# A further step forward in measuring journals' technological factor

**Vicente P. Guerrero-Bote; Henk F. Moed; Félix De-Moya-Anegón**

How to cite this article:

**Guerrero-Bote, Vicente P.**; **Moed, Henk F.**; **De-Moya-Anegón, Félix** (2021). "A further step forward in measuring journals' technological factor". *Profesional de la información*, v. 30, n. 4, e300406.

*https://doi.org/10.3145/epi.2021.jul.06*

Manuscript received on 24th June 2021
Accepted on 9th July 2021

**Vicente P. Guerrero-Bote** ✉
*https://orcid.org/0000-0003-4821-9768*

*SCImago Research Group*
*Universidad de Extremadura*
Plazuela Ibn Marwam, s/n
06071 Badajoz, Spain
*guerrero@unex.es*

**Henk F. Moed**
*https://orcid.org/0000-0003-2446-905X*

*Sapienza University of Rome*
*Department of Computer, Automatic and Management Engineering*
Rome, Italy
*henk.moed@uniroma1.it*
*hf.moed@gmail.com*

**Félix De-Moya-Anegón**
*https://orcid.org/0000-0002-0255-8628*
*SCImago Research Group*
Madrid, Spain
*felix.moya@scimago.es*

**Abstract**

A new indicator of technological impact of a scientific journal is presented, based on citations in patent applications to papers published in that journal. Key characteristics are that patent citations are aggregated by patent family, and that not all citations are considered equal. To each family citation a weight is assigned that is proportional to the Gross Domestic Product (GDP) of the countries in which protection is requested, to take into account the costs and expected benefits of patenting, and inversely proportional to the number of cited references in a patent family, as a way to correct for differences in citation frequencies in patent applications among technological fields. Around one third of journals indexed in *Scopus* have at least one citation from patent applications in a 5-year citation window. The distribution of the technological impact scores among journals can be modelled as a power law distribution, with the slope being a little smoother than that of common scientific impact indicators SJR and JIF. However, the correlations between technological and scientific impact indicators are mostly low or moderate, which shows that they measure quite different aspects of journal or research performance.

**Keywords**

Technological impact; Indicators; Journal technological factor; Journals; Scholarly journals; Patents; *Patstat*; Citations; Research performance; Bibliometrics; Non-patent references; NPR; Non-patent literature; NPL.

**Funding**

This work has been financed by the *Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016* and the *Fondo Europeo de Desarrollo Regional* (*Feder*) as part of research project CSO2016-75031-R.

## 1. Introduction

The idea that science is a structured knowledge system, and that research is not only dedicated to contributing new knowledge to the body of scientific understanding, but also to the application of this knowledge to solve practical problems, is currently widely accepted both in science and in the wider society.

Nowadays, a large part of R&D investment comes from public funds, which entails the need for transparency and accountability, and, therefore, the need to evaluate R&D performance. Such evaluation is also useful as it helps to improve "the institution of science" (**Merton**, 1957).

Research activity can be conceived of as a cycle in which inputs lead to the creation of new knowledge that generates outputs –scientific articles, communications in congresses, doctoral theses, monographs, patents, products– which, in their turn, can be transformed into inputs of a new cycle (**Salkind**; **Rainwater**, 2003; **Callon**; **Courtial**; **Penan**, 1993). Hence, it is clear that the evaluation of research presupposes the evaluation of its outputs.

However, it is well known that not all published works have the same value. To analyse the quality of the works, Scientometrics had developed the notion of scientific impact, such as the impact generated by the publication of a work upon surrounding research activities in the scientific community. And to calculate it, the scientometric method uses the citation (**Price**, 1963), based on the assumption that, despite differences in citers' motivations (**Bornmann**; **Daniel**, 2008), a citation reflects a recognition of a previous work (**Moed**, 2005). Citations of one scientific paper to another provide links between people, ideas, publications, and institutions, and such links constitute a network that can be analysed quantitatively.

The scientometric method has its origin in the work of Eugene Garfield who created the *Science Citation Index* (*SCI*) as a bibliographic database that included cited references given in source documents (**Garfield**, 1955). It must be noted that his initial purpose was to facilitate and improve information retrieval rather than research evaluation.

To determine the journals to be included in the SCI in an objective way, the Impact Factor (**Garfield**, 1972) was developed which became the most widely used bibliometric tool for the evaluation of scientific journals. Ever since, the impact factor has been criticized for enabling journal editors to give all citations the same weight regardless of their origin, for taking into account only a part of the global scientific literature, and for its lack of comparability across scientific disciplines.

To solve some of these problems, other citation-based journal indicators have emerged, such as the SNIP (*Source Normalized Impact per Paper*), one of the journal indicators derived from *Elsevier*'s *Scopus* (**Moed**, 2010; **Waltman** *et al.*, 2013), and based on a subject field normalization correcting for differences in citation frequencies among disciplines.

Recursive methods have emerged as well, based on the *Google Pagerank* principle, that do not assign the same value to all citations, but a weight that depends on the prestige of the journals in which a citation is made. The *Eigenfactor* (**Bergstrom**, 2007) published by *Clarivate Analytics* in the *Web of Science*, and *SCImago Journal Rank* (**González-Pereira**; **Guerrero-Bote**; **De-Moya-Anegón** *et al.*, 2010; **Guerrero-Bote**; **De-Moya-Anegón**, 2012) derived from *Scopus*, are typical applications of this *Pagerank* principle. It has scientometric roots: it was developed by Francis Narin and co-workers (**Pinski**; **Narin**, 1976) to calculate journal influence weights.

All these indicators use citations from the scientific literature to evaluate scientific journals, so they measure the impact or prestige of those publications within the scientific community. But citation analysis can be expanded by analysing citations in patents to scientific papers, and measuring in this way the influence of scientific-work on technological development (**De-Moya-Anegón** *et al.*, 2020). Thus, Narin and his collaborators began using citations from patents to measure technological impact in the 1980s (**Narin**; **Olivastro**, 1992; **Narin**; **Hamilton**; **Olivastro**, 1997), and there has been more recent work in the same line (**Huang**; **Huang**; **Chen**, 2014; **Liaw** *et al.*, 2014). However, the use of patent citations to calculate indicators is a complex endeavour. The distribution of patent citations is highly skewed, both on the citing side –a large part of the patents contain no references to the scientific literature (**Verbeek** *et al.*, 2002)– and on the cited side –the scientific papers cited in patents are published in a small number of journals (**Callaert** *et al.*, 2006)–. In a detailed review of the scientific literature on the issue, **Van-Raan** (2017) found that citations from patents vary depending on the inventors, the examiners, the technology field, the patent office, and the company.

Of particular importance is the difference due to the patent office, especially if data from more than one patent office are to be used. For example, applicants applying for protection at the *USPTO* should include as many references as possible so as to avoid future problems, while it is unnecessary to include many references in applying for protection at the *EPO*, and most references are added by the examiners. Thus, there will be no differences if the data to be included are from a single patent office (**Huang**; **Huang**; **Chen**, 2014; **Liaw** *et al.*, 2014), but if data from different patent offices are to be included then it is necessary to design a normalization mechanism.

Publications associated with patents are grouped into patent families (**Martínez**, 2011). For example, all applications for protection of an invention in different countries are grouped into one family. As they are evaluated by different examiners and in different national authorities, applications corresponding to the same invention may end up having different cited references.

**Guerrero-Bote**, **Moed**, & **De-Moya-Anegón** (2021) developed a methodological approach to handling such differences in references among patent applications from the same family. It unifies all references in patent applications from the same family, so that each patent application contains the same references. The rationale behind this approach is that patent applications from the same family relate to the same invention, and therefore should have the same cited references.

Conceptually, a patent family can be conceived of as a single patent application in which the reference list contains all unique cited references in the total set of patent applications belonging to the family, and the list of designated countries consists of the all unique designated countries in the total set of family members.

The objective of the present communication is to design an indicator of technological impact of a scholarly journal, based on citations in patent applications to papers published in that journal. This measure is denoted a journal's Technology Factor (TF). A key characteristic is that not all citations in patent applications are considered equal. To each family citation a weight is assigned that is proportional to the Gross Domestic Product (GDP) of the countries in which protection is requested, and inversely proportional to the number of cited references in a patent family.

The rationale for weighting a citation of a patent family with the GDP of designated countries is that applying for a patent in a more advanced country with a larger market is more expensive and difficult than it is in a smaller or developing country. Since these costs and efforts are justified because greater benefits can be expected, the value of a patent application should be related to the countries in which protection has been applied for.

