A new journal citation impact measure that compensates for disparities in citation potential among research areas

Henk F. Moed
Elsevier Radarweg 29, 1043 NX Amsterdam, The Netherlands
Email: h.moed@elsevier.com

This paper presents a new journal indicator of citation impact of a scientific-scholarly journal. It builds further upon Eugene Garfield’s groundbreaking ideas presented in many of his early and later publications, by combining his concept of a journal impact factor with his notion that “evaluation studies using citation data must be very sensitive to all divisions, both subtle and gross, between areas of research; and when they are found, the study must properly compensate for disparities in citation potential”. The proposed indicator is based on a tailor made delimitation of a journal’s subject field, and takes into account the frequency and immediacy of citation and database coverage in a subject field.

Introduction

Eugene Garfield’s original and illuminative studies of the scientific communication system have generated an enormous impact not only upon library and information scientists, but also upon producers of scientific literature databases, journal publishers and editors, librarians, research managers and research performance assessors. His studies were based on the Science Citation Index (SCI), one of his key information products published by the Institute for Scientific Information, a company he founded in 1960. A good illustration of the use of the SCI and related indexes for studying the structure of the scholarly communication system is the following paragraph in which he relates citation counts in the SCI to a concept of ‘quality’ of a scientific journal.

“Since authors refer to previous material to support, illustrate, or elaborate on a particular point, the act of citing is an expression of the importance of the material. The total number of such expressions is about the most objective measure there is of the material’s importance to current research. The number of times all the material in a given journal has been cited is an equally objective and enlightening measure of the quality of a journal as a medium for communicating research results”.

Relationships among journals were analyzed in terms of citations from one journal to another, and core journals and more peripheral ones were identified. Equally important, his analyses provided the basis for a unique ‘internal’ monitor of the adequacy of coverage of the SCI itself.

The journal statistics Garfield explored were soon isolated from the study context and published by ISI in rankings of journals by impact factor, probably the bibliometric construct most widely used in the scholarly and publishing community. Nowadays journal publishers and editors use journal metrics to obtain answers to questions as to how a journal compares with its competitors. Librarians use it to determine which journals are likely to be the most useful for the researchers in their institution. Research managers apply journal citation measures when dealing with the question how they should devise an effective publication strategy for the papers written in their group, and many research performance assessors use such measures as tools for assessing the research performance of an individual researcher or research group.

The potentialities of journal citation measures in the contexts outlined above are intensively discussed among practitioners in the field of library and information science and quantitative science studies, but also in the broader scientific community and in the various user groups. Perhaps the following three statements represent a consensus among most if not all participants (e.g., Moed3, Glanzel4).
1. Journal performance is a complex, multidimensional concept that cannot be fully captured in a single metric, page 249.

2. In the construction and interpretation of journal citation measures. It is crucial to take into account differences in communication and citation practices between research fields.

3. Although journal quality is an aspect of research performance in its own right, journal impact measures should not be used as surrogates of actual citation impact of a group’s publications.

The extensive use of the journal impact factor (JIF) for purposes for which it was not designed, raised a series of criticisms, all aiming to adapt the measure to the new user environments and to propose new types of indicators. But the impact factor is still the norm, it is the metric ‘to beat’. During the past years, numerous other approaches to the measurement and ranking of journal impact were explored. Without claiming completeness, important approaches are as follows.

1. In order to account for differences in citation practices between research fields – both the frequency and the immediacy of citation –, Pudovkin and Garfield developed a ranking procedure of scientific journals similar to percentile ranking, that generate rank-normalised impact factors of scientific journals based on citation analysis.

2. One of the limitations of traditional citation analysis is that all citations are considered ‘equal’. Following Pinski and Narin, several authors applied algorithms based on Google PageRank to the journal-to-journal citation network whereby the subject field, quality and reputation of a journal have a direct effect on the value of the citations it gives to other journals.

3. The Hirsch Index, originally developed as a tool to evaluate individual scientists, has become a popular bibliometric indicator ever since its launch in 2005. Braun, Glanzel and Schubert and others explored the use of Hirsch Indices calculated for scientific journals.

4. Citation distributions of scientific journals are known to be skewed. Stringer, Sales-Pardo and Amaral developed a model for the asymptotic number of citations collected by papers published in a journal and obtained evidence that the corresponding citation distribution for articles cited at least once is approximately normal, enabling one to quantify both the typical impact and the range of impacts of papers published in a journal.

5. In view of the complex nature of journal citation distributions, the development and testing of models describing such distributions – which has a long tradition in library and information science – is ongoing. Glanzel modelled a journal citation distribution as a negative binomial distribution, and characterized a journal’s impact by estimating the two parameters of this distribution.

