number of citations received by \( S_1 \) source items in the year \( Y \);

\( C_2 \) denotes the number of citations received by \( S_2 \) source items in the year \( Y \);

\( S_1 \) denotes the number of source items published in the journal J in the year \( Y-1 \); and

\( S_2 \) denotes the number of source items published in the journal J in the year \( Y-2 \).

Suppose the journal J has published 32 and 36 source items in the years 1907 and 1908 respectively. These source items have received respectively 40 and 28 citations in 1909.

Now, the impact factor of the journal J will be \((40+28)/(32+36) = 1\).

It may be noted that the JCR impact factor of a journal cannot be determined till the journal has completed two calendar years of its life and all its issues have been published in time. For a journal that has started publishing in the year 2010, its impact factor will be known only in 2012 provided that the journal has been very regular in publishing its issues.

**Features**

The features of impact factor can be enumerated as below:

(i) It is a pure number and does not have any unit.

(ii) It is not a constant like the specific gravity of water.

(iii) The number is expressed up to three digits after decimal in JCR, e.g. 2.319

(iv) It is year-specific. Because of various factors it generally varies from year to year.

(v) It is database-specific. The impact factor of the journal J given by the database Da is most likely to be different from the impact factor given by the database Db.

(vi) The value of impact factor generally lies between 0 and 50. It may go beyond 50 due to various reasons.

(vii) By and large review periodicals tend to have greater impact factors than research periodicals.

(viii) It also varies from subject to subject.

(ix) The impact factor indicates the standing of the journal in the world.

(x) The impact factor may be considered as an indication of the quality of the journal in most cases.

**Uses**

With the passage of time impact factor is finding more and more uses. Some of the uses are being discussed here.

**Selection of Journals** – While selecting journals for acquisition in a library, librarians tend to prefer selection of journals on the basis of impact factors as impact factors clearly show the standing of the journals in the world. Nowadays authorities are in general aware of impact factors. As such, no eyebrow is raised in case journals are selected on the basis of impact factors.

**Discontinuation of Journals** - Many a time libraries are to discontinue some journals because of budget constraints and various other factors. In such a situation the problem the librarians face is how to decide which journals are to be discontinued out of so many. At the National Centre of Bibliometrics we used to receive requests from librarians to help them in such matters. To solve their problems, invariably we used to arrange the journals according to their impact factors. The journals figuring at the bottom of the ranked list usually were chosen for deselection.

**Placing a Paper** – Every researcher after completing the paper wants a reputed journal to place his/her paper so that the paper comes to the notice of the researchers
of his field all over the world. In such cases the list of journals arranged subject-wise according to impact factor in *Journal Citation Reports* proves to be of immense help.

**Evaluation of Recent Research Output**

In late 1980s, Dr A P Mitra, F. R. S. Director General of Council of Scientific and Industrial Research (CSIR) desired that INSDOC should evolve an objective method for the evaluation of recent research output of CSIR. That means the standard of each paper published by CSIR scientists last year is to be determined objectively.

Citation analysis of the papers was ruled out because it requires a minimum time for the papers to get cited starting from six months to years. The standard of a paper can be more or less optimally judged on the basis of the citation scenario that evolves in about five years time. This is what we felt after viewing the citation scenario of thousands and thousands of papers. The director general of a group of laboratories or the director of a laboratory cannot wait that long since in many cases they are appointed for a term of five years only. Just after taking over they like to know the standing of the laboratory in terms of its publications, usually the standard of the research papers. If they come to know quickly that the standard is not up to the mark, they can ensure corrective measures before long.

As we were to determine the standard of the papers published last year, some other method was necessary. We deliberated on this and the opinions of the experts like Subbiah Arunachalam and P S Nagpaul were also sought. We finally came to the conclusion that impact factors of the journals were the only way out. The method of evaluation using impact factor was developed [Sen, Shail] and the performance of the laboratories were beautifully portrayed. The work was appreciated not only by Dr A P Mitra but also his successors.

Till now, to my knowledge, there is no other method whereby evaluation of recent research output can be done so objectively, effectively and quickly. Peer review is another method whereby evaluation can be done. The method is time consuming and not always 100 per cent objective.

**Awarding Scientists**

Some laboratories have the tradition of awarding scientists annually for their research contributions. The scientist contributing the best paper is usually awarded. Determination of the best paper is usually done through peer review which is not always objective. I heard that a laboratory has gone for the impact factor. The paper which was published in the highest impact journal was awarded.

**Varieties**

Over the years varieties of impact factors have evolved. Only some of them are being discussed here.

**Normalised Impact Factor**

The demand of the circumstances led to the evolution of this particular variety of impact factor. In late 1980s when we were working on the bibliometric evaluation of CSIR research output, the average impact factor of some of the laboratories like National Aeronautical Laboratory (NAL) was found to be quite low compared to National Chemical Laboratory (NCL). For example, in 1989 research output, the average impact factor of NAL was 0.33 and that of NCL was 1.26. In fact, the standard of research of NAL was quite high. The average impact factor was not showing the correct picture. In reality, low impact factor of journals pertaining to aerospace engineering and technology (AET) was responsible for this. In 1989, the highest impact factor of research journals in AET was only 0.454 (*IEEE T Aero Elec Sys*), and that of chemistry was 5.049 (*Angewandte Chemie – International Edition*). This particular situation drove us to normalise the impact factors.

The methodology of normalization is being described here very briefly as the same has been described earlier. The highest impact factor in any individual subject category of JCR was considered as 10. As a result the normalised impact factor of *IEEE T Aero Elec Sys* became 10 multiplying 0.454 with 22.026, and that of *Angewandte Chemie – International Edition* also became 10 multiplying 5.049 with 1.981. The impact factors of the journals of aerospace engineering and technology were multiplied with 22.026 to get normalised impact factors of all the journals belonging to this subject category. Similarly, impact factors of journals belonging to other subject categories were determined. When the average normalised impact factors were determined for NAL and NCL, they were found to be 3.16 and 3.85 respectively which were much more realistic.

The number with which the highest impact factor of a subject category was multiplied was termed as multiplier.
The symbol for normalised impact factor of a journal was suggested as $\text{nIf}(J)$.  

The normalised impact factor of a journal can be computed using the following formula.

$$
\text{nIf}(J) = \frac{I_{(J)} \times 10}{h_{If}}
$$

Where $\text{nIf}(J)$ stands for the normalised impact factor of the journal $J$ for the year $Y$,

$I_{(J)}$ is the impact factor of the journal $J$ for the year $Y$,

$h_{If}$ is the highest impact factor of the subject category in the year $Y$ to which the journal $J$ belongs.

**Features**

(i) Normalised impact factor is also a number, and does not have any unit.

(ii) It is also year specific and database specific.

(iii) The value of normalised impact factor lies between 0 and 10.

(iv) Regardless of subject category, the normalised impact factor of a journal provides an indication of its standard. If the normalised impact factor is close to 10, the periodical definitely is of high standard; if it is around 0, it is of poor standard.

**References**