The idea behind weighting a patent family citation with the reciprocal value of the total number of cited references is that in this way one can correct for differences in citation frequencies in patent applications among technological fields. In some fields patent applications contain more references than in others. This weighting approach can be conceived of as a form of "source" (**Moed**, 2010) or "citing side" (**Zitt**; **Small**, 2008) normalization.

## 2. Data

The present study uses a bibliographic database of scientific literature, a patent database, and a linking procedure that allows identification of the publication records in bibliographic database records that correspond to the cited references in the patent database.

As a bibliographic database we have used *Scopus*, created and maintained by *Elsevier*. It includes more than 31 000 scholarly sources (**Hane**, 2004; **Pickering**, 2004). The characteristics are analysed in several studies (**Archambault** *et al.*, 2009; **Leydesdorff**; **De-Moya-Anegón**; **Guerrero-Bote**, 2010; **De-Moya-Anegón** *et al.*, 2007). Numerous scientometric studies have used *Scopus* (e.g., **Gorraiz**; **Gumpenberger**; **Wieland**, 2011; **Jacsó**, 2011; **Guerrero-Bote**; **De-Moya-Anegón**, 2015; **De-Moya-Anegón** *et al.*, 2018). In *Scopus*, the sources are categorized by main subject areas or disciplines, and by specific Subject Areas or subject categories. There are more than 300 subject categories that are grouped into 26 main subject areas. In addition, there is a main subject area denoted "Multidisciplinary" that contains general journals such as *Nature* or *Science*.

*Patstat* ("EPO worldwide PATent STATistical Database") is a worldwide patent database created by *EPO* in 2008 as a tool for statistical research into patents. *Patstat* has become a de facto standard (**Kang**; **Tarasconi**, 2016) because it has worldwide coverage, it includes more information than other patent databases, and has some special features that other databases do not have. *Patstat* includes

> "…bibliographic information about applications and publications as well as legal information about patents… An application has at least one publication, otherwise it would still be treated as confidential and would not be accessible in any database… Applications which cover the same or similar invention are grouped into families… There are several types of publications, each for a different purpose. Typically the first application is published 18 months after its filing date or its priority date. The granted patent specification is published when patent protection has been granted. There are other kinds of publications, e.g., corrections or publications of search reports, limitations, etc." (*European Patent Office*, 2018).

In this way, one may find very old requests that have not been granted and possibly never will be.

Likewise, each request has a filing date and an earliest publication date, and logically each family will have multiple values for those two dates. When considering families, the earliest values of both dates are used. Two main limitations should be stressed. The first is that *Patstat* contains data from the patent offices that are oriented towards the process of patent examination, and not towards the calculation of bibliometric or informetric indicators. The second limitation is that, although it contains data from all over the world, the data is exchanged with other offices through agreements that can leave some gaps (*European Patent Office*, 2018).

Due to the first limitation, the references in patent applications to Non-Patent Literature (NPL references) are *not* standardized. This poses a major problem in linking them to the publication records in the bibliographic database. **Jefferson** *et al.* (2018) used *PubMed* and *Crossref* as publication databases, but do not indicate how they resolved cases in which more than one DOI were recovered for the same NPL, and how confident they were that the retrieved document corresponds to the reference.

To link the NPL references from the *Patstat* database with the *Scopus* scientific papers, a procedure was applied that was developed by the *SCImago Research Group*. This procedure consists of two phases: a broad generation of *candidate* couples of a patent application cited reference and a linked scientific paper record, and a second phase of validation of these couples. This procedure has been implemented with reasonable results and acceptable costs (**Guerrero-Bote**; **Sánchez-Jiménez**; **De-Moya-Anegón**, 2019). The application of this procedure was limited to scientific publications in *Scopus* published after 2002.

"Like most journal indicators, the Technological Factor (TF) of a journal is calculated on an annual basis, using citations from families of patent applications published in a particular year to papers published in the 5 previous years"

## 3. Methods

For this indicator, we use patent applications which constitute the main record of the *Patstat* database. Not all patent applications are granted and some will never be granted. However, it seems important to us to have all the applications, since if there are several applications that refer to the same invention, it is an indication of the interest that it may have.

Like most journal indicators, the Technological Factor (*TF*) of a journal is calculated on an annual basis, using citations from families of patent applications published in a particular year to papers published in the 5 previous years. It is defined as follows:

$$TF = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{p} w_{ij}$$

in which $w_{ij}$ is the weight of the citation from patent family *j* applied in a particular year and received by the *i*-th document published in the journal in the 5 preceding years; *p* is the total number of patent families citing the journal, and *n* is the total number of citable documents from the journal published in the 5 preceding years (**Guerrero-Bote**; **Moed**; **De-Moya-Anegón**, 2021).

Furthermore, as outlined in **Guerrero-Bote**; **Moed**; **De-Moya-Anegón** (2021), the weight associated with a citation from a patent family is:

$$w_{ij} = \frac{\sum_{k=1}^{c_j} GDP_{jk}}{r_j}$$

in which $GDP_{jk}$ is the portion of the world GDP of the *k*-th country in the year in which the *j*-th patent family applied for protection in it; $C_j$ is the set of countries in which the *j*-th patent family requests protection, and $r_j$ is the total number of citations to scientific papers in the *j*-th patent family.

In this way we use the earliest filing year of the patent family to determine the weight of the family's citations, specifically the portion of the *GDP* of the countries where it requests protection. For everything else we use the earliest publication year of the patent family.

Since the weights of the patent families are all less than or equal to one, and are divided by the references, the *TF* values are very small. For this reason, we normalize these values by calculating a Relative Technological Factor (*TFR*), by dividing the *TF* value of any journal with the mean *TF* value ( $\overline{TF}$ ) over all journals that have *TF* > 0 in a particular year:

$$TFR = \frac{TF}{\overline{TF}}$$

Hence, those journals that received no citations from patent families in a year are not taken into account. In this way *TFR* has a meaning in itself, since the world average is represented by the value unity. *TFR* = 0.8 means that a journal's *TFR* value is 20 per cent below the world average.

## 4. Results

Based on the Spring 2019 version of *Patstat*, Figure 1 shows on the secondary (right-hand) vertical axis the total number of patent families as a function of the application year. On the left-hand vertical axis, Figure 1 plots the number of *two* categories of patent families: families containing at least one citation to a scientific paper; and families containing at least one citation to a 1-5 year old paper, and the number of papers in scientific journals that were cited in patent families, as well as the number of papers that were cited within a 5-year citation window, i.e., cited when they are 1-5 years old relative to the application year of the citing patent families.

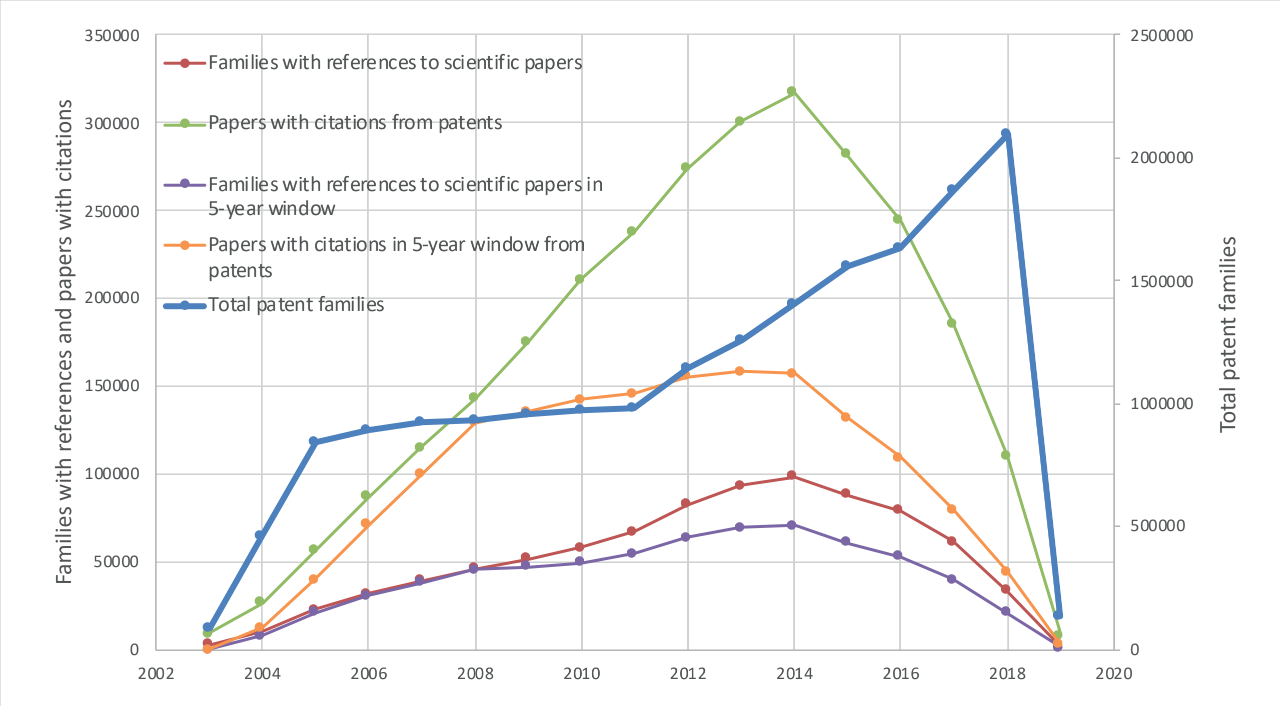


Figure 1. Number of patent families, citations in patent families to scientific papers, and cited papers as a function of the year of publication of the patent's application