6. The approaches indicated above are all based on citation counts. But more and more studies explore the use of data on downloads of papers in full text format from electronic publication archives. Journal ‘usage’ factors are constructed and calculated, and their correlation with citation-based measures is examined.

This paper presents a new journal indicator that is a member of the family of journal citation indicators that use the concept of the ISI journal impact factor as a starting point, and aim at taking into account important factors that do affect a journal’s citation rate, but that have little to do with its performance. It builds further upon Eugene Garfield’s ideas presented in many of his early and later publications. The new indicator was developed at the Centre for Science and Technology Studies (CWTS) at Leiden University.

The structure of this paper is as follows. The details of the SNIP indicator are presented in Section 2. This section also explains the concept of source normalization, and shows how it can be used to construct an indicator of journal citation impact. Section 3 highlights general
MOED: A NEW JOURNAL CITATION IMPACT MEASURE THAT COMPENSATES FOR DISPARITIES

A journal’s raw impact per paper

Citation potential in its subject field

A field’s frequency & immediacy of citation

Database coverage

Journal scope, focus

Source normalized impact per paper (SNIP)

Source normalization

Many authors have underlined that it is improper to make comparisons between citation counts generated in different research fields, because citation practices can vary significantly from one field to another. For instance, articles in biochemistry often contain over 50 cited references, while a typical mathematical paper has perhaps only 10. This difference explains why biochemical papers are cited so much more often than mathematical ones. Eugene Garfield’s view on this is clear. “Evaluation studies using citation data must be very sensitive to all divisions, both subtle and gross, between areas of research; and when they are found, the study must properly compensate for disparities in citation potential”.

One way to overcome differences in citation potential is calculating ‘relative’ indicators that calculate the ratio of a journal’s citation impact per paper and the world citation average in the subject field covered by the journal. This approach can be denoted as target normalization. A second approach is based on the idea of source normalization, labelled by Zitt and Small as “citing-side normalization”. Its base idea is that the actual citation rate of a set of target papers in a subject field is ‘normalized’ or ‘divided’ by a measure of the frequency at which articles in that field cite other documents. The indicator presented in the current paper is based on this idea.

A journal’s source normalized impact per paper, SNIP, is a ratio of two measures: it divides the number of times a journal’s articles are cited – i.e., its raw impact per paper – to the frequency at which papers in the journal’s subject field cite other materials – the subfield’s citation potential. In other words, it measures a journal’s citation rate per cited reference in – or per citation given in – documents in the journal’s subject field. Figure 1 summarizes the principal characteristics of this indicator:

1. It aims at taking into account only peer reviewed articles published in a journal, both as cited and as citing documents.
2. A journal’s subject field is defined in a dedicated manner, taking into account a journal’s scope.
3. It accounts for differences in the frequency and immediacy of citation between research fields, by applying the idea of source normalization.
4. It corrects for differences in the extent to which the publication database used in the calculations covers a subject field.

Raw impact per paper

This measure is similar to the ISI (currently Thomson Reuters) journal impact factor. It is defined as the average citation rate in a particular year (the citing or impact factor year) of papers published in a journal during the three preceding years. By considering three cited preceding years rather than two – as is the case in the ISI impact factor –, results for journals in fields showing a slowly maturing impact that does not peak after one or two years can be expected to be more reliable. ‘Papers’ include articles, reviews and proceedings papers. Informal, non-peer reviewed communications such as editorials and letters to the editor are discarded both as cited and as citing sources.

A journal’s subject field

A journal’s subject field is defined as the collection of papers citing the journal. In this way a subfield is defined
by the (formal) users of a journal, whose behaviour can be expected to properly reflect the journal’s scope. One technical aspect should be clarified here. As outlined below, the time frame of three preceding cited years plays a key role not only in the calculation of the raw impact per paper, but also in that of a journal’s citation potential. But in the SNIP methodology, the articles constituting a journal’s subject field do not necessarily have to cite 1-3 year old papers published in the journal, because this would introduce a bias in favour of articles citing recent materials above older documents. Therefore, a ten-year time window is applied: a journal’s subject field consists of the articles citing at least one 1-10 year old paper published in the journal.

Citation potential

Citation potential in a subject field captures how frequently authors in that field cite other documents in their reference lists. A first operationalization would simply be the average number of cited references in articles covering a field. But the SNIP methodology takes into account three additional factors.

1. It counts only cited references that are one to three years old. In this way, raw impact per paper and citation potential relate to the same time window. This is appropriate, because the probability for a 1-3 year old article in a journal to be cited is proportional to the average number of 1-3 year old cited references contained in papers in the journal’s subject field. If one would use the total number of cited references as the probability factor, journals in fields in which articles tend to have long reference lists and citations a low immediacy, such as taxonomy, would be disadvantaged.