Figure 1 shows a monotonic increase in the total number of patent families up until 2018. The sharp fall in 2019, is logical since the *Patstat* version published in the Spring of 2019 contains very few patent applications published in 2019.

The curve corresponding to the families that contain references to scientific papers indexed in *Scopus* is very different from that based on all patent families: not only the absolute numbers are much lower, but also the decline phase starts earlier, namely, as from 2014. Focusing on patent families citing papers within the 5-year window, the curve is similar to the one based on cited papers of all ages, although the absolute numbers are lower from 2008 onwards. As the linking process takes into account only papers published since 2003, in the first years these curves coincide.

The two curves related to number of scientific papers receiving citations from patent families have similar characteristics, to the extent that their absolute numbers by far exceed the numbers of patent families, and both show a decline from 2014 onwards. Two observations must be made. The first relates to the lack of coincidence of the two curves during 2003-2008. Since the analysis takes into account only patent application-citations to papers published after 2003, one would expect these two curves to fully coincide, but there is a constant difference between them. This constant difference is due to the papers that are cited by patent applications published in the same year as the paper or in previous years, within whose citation windows the papers are not included.

A second observation is the decline in the number of citations of patent applications to papers and the number of cited papers after 2014. To further investigate this decline, the results presented in the present communication, based on the *Patstat* version of Spring 2019, are compared with those deriving from the *Patstat* version of one year earlier, Spring 2018.

This comparison shows that the *temporal development* of the curves based on the Spring 2018 dataset is *qualitatively similar* to that of the corresponding curves for the Spring 2019 set (apart from the obvious difference in 2018), but that in the Spring 2019 version the *absolute* numbers of patent families with citations to papers and of cited papers are *larger* than they are in the Spring 2018 version, while the total numbers of patent families derived from the two databases are statistically similar. Also, the increase in numbers of families with citations and cited papers in the 2019 database compared to the 2018 version is greater in recent application years. For example, for the application year 2012 the increase amounts to 12%, for 2016 it is about 50%, and for 2017 over 100%. This might be due to the fact that many evaluation processes are open and that these evaluation processes incorporate references to scientific papers in the families of patent applications. This means that the indicators of those last years may vary in successive years.

Table 1 shows that the number of journals with TF greater than zero is very small in the first years. Indeed, in 2003 there are none. This is because the process of matching cited references in patent applications and scientific papers in *Scopus* was carried out for papers published after 2003, so no journal can have citations from patent applications to papers published in the citation window in 2003 (the citation window would be 1998-2002). In the year 2004 there are few because only one year of the window (1999-2003) is covered.

The maximum number of journals with TF greater than zero is obtained in 2014, the same year as the year in which according to Figure 1 the numbers of patent families containing citations to papers and numbers of cited papers reach their maximum value. As outlined above, the decline after this year is probably due to unfinished patent evaluation processes on the date that the *Patstat* database used in this paper was created. Consequently, indicator calculations for the same application years but based on future *Patstat* versions may yield slightly different results.

> The Technology Factor (TF) average remains fairly stable, especially in the years near the year in which the maximum is reached

The TF average remains fairly stable, especially in the years near the year in which the maximum is reached. Calculating the average of TFR considering only the journals with TF > 0 would have resulted in a value of unity in all years. But taking all journals into account in the calculation of an average TFR, as in the fifth column of Table 1, the value obtained is the fraction of journals with TF > 0. The TFR Max oscillates quite a bit.

Table 1. Journals with TF greater than zero (receiving citations from patent applications in the 5-year window)

| Year | N. journals | Journals with TF | TF Avg. | TFR Avg. | TFR Max |
|---|---|---|---|---|---|
| 2003 | 17903 | 0 | 0 | 0 | 0 |
| 2004 | 18429 | 2561 | 0.00286 | 0.14 | 91.52 |
| 2005 | 19104 | 4691 | 0.00521 | 0.25 | 69.66 |
| 2006 | 20566 | 6129 | 0.00562 | 0.30 | 81.87 |
| 2007 | 22124 | 7130 | 0.00546 | 0.32 | 251.10 |
| 2008 | 23856 | 7976 | 0.00496 | 0.33 | 48.44 |
| 2009 | 25418 | 8451 | 0.00437 | 0.33 | 126.92 |
| 2010 | 27274 | 9082 | 0.00371 | 0.33 | 52.50 |
| 2011 | 28770 | 9564 | 0.00351 | 0.33 | 55.79 |
| 2012 | 30153 | 10168 | 0.00331 | 0.34 | 57.24 |
| 2013 | 30825 | 10566 | 0.00335 | 0.34 | 103.36 |
| 2014 | 31315 | 10663 | 0.00339 | 0.34 | 72.84 |
| 2015 | 32073 | 10012 | 0.00292 | 0.31 | 55.64 |
| 2016 | 32625 | 9299 | 0.00277 | 0.29 | 116.84 |
| 2017 | 32948 | 8278 | 0.00217 | 0.25 | 203.66 |
| 2018 | 31788 | 6457 | 0.00121 | 0.20 | 204.22 |

TF: Technology Factor; N. Journals: number of journals indexed in Scopus; Journals with TF: journals with at least one citation from patent applications in the 5-year window; TF Avg.: annual average of TF (among journals with TF > 0); TFR Avg.: annual average of TFR; TFR Max: the maximum TFR value in a particular year.

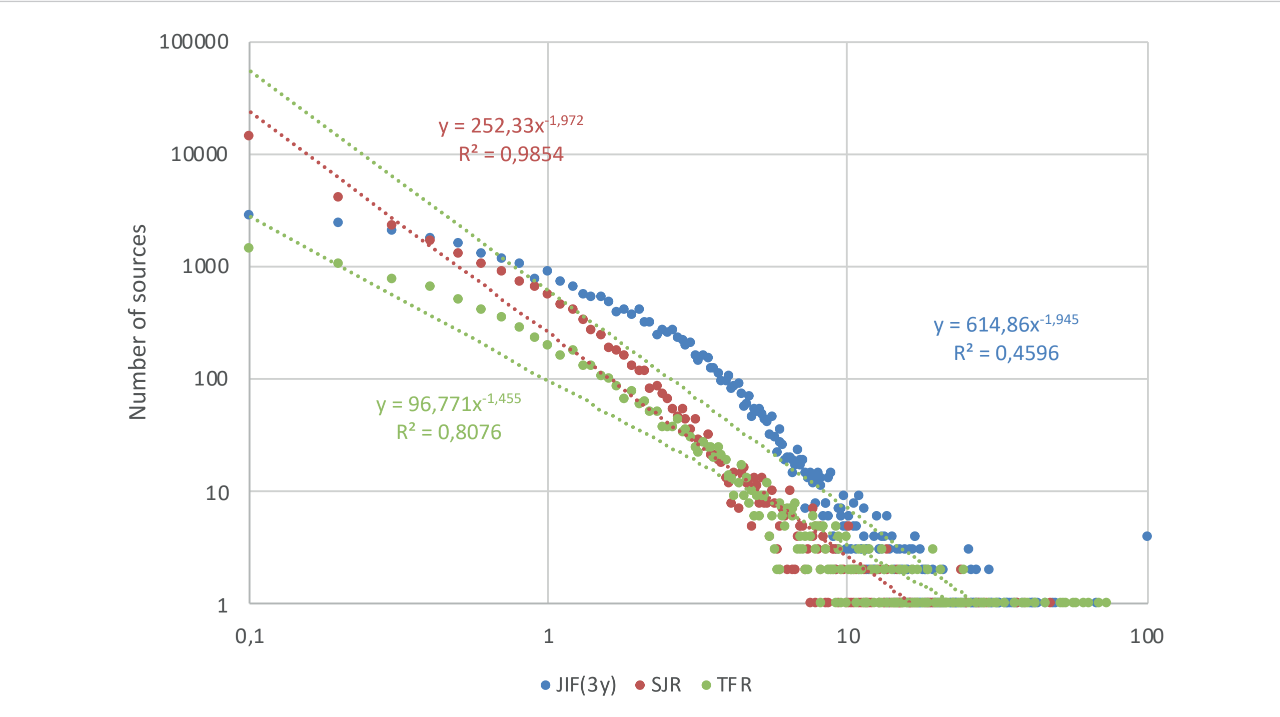


Figure 2. Frequency histogram for 0.1 intervals for TFR, SJR, and JIF (3y). Both axes are logarithmic.

Figure 2 shows a frequency histogram for the three indicators. The axes are on a logarithmic scale. As one can see, the three distributions are power law, and the one with the lowest slope, the lowest exponent, is TFR.

Table 2. Pearson correlation coefficients of journal indicators in 2014 (n = 33,315 sources)

| Pearson | Citables (3y) | Citables (5y) | JIF (3y) | SJR | Cit. Families (5y) | Weight | Pat. Cited. Items | % Pat. cited. items | TFR |
|---|---|---|---|---|---|---|---|---|---|
| Citables (3y) | 1 | 0,97 | 0,08 | 0,08 | 0,02 | 0,56 | 0,56 | 0,04 | 0,01 |
| Citables (5y) | 0,97 | 1 | 0,09 | 0,10 | 0,03 | 0,63 | 0,62 | 0,05 | 0,01 |
| JIF (3y) | 0,08 | 0,09 | 1 | 0,79 | 0,31 | 0,20 | 0,24 | 0,33 | 0,16 |
| SJR | 0,08 | 0,10 | 0,79 | 1 | 0,23 | 0,18 | 0,24 | 0,25 | 0,07 |
| Cit. Families (5y) | 0,02 | 0,03 | 0,31 | 0,23 | 1 | 0,22 | 0,19 | 0,91 | 0,79 |
| Weight | 0,56 | 0,63 | 0,20 | 0,18 | 0,22 | 1,00 | 0,91 | 0,26 | 0,19 |
| Pat. Cited. Items | 0,56 | 0,62 | 0,24 | 0,24 | 0,19 | 0,91 | 1 | 0,25 | 0,11 |
| % Pat. Cited. Items | 0,04 | 0,05 | 0,33 | 0,25 | 0,91 | 0,26 | 0,25 | 1 | 0,72 |
| TFR | 0,01 | 0,01 | 0,16 | 0,07 | 0,79 | 0,19 | 0,11 | 0,72 | 1 |

Citables (3y): citable documents (articles, reviews, conference papers, and short surveys) published in the journal in the three preceding years. Citables (5y): idem published in the preceding 5 years. JIF (3y): Journal Impact Factor calculated with a 3-year window. SJR: SCImago Journal Rank. Cit. Families (5y): number of citing families published in 2014 that cite papers published in the previous 5 years, relativized by dividing it by the number of Citables (5y). Weight: weight accumulated by the journal from the citations of the patent families published in 2014 citing papers published in the previous 5 years. Pat. Cited. Items: number of papers published in the preceding 5 years cited in the patent families published in 2014. % Pat. Cited Items: idem, but in percentage terms. TFR: Relative Technological Factor.

Table 2 shows the Pearson correlation coefficients between various indicators. We have to clarify that the *Citing families* indicators are relativized by dividing their value by the *Citables (5y)*, but not *Weight* (which if divided by the number *Citables (5y)* would completely correlate with TFR). Likewise, neither is the *Pat* column relativized by dividing its values by *Cited items* (in that case it would fully correlate with *% Pat. cited items*).

One can distinguish four subgroups. First of all, there are the two production indicators (*Citable* for three and five years). A next subgroup contains the scientific impact indicators (SJR and JIF *(3y)*), and a third subgroup the absolute indicators related to patent applications (*Pat. cited items* and *Accumulated weight*). Finally, a fourth group contains the relative patent application-based indicators (Family citations to the previous five years relativized, % Pat. cited items and TFR). A moderate correlation is seen between the indicators of production and the absolute patent application based indicators, a quite logical result as both subgroups contain size independent indicators. All other correlations tend to be weak.

The correlations of the TFR with the SJR and the JIF (3y) are weak. When the correlations are made by the 27 areas or the 310 specific areas or categories, higher averages are obtained, as can be seen in Table 3.

Table 3. Correlation coefficient averages of the TFR with the SJR and the JIF (3y) and their standard deviations, by areas and by specific areas or categories

| | TFR-SJR | | TFR-JIF(3y) | |
|---|---|---|---|---|
| | Areas | Categories | Areas | Categories |
| Average | 0.22 | 0.27 | 0.28 | 0.32 |
| Std. dev. | 0.18 | 0.25 | 0.18 | 0.25 |

In all cases, TFR shows a somewhat stronger correlation with the JIF (3y) than it does with SJR, perhaps because the TFR is an average equal to the JIF. Nevertheless, both correlations should be qualified as weak. Although these correlation averages are greater than the overall correlations, there are main subject areas with fairly low value correlation coefficients such as Business, Management and Accounting, Earth and Planetary Sciences, Mathematics, and Social Sciences, while Multidisciplinary has correlations above 0.7 and Chemistry above 0.45. Analysing at the level of subject categories, some categories such as Drug Guides, Occupational Therapy, and Pharmacy even show negative correlations, but others, including Emergency Nursing, Nuclear Energy and Engineering, Ceramics and Composites, show strong positive correlations with Pearson correlation coefficients above 0.8.

As we already mentioned in the Introduction, the distribution of citations from patent applications is very skewed: many patent applications do not contain any citations to scientific papers (**Verbeek** *et al.*, 2002), and the papers cited are concentrated in a small number of journals (**Callaert** *et al.*, 2006). This means that the TFR cannot be expected to be as statistically stable as other journal indicators, and even less so in recent years where not all citations from patent applications have yet been obtained. Calculating the Pearson correlation coefficient between TFR scores obtained by journals in different years, Table 4 reveals positive, statistically significant values for most pairs of years, in the range between 0.50 and 0.75 (for JIF (3y), this correlation is greater, oscillating between 0.7 and 0.96, and for SJR even greater, between 0.8 and 0.97). However, lower correlations are obtained in recent years.

"The three distributions are power law, and the one with the lowest slope, the lowest exponent, is Relative Technological Factor (TFR)"

A lower correlation with the adjacent years is also observed, which is a bit strange.

Table 4. Pearson coefficients for the correlation between journals' TFR values in different years

| Year | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2008 | 1 | 0.50 | 0.59 | 0.71 | 0.71 | 0.65 | 0.64 | 0.64 | 0.64 | 0.58 | 0.52 |
| 2009 | 0.50 | 1 | 0.61 | 0.56 | 0.76 | 0.72 | 0.70 | 0.70 | 0.70 | 0.65 | 0.54 |
| 2010 | 0.59 | 0.61 | 1 | 0.49 | 0.46 | 0.65 | 0.67 | 0.60 | 0.64 | 0.63 | 0.58 |
| 2011 | 0.71 | 0.56 | 0.49 | 1 | 0.54 | 0.47 | 0.76 | 0.67 | 0.67 | 0.63 | 0.53 |
| 2012 | 0.71 | 0.76 | 0.46 | 0.54 | 1 | 0.57 | 0.52 | 0.66 | 0.64 | 0.60 | 0.51 |
| 2013 | 0.65 | 0.72 | 0.65 | 0.47 | 0.57 | 1 | 0.58 | 0.52 | 0.70 | 0.63 | 0.55 |
| 2014 | 0.64 | 0.70 | 0.67 | 0.76 | 0.52 | 0.58 | 1 | 0.57 | 0.39 | 0.57 | 0.49 |
| 2015 | 0.64 | 0.70 | 0.60 | 0.67 | 0.66 | 0.52 | 0.57 | 1 | 0.58 | 0.30 | 0.50 |
| 2016 | 0.64 | 0.70 | 0.64 | 0.67 | 0.64 | 0.70 | 0.39 | 0.58 | 1 | 0.32 | 0.14 |
| 2017 | 0.58 | 0.65 | 0.63 | 0.63 | 0.60 | 0.63 | 0.57 | 0.30 | 0.32 | 1 | 0.54 |
| 2018 | 0.52 | 0.54 | 0.58 | 0.53 | 0.51 | 0.55 | 0.49 | 0.50 | 0.14 | 0.54 | 1 |

Since 2014 is the most complete year that we have in our data, in the remaining part of this section, analyses will be presented for this year. The first relates to a breakdown of sources into source types.