2. SNIP’s measure of citation potential only takes into account cited references published in sources that are indexed for the database (i.e., Scopus). For instance, citations to books are not included. To the extent that journal metrics is used to assess journals indexed for a database, it is appropriate to count only cited references actually published in indexed journals. In this way, one corrects for differences in database coverage across research fields. If not, the citation impact of indexed journals in fields in which database coverage is not as high as it is in (bio-)medical and physical sciences, such as mathematics, engineering, social sciences and humanities, would be systematically undervalued.

3. Citation potential is itself normalized by calculating a relative database potential. In a first step citation potential is calculated for all journals in the database. Next, the median journal in terms of citation potential is identified, and the value of its citation potential is used as a normalization factor, by dividing the citation potential of each journal by that of the median journal. Hence, journals for which the subject field shows a citation potential above that for the median journal in the database have a relative citation potential above one.

General features

Comparing a journal’s SNIP with its raw impact per paper, – which in its turn is to some extent similar to the Thomson impact factor – , the following features can be noted.

1. If a journal covers a subject field in which the citation potential is higher than that for the median journal in the database (in other words, the relative citation potential is above one), its SNIP value is lower than that of its raw impact per paper. For instance, for journals in the field of molecular biology SNIP tends to be lower than the raw impact per paper.

2. On the other hand, for journals in subject fields in which citation potentials are lower than that for the median journal, the SNIP value exceeds that of its raw impact per paper. For example, for journals in the field of mathematics SNIP tends to be higher than the raw impact per paper.

3. SNIP is constructed in such a way that by definition for 50 per cent of journals SNIP is higher than the raw impact per paper, while for another 50 per cent it is lower. In other words, taking the raw impact per paper as the norm, and characterizing SNIP, 50 per cent of journals goes up, and another 50 per cent goes down.
Selected outcomes

Examples of individual journals

Table 1 presents outcomes for selected journals. Data relate to the citing year 2007. The first journal, Journal of Electronic Materials, is the median journal in the database in terms of database potential. Therefore, the relative database potential of the subject field covered by this journal is 1.0. Outcomes of the next two journals, Inventiones Mathematicae and Molecular Cell illustrate the SNIP methodology quite clearly. The raw impact per paper of the latter journal is almost nine times that of the former (13.0 against 1.5), but the SNIP values are statistically similar (4.0 versus 3.8), as the citation potential of the molecular biological journal is 8 times that of the mathematical periodical (3.2 versus 0.4).

The second case relates to the sections of the Journal of Gerontology. The raw impact per paper of the section on biological and medical sciences is higher than that of the psychology and social sciences section, but for the SNIP values it is the other way around. Finally, the case of two journals in the field of atomic and molecular physics shows that the more topical journal Nanoparticles Research has a raw impact per paper that is about twice that of the more general Journal of Molecular Spectroscopy, but if one corrects for differences in citation potential between the subject fields

Table 1 — SNIP values for specific journals

<table>
<thead>
<tr>
<th>Journal</th>
<th>Raw Impact per Paper</th>
<th>Relative Citation Potential</th>
<th>SNIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of Electronic Materials</td>
<td>1.5</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Inventiones Mathematicae</td>
<td>1.5</td>
<td>0.4</td>
<td>3.8</td>
</tr>
<tr>
<td>Molecular Cell</td>
<td>13.0</td>
<td>3.2</td>
<td>4.0</td>
</tr>
<tr>
<td>Journal of Gerontology A – Biological and Medical Sciences</td>
<td>3.7</td>
<td>2.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Journal of Gerontology B – Psychology and Social Sciences</td>
<td>2.7</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Journal of Molecular Spectroscopy</td>
<td>1.1</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Nanoparticle Research</td>
<td>2.3</td>
<td>1.9</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table 2 — Statistics on SNIP for selected subject fields