In *Scopus*, sources are classified into 4 types: Scientific Journals, Book Series, Conference Proceedings, and Trade Journals. Table 5 shows averages of a series of indicators broken down into the four types of *Scopus* sources. In terms of number of sources, the major type is Scientific Journals, followed by Conference Proceedings. However, the type Scientific Journals is third in terms of average TFR as well as number of citations from patent applications per paper. When accumulated by source, the highest averages are obtained (Weight by Source and Cit.Fam. By Source Avg.). The source type with the highest average TFR is Conference Proceedings, followed by Trade Journals. It is also the type with the greatest percentage of cited papers. This is quite logical due to the importance of this type in such scientific areas as Engineering or Computer Science which are the disciplines that contribute most to patents. Unsurprisingly, the disciplines with rapid obsolescence of the scientific literature are those that seek both faster publication and the protection of inventions that have to be exploited quickly.

> The correlations of the Relative Technological Factor (TFR) with the SJR and the JIF (3y) are weak

Table 5. Indicators by type of sources (for the year 2014)

| Source type | SJR Avg. | TFR Avg. | TFR Max. | JIF(3y) Avg. | Cit. Fam by paper Avg. | % Cited | N. Sources | N. Papers 2014 Avg. | Citables (5y) Avg. | Cit. Fam by Source Avg. | Weight by Source Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Scientific Journals | 0.68 | 0.19 | 61.84 | 1.49 | 0.02 | 1.96 | 22811 | 97.11 | 387.43 | 6.61 | 0.51 |
| Book Series | 0.43 | 0.10 | 11.06 | 1.86 | 0.05 | 3.13 | 859 | 155.98 | 521.99 | 3.35 | 0.36 |
| Conference Proceedings | 0.19 | 0.85 | 72.84 | 0.65 | 0.06 | 4.03 | 7224 | 20.10 | 183.30 | 3.42 | 0.43 |
| Trade Journals | 0.13 | 0.29 | 51.27 | 0.13 | 0.02 | 1.10 | 421 | 60.78 | 285.60 | 1.16 | 0.20 |

SJR Avg.: Scimago Journal Rank average, TFR Avg.: Relative Technological Factor average; TFR Max.: Maximum value of the Relative Technological Factor; JIF (3y): Journal Impact Factor (3 years) average; Cit. Fam by paper average: Average number of patent family citations per paper; % Cited: Percentage of Cited Papers; N. Papers average in 2014: Average number of papers in 2014; Citables (5y) Avg.: Average number of citable documents published in the preceding five years; Cit. Fam. By Journal Avg.: Average citations of patent families per source; Weight by Source Avg.: Average number of weighted patent family citations per source.

The next analysis presents a breakdown by main subject category. Figure 3 is a scatter plot of the average TFR versus the average SJR for 26 main subject areas. In addition, both the number of sources with TFR > 0 and the different types of sources are shown in concentric circles, the radii of which are proportional to the number of sources. For details, the reader is referred to the figure legend. One observes that most of the subject areas have an average TFR value less than unity. This is quite logical since TFR is a normalized indicator so that the overall average of the sources with TFR > 0 is itself unity, and, as one sees in Table 1, only about a third of the sources receive citations from patent applications.

Computer Science is the area with the largest average TFR, being the only area whose value exceeds unity. It is also the only area in which the type Conference Proceedings ranks first in terms of number of sources, well above Scientific Journals. Engineering shows a similar picture. However, in other areas, conference proceedings have a modest TFR value. Computer Science and Engineering are also the areas with the lowest SJR, precisely because of the large number of conference proceedings that are being published in these fields, which thus appear to be relatively less frequently cited in the scientific literature but more often cited by the patent families. One observes that the trend line is practically flat.

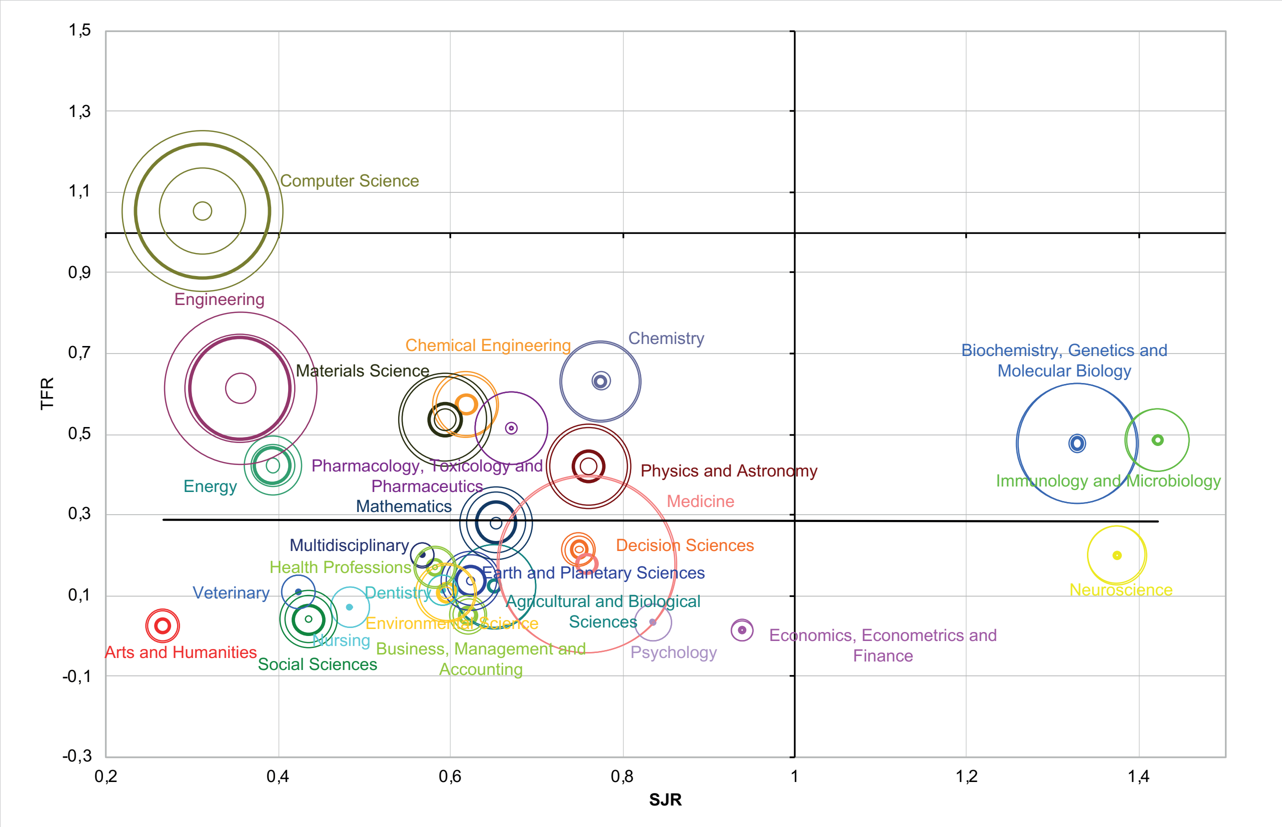


Figure 3. Scatter plot of main subject areas by average TFR and average SJR. For each point, the radius of the largest circle is proportional to the number of sources with TFR > 0, the radius of the thickest circle is proportional to the number of conference proceedings with TFR > 0, the radius of the next-to-thinnest circumference circle is proportional to the number of journals with TFR > 0, and radius of the smallest circle is proportional to the number of sources of other types with TF > 0 (Book Series and Trade Journals). Data refer to the year 2014.

This can be seen in detail in Table 6, where the total number of Citing Families, the Accumulated Weight, the Avg. TFR, the sum of the TFR, and the percentage of sources with TFR > 0 are given for each area, and then, for each type of source, the avg. TFR, the sum of the TFR, and the percentage of sources with TFR > 0. All this for the year 2014.

One observes from Table 6 that, although Computer Science, Chemistry, and Engineering are the areas with the highest average TFR, the ones that accumulate the most citing families are Biochemistry, Genetics and Molecular Biology and Medicine, and the one that accumulates the most weight is Engineering followed by Computer Science. Economics, Econometrics and Finance, Arts and Humanities, Psychology, and Social Sciences have the lowest average TFR. The low scores for Arts and Humanities are consistent with the findings related to this main subject field presented by **De-Moya-Anegón** *et al.* (2020).

It is striking that the Arts and Humanities scores in citing families and accumulated weight are not the last, although they are in the low central area.

Figure 4 shows 4 scatter plots of the Chemistry main area and of the subject categories Media Technology, Nanoscience and Nanotechnology, and Library and Information Sciences. The TFR is represented on the horizontal axis, while two vertical axes are used –the one on the left for SJR and the one on the right for JIF (3y). Thus, each source is represented by two points which are vertically displaced, blue corresponds to SJR and yellow to JIF (3y). The points that appear on the vertical axis on the left correspond to sources whose TFR is zero because they have no citation from patent family applications.

Figure 5 presents scatter plots of the TFR and SJR indicator values of journals not only in the main subject field Chemistry, but also at a lower level of field aggregation in three subject categories: Media Technology, Nanoscience and Nanotechnology, and Library and Information Sciences.

> The source type with the highest average Relative Technological Factor (TFR) is Conference Proceedings

Table 6. Multiple indicators by main subject area and type of source

| Scientific area | Citing Families | Accumulated Weight | TFR Avg. | TFR Sum. | %Sources with TFR | SJ. TFR | SJ. Sum TFR | SJ. %Src. w. TFR | CP. TFR | CP. Sum TFR | CP. %Src. w. TFR | BS. TFR | BS. Sum TFR | BS. %Src. w. TFR | TJ. TFR | TJ. Sum TFR | TJ. %Src. w. TFR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Multidisciplinary | 4813 | 247.29 | 0.20 | 23.2 | 46.55 | 0.20 | 23.1 | 46.49 | 0.04 | 0.0 | 100 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 |
| Agricultural and Biological Sciences | 9067 | 698.54 | 0.12 | 252.6 | 35.56 | 0.13 | 244.2 | 36.52 | 0.06 | 3.1 | 25.00 | 0.10 | 3.7 | 18.92 | 0.06 | 1.6 | 10.71 |
| Arts and Humanities | 1980 | 125.65 | 0.02 | 88.2 | 2.94 | 0.01 | 35.2 | 2.65 | 0.96 | 53.0 | 38.18 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 |
| Biochemistry, Genetics and Molecular Biology | 53134 | 3097.68 | 0.48 | 950.6 | 75.70 | 0.49 | 913.1 | 78.02 | 0.24 | 10.8 | 28.89 | 0.42 | 24.7 | 38.98 | 0.23 | 2.0 | 66.67 |
| Business, Management and Accounting | 501 | 60.89 | 0.05 | 75.6 | 9.54 | 0.03 | 37.6 | 8.56 | 0.18 | 31.2 | 17.44 | 0.003 | 0.2 | 1.64 | 0.08 | 6.6 | 12.82 |
| Chemical Engineering | 17994 | 1548.21 | 0.57 | 430.5 | 58.85 | 0.71 | 369.6 | 71.51 | 0.31 | 46.3 | 26.49 | 0.10 | 2.7 | 33.33 | 0.24 | 11.9 | 38.00 |
| Chemistry | 33928 | 3004.23 | 0.63 | 570.1 | 72.87 | 0.65 | 514.3 | 78.45 | 1.16 | 30.1 | 38.46 | 0.32 | 16.2 | 24.00 | 0.25 | 9.5 | 44.74 |
| Computer Science | 28282 | 3347.76 | 1.05 | 5970.9 | 47.62 | 0.67 | 921.2 | 54.68 | 1.19 | 5018.0 | 45.30 | 0.51 | 25.3 | 52.00 | 0.34 | 6.4 | 36.84 |
| Decision Sciences | 571 | 60.42 | 0.21 | 92.4 | 25.87 | 0.11 | 33.7 | 25.24 | 0.55 | 58.0 | 26.67 | 0.03 | 0.3 | 30.00 | 0.34 | 0.3 | 100 |
| Earth and Planetary Sciences | 1719 | 219.54 | 0.14 | 197.4 | 24.13 | 0.10 | 104.0 | 24.07 | 0.29 | 91.3 | 28.75 | 0.01 | 0.4 | 7.55 | 0.06 | 1.6 | 6.90 |
| Economics, Econometrics and Finance | 91 | 10.17 | 0.01 | 12.3 | 4.49 | 0.01 | 9.8 | 4.46 | 0.09 | 2.4 | 11.54 | 0.00 | 0.0 | 1.89 | 0.00 | 0.0 | 0.00 |
| Energy | 5449 | 664.75 | 0.42 | 340.3 | 42.43 | 0.47 | 171.3 | 50.82 | 0.42 | 165.2 | 35.73 | 0.04 | 0.3 | 33.33 | 0.09 | 3.5 | 32.50 |
| Engineering | 40566 | 4722.82 | 0.61 | 3130.9 | 47.00 | 0.50 | 1178.3 | 52.24 | 0.76 | 1842.4 | 43.81 | 0.11 | 8.1 | 36.49 | 0.47 | 102.0 | 29.36 |
| Environmental Science | 4960 | 492.06 | 0.11 | 156.8 | 24.69 | 0.11 | 132.6 | 25.89 | 0.10 | 19.1 | 19.79 | 0.01 | 0.4 | 11.11 | 0.14 | 4.6 | 23.53 |
| Immunology and Microbiology | 12727 | 683.02 | 0.49 | 260.6 | 75.37 | 0.49 | 256.4 | 75.81 | 0.00 | 0.00 | 0.00 | 0.42 | 4.2 | 60.00 | 0.00 | 0.0 | 0.00 |
| Materials Science | 28720 | 2915.83 | 0.54 | 795.7 | 59.49 | 0.61 | 631.7 | 70.34 | 0.47 | 135.9 | 36.77 | 0.17 | 4.8 | 50.00 | 0.18 | 23.2 | 25.20 |
| Mathematics | 7506 | 886.25 | 0.28 | 512.7 | 29.65 | 0.14 | 184.3 | 28.14 | 0.65 | 322.3 | 33.81 | 0.14 | 6.1 | 27.91 | 0.00 | 0.0 | 0.00 |
| Medicine | 52426 | 3078.91 | 0.18 | 1262.2 | 46.45 | 0.17 | 1196.5 | 46.80 | 0.44 | 50.2 | 35.09 | 0.19 | 15.0 | 33.33 | 0.08 | 0.5 | 33.33 |
| Neuroscience | 5026 | 229.99 | 0.20 | 108.0 | 63.03 | 0.18 | 95.9 | 63.74 | 1.57 | 11.0 | 42.86 | 0.11 | 1.1 | 40.00 | 0.00 | 0.0 | 0.00 |
| Nursing | 1021 | 71.79 | 0.07 | 42.9 | 24.83 | 0.07 | 42.7 | 24.92 | 0.00 | 0.00 | 0.00 | 0.04 | 0.2 | 25.00 | 0.00 | 0.0 | 0.00 |
| Pharmacology, Toxicology and Pharmaceutics | 14522 | 991.45 | 0.51 | 401.9 | 70.42 | 0.52 | 387.3 | 71.85 | 0.00 | 0.00 | 0.00 | 0.60 | 10.8 | 33.33 | 0.26 | 3.9 | 53.33 |
| Physics and Astronomy | 22493 | 2354.47 | 0.42 | 561.7 | 54.38 | 0.44 | 430.6 | 62.73 | 0.38 | 116.1 | 30.19 | 0.11 | 3.7 | 30.30 | 0.57 | 11.4 | 60.00 |
| Psychology | 483 | 29.02 | 0.03 | 36.9 | 12.72 | 0.03 | 30.8 | 12.69 | 1.98 | 6.0 | 66.67 | 0.01 | 0.2 | 5.56 | 0.00 | 0.0 | 0.00 |
| Social Sciences | 1400 | 137.78 | 0.04 | 243.5 | 5.32 | 0.02 | 97.3 | 4.28 | 0.35 | 141.6 | 22.69 | 0.01 | 3.5 | 0.93 | 0.03 | 1.0 | 3.12 |
| Veterinary | 803 | 62.43 | 0.11 | 24.2 | 47.51 | 0.11 | 24.2 | 47.73 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 |
| Dentistry | 327 | 26.73 | 0.11 | 20.1 | 47.78 | 0.11 | 20.1 | 47.78 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 | 0.0 | 0.00 |
| Health Professions | 1320 | 121.66 | 0.17 | 92.2 | 32.66 | 0.13 | 60.8 | 31.59 | 0.53 | 31.0 | 40.68 | 0.10 | 0.2 | 50.00 | 0.07 | 0.2 | 33.33 |