<table>
<thead>
<tr>
<th>Scopus Journal Subject Category</th>
<th>Nr Journals</th>
<th>Spearman Rank Correlation</th>
<th>SNIP P25</th>
<th>SNIP P50</th>
<th>SNIP P75</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fields with the largest Spearman R between SNIP and raw impact per paper</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatology</td>
<td>44</td>
<td>0.98</td>
<td>0.09</td>
<td>0.50</td>
<td>1.12</td>
</tr>
<tr>
<td>Biochemistry (medical)</td>
<td>42</td>
<td>0.98</td>
<td>0.08</td>
<td>0.45</td>
<td>0.90</td>
</tr>
<tr>
<td>Internal Medicine</td>
<td>102</td>
<td>0.98</td>
<td>0.10</td>
<td>0.36</td>
<td>0.92</td>
</tr>
<tr>
<td>Hepatology</td>
<td>35</td>
<td>0.98</td>
<td>0.09</td>
<td>0.71</td>
<td>1.23</td>
</tr>
<tr>
<td>Endocrinology, Diabetes and Metabolism</td>
<td>163</td>
<td>0.98</td>
<td>0.13</td>
<td>0.65</td>
<td>1.19</td>
</tr>
<tr>
<td><strong>Fields with the lowest Spearman R between SNIP and raw impact per paper</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arts and Humanities (miscellaneous)</td>
<td>43</td>
<td>0.63</td>
<td>0.22</td>
<td>0.39</td>
<td>0.63</td>
</tr>
<tr>
<td>Algebra and Number Theory</td>
<td>51</td>
<td>0.65</td>
<td>0.81</td>
<td>1.03</td>
<td>1.25</td>
</tr>
<tr>
<td>Discrete Mathematics and Combinatorics</td>
<td>36</td>
<td>0.68</td>
<td>0.77</td>
<td>1.19</td>
<td>1.49</td>
</tr>
<tr>
<td>Behavioral Neuroscience</td>
<td>60</td>
<td>0.75</td>
<td>0.99</td>
<td>1.25</td>
<td>1.65</td>
</tr>
<tr>
<td>Archeology (arts and humanities)</td>
<td>89</td>
<td>0.77</td>
<td>0.10</td>
<td>0.25</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Table includes only subject categories containing more than 30 journals. SNIP 25 means: 25 % of journals has a SNIP value below the value indicated (bottom quartile). SNIP P50 (i.e., the median of the distribution): 50 % of journals has a SNIP value below the value indicated. SNIP P75: 25 % of journals has a SNIP value greater than the value indicated (=top quartile)
covered by these two journals, SNIP values are almost identical.

Statistics for selected subject fields

Table 2 shows for selected Scopus journal subject categories the number of journals included, and the Spearman Rank correlation coefficient between SNIP and Raw Impact per Paper. Moreover, it gives the 25th, 50th (i.e. median) and 75th percentile of the distribution of SNIP values across journals in a particular subject category. For instance, a value in the column SNIP P25 means that 25 per cent of journals has a SNIP value below the value indicated (this value marks the bottom quartile of the distribution). Results are presented for the 5 categories with the lowest Spearman rank correlation coefficients between SNIP and raw impact per paper; and for the 5 fields showing the highest rank correlation coefficient. Data relate to the citing year 2008.

It must be underlined that SNIP and the raw impact per paper indicator are strongly dependent, as the first is defined as the ratio of the second indicator and the citation potential in a journal’s subject field. The rank correlation coefficient should be interpreted as a purely descriptive statistic indicating the degree of similarity among journals in a subject category between a ranking based on SNIP and one based on raw impact per paper.

Table 2 shows that the 5 subject fields for which the rank correlation between SNIP and raw impact per paper is largest are all in medical sciences, whereas among the 5 fields with the lowest correlation two are from arts & humanities and two from mathematics. For a complete overview, the reader is referred to an information document at the CWTS website on journal indicators.

Conclusion

The article that introduced the SNIP indicator26 gives a list of what the author believes to be strong points of SNIP, a list of issues that should be taken into account when interpreting SNIP values, and problems that have to be further analyzed. These points are not repeated here in detail. Summarizing, strong points of the SNIP methodology are as follows. The delimitation of a journal’s subject field does not depend upon some pre-defined categorization of journals into subject categories; it allows a sensible calculation of journal impact of multi-disciplinary or general journals; it corrects for differences in citation frequency and immediacy and in database coverage, between journals not only from different disciplines, but also within a discipline; and it aims at taking into account only peer reviewed articles.

Important points to keep in mind are that SNIP values tend to be higher for journals publishing review articles or showing a high journal self citation rate. Moreover, the source normalization applied in SNIP does not take into account the growth of the literature in a field, nor the extent to which papers in a field are cited from other fields. More sophisticated methods to define subject fields using citation analysis could be explored, along with the biases they may cause. And the relationship between rankings of journals based on SNIP and peer judgements on these journals should be further clarified.

As from January 2010, SNIP is included in Scopus, jointly with the SJR Indicator developed by the Scimago Research Group27. SNIP and a series of related indicators are for all Scopus journals and for the past ten years freely available a website created and hosted by CWTS28. Although it must be underlined that there is no such thing as a single ‘perfect’ indicator of journal performance, responses from the user communities and more research on the SNIP methodology will further assess and enhance SNIP’s validity and usefulness. In this way, a next step is made along a trajectory that started with Eugene Garfield’s invention and creation of the Science Citation Index, and the publication of his groundbreaking studies of the system of scientific-scholarly communication.

References

2. Ibid p. 23-24


24. Van Leeuwen op. cit.


27. Gonzales op. cit