Citing Families: Number of patent family citations; Accumulated Weight: Weighted patent families citations accumulated; TFR Avg.: Relative Technological Factor average; TFR Sum.: Sum of the value of the Relative Technological Factor; %Sources with TFR: Percentage of sources with TF and TFR > 0; SJ. TFR: Relative Technological Factor average of Scientific Journals; SJ. Sum TFR: Sum of the value of the Relative Technological Factor of Scientific Journals; SJ. %Src. w. TFR: Percentage of Scientific Journals with TF and TFR > 0; CP. TFR: Relative Technological Factor average of Conference Proceedings; CP. Sum TFR: Sum of the value of the Relative Technological Factor of Conference Proceedings; CP. %Src. W. TFR: Percentage of Conference Proceedings with TF and TFR > 0; BJ. TFR: Relative Technological Factor average of Book Series; BS. Sum TFR: Sum of the value of the Relative Technological Factor of Book Series; BJ. %Src. W. TFR: Percentage of Book Series with TF and TFR > 0; TJ. TFR: Relative Technological Factor average of Trade Journals; TJ. Sum TFR: Sum of the values of the Relative Technological Factor of Trade Journals; TJ. %Src. W. TFR: Percentage of Trade Journals with TF and TFR > 0.

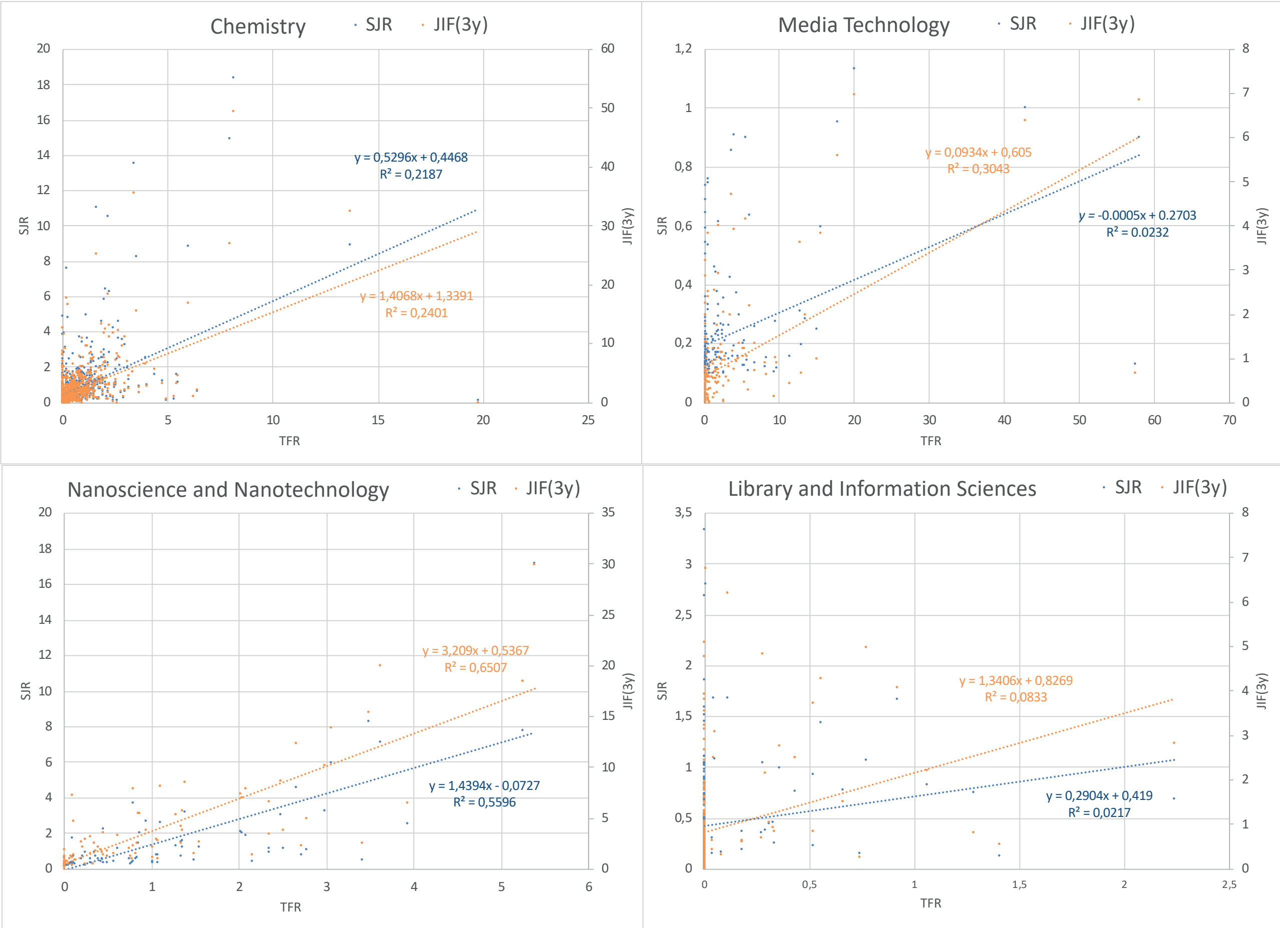


Figure 4. Scatter plots of TFR *vs* SJR and JIF (3y) of a main subject area and three subject categories.

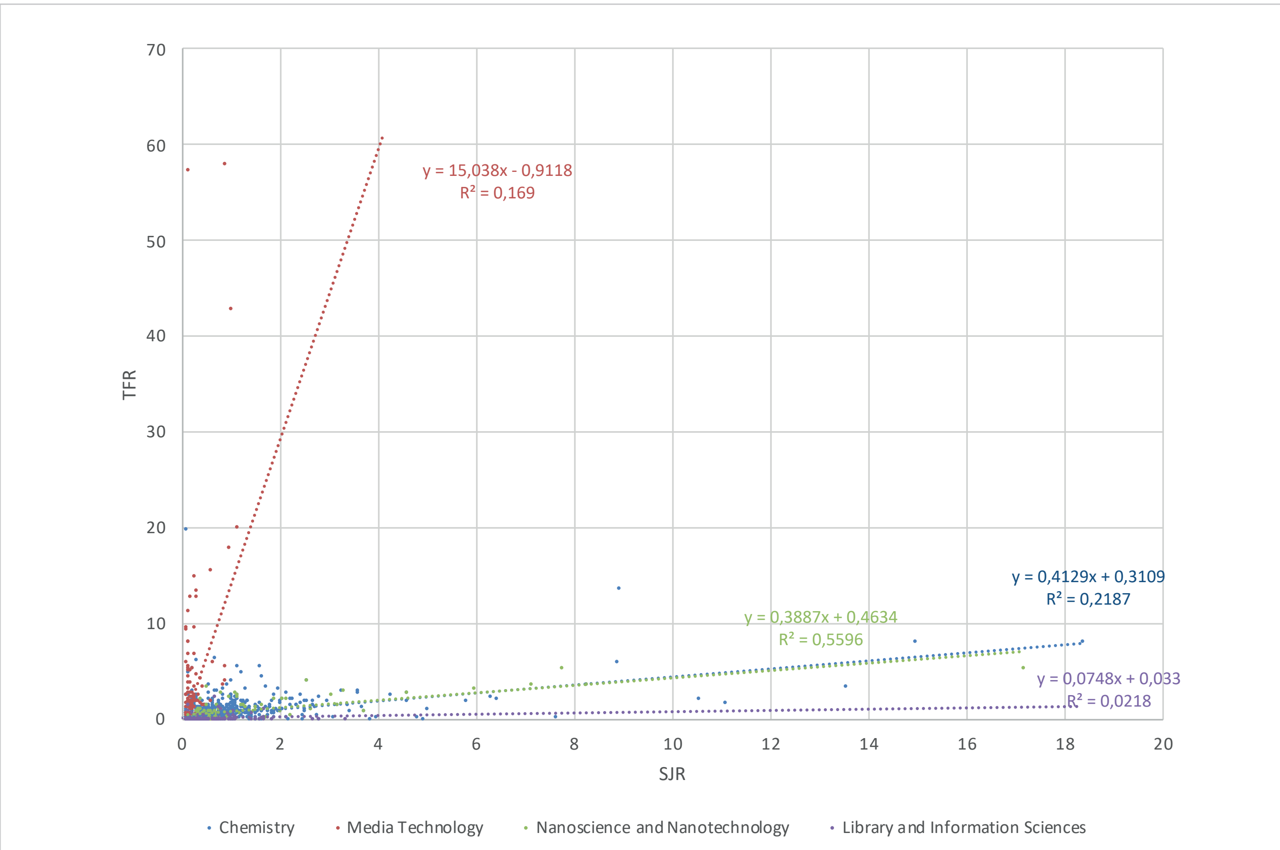


Figure 5. Scatter plot of TFR versus SJR for a main subject field and three subject categories

The Chemistry area has the second highest average TFR (0.63) and has a moderate correlation with SJR (0.47) and with JIF (3y) (0.49). It has 658 publications with positive TFR, which is 73% of the total number of sources in this main subject field. Most of the contribution to the TFR comes from Scientific Journals.

> Media Technology is the category with the highest average Relative Technological Factor (TFR) (2.39)

Media Technology is the category with the highest average TFR (2.39) and has a moderate correlation with SJR (0.41) and JIF (3y) (0.55). It is included within the Engineering area. It has 110 sources that receive citations from patent applications, which represents 53% of the total number of sources in Media Technology. Most of the contribution to the average TFR corresponds to Conference Proceedings followed by Scientific Journals.

Nanoscience and Nanotechnology is a category included in the Materials Science area with an average TFR close to unity (0.99), but which has a high correlation with SJR (0.75) and with JIF (3y) (0.81). All sources that have TFR are journals, in total 76, representing 83% of the total number of sources in this category.

To show data of a category that did not have a very high average TFR, Library and Information Sciences, included within the Social Sciences area, was considered a good example. This category has a low average TFR, and in it the TFR shows little correlation with either SJR or JIF (3y). Only 29 sources, 13.5%, have citations from patent applications. Although the average TFR is low, this result is consistent with the evidence of the technological impact of this area found by **Halevi** and **Moed** (2012). The vast majority of sources with TFR are Scientific Journals. Table 7 presents a list of sources that obtained at least one citation from patent applications made in 2014. Only 29 sources have TFR > 0, which is 13%. One can also see that most of the TFR is provided by scientific journals.

Table 7. TFR ordered list of journals obtaining citations from patent applications submitted in 2014

| Source | SJR | JIF(3y) | TFR |
|---|---|---|---|
| *Information retrieval* | 0.69 | 2.808 | 2.24 |
| *Journal of information and organizational sciences* | 0.12 | 0.541 | 1.41 |
| *Library collections, acquisition and technical services* | 0.75 | 0.816 | 1.29 |
| *Journal of information science* | 0.82 | 2.203 | 1.06 |
| *IEEE Transactions on information theory* | 1.67 | 4.060 | 0.92 |
| *International journal of information management* | 1.06 | 4.981 | 0.78 |
| *Information design journal* | 0.15 | 0.270 | 0.74 |
| *Journal of documentation* | 0.78 | 1.500 | 0.66 |
| *Journal of chemical information and modeling* | 1.43 | 4.263 | 0.56 |
| *Journal of cheminformatics* | 0.93 | 3.711 | 0.52 |
| *International journal of metadata, semantics and ontologies* | 0.23 | 0.836 | 0.52 |
| *Information processing and management* | 0.76 | 2.504 | 0.44 |
| *International journal of geographical information science* | 0.99 | 2.761 | 0.36 |
| *Journal of information science and engineering* | 0.25 | 0.833 | 0.34 |
| *Library trends* | 0.46 | 0.924 | 0.33 |
| *Information research* | 0.45 | 1.054 | 0.31 |
| *Language resources and evaluation* | 0.38 | 2.164 | 0.29 |
| *European journal of information systems* | 1.05 | 4.824 | 0.28 |
| *International journal of data mining and bioinformatics* | 0.36 | 0.696 | 0.27 |
| *Annals of library and information studies* | 0.36 | 0.614 | 0.18 |
| *Progress in informatics* | 0.19 | 0.630 | 0.18 |
| *Government information quarterly* | 1.68 | 6.192 | 0.11 |
| *Journal of information and computational science* | 0.17 | 0.328 | 0.08 |
| *Scientometrics* | 1.08 | 3.076 | 0.05 |
| *Library and information science research* | 1.68 | 2.486 | 0.05 |
| *16th Americas conference on information systems 2010, AMCIS 2010* | 0.15 | 0.440 | 0.04 |
| *Lecture notes in control and information sciences* | 0.31 | 0.625 | 0.04 |
| *Information systems research* | 2.8 | 6.746 | 0.007 |
| *Proceedings of the ASIST annual meeting* | 0.28 | 0.533 | 0.002 |

## 5. Conclusions

This study has presented the development of the Relative Technology Factor (TFR), a new metric of the technological impact of scholarly sources that aims to measure their contribution to technological progress through the analysis of citations to the scientific literature from patent applications.

Methodologically, the TFR assigns to a citation from a patent family a weight that depends upon the geographical coverage of the protection requested in a patent application, the economic benefit it is expected to generate, and the patent family's propensity to cite the scientific literature. The cited reference list in a patent family is created by accumulating and de-duplicating the references in the various members of the family.

A 5-year citation window is used, because the patent evaluation and publication process takes a long time for a paper to be cited. The said evaluation and publication process means that the values are not stable until several years after the patent applications' publication year, although they can be calculated by the end of that year.

To give meaning to the value of the indicator, it is normalized so that the annual average of the journals that have citations from patent applications is unity. In this way, with this single value one can know how a journal performs relative to a 'world' average across all journals.

The statistical characterization of the TFR and its comparison with the SJR and the JIF (3y) shows that the distribution of TFR scores among sources can be modeled as a power law distribution, with the slope being a little smoother than that of the SJR and JIF (3y). However, the correlations between the three indicators are mostly low or moderate, which shows that they measure quite different aspects of journal or research performance. Slightly higher correlation coefficients are obtained at the level of main subject fields, with some subject categories showing a strong, positive correlation and others a very low correlation. This is coherent with the studies of **Huang**, **Huang**, & **Chen** (2014) and **Liaw** *et al.* (2014).

> Relative Technological Factor (TFR) assigns to a citation from a patent family a weight that depends upon the geographical coverage of the protection requested in a patent application, the economic benefit it is expected to generate, and the patent family's propensity to cite the scientific literature

Only around a third of the journals have at least one citation from patent applications. In other words, two-thirds have a zero value of TFR. Those with values greater than zero are not necessarily those with higher SJR or JIF (3y).

## 6. Limitations

- The indicator for each year is calculated on the basis of the patent applications published in that year. However, the evaluation process does not stop incorporating references in the following years. This means that if the indicators for 2019 are calculated this year with the *Patstat* version of spring 2020 that already contains the 2019 applications, these indicators may vary in the following years, because the families of applications published in 2019 will continue incorporating references in the coming years.
- The indicator has been calculated for all the types of sources present in *Scopus*, as is done with SJR and JIF (3y). This assumes that some sources, such as conference proceedings, may have a periodicity longer than annual, but nevertheless may have TFR > 0 in years in which they do not have any publication. The opposite may occur as well. For example, a biannual Conference that took place in 2019 may be cited by patent applications published in 2020 while in 2020 it has no publication, and it could be the case that none of the applications published in 2019 cites it (i.e., its publication of 2017).

> The values are not stable until several years after the patent applications' publication year

## 7. References

**Archambault, Éric**; **Campbell, David**; **Gingras, Yves**; **Larivière, Vincent** (2009). "Comparing bibliometric statistics obtained from the Web of Science and Scopus". *Journal of the American Society for Information Science and Technology* (*Jasist*), v. 60, n. 7, pp. 1320-1326.
*https://doi.org/10.1002/asi.21062*

**Bergstrom, Carl T.** (2007). "Eigenfactor: Measuring the value and prestige of scholarly journals". *College and research libraries news*, v. 68, n. 5, pp. 314-316.
*https://doi.org/10.5860/crln.68.5.7804*

**Bornmann, Lutz**; **Daniel, Hans-Dieter** (2008). "What do citation counts measure? A review of studies on citing behavior". *Journal of documentation*, v. 64, n. 1, pp. 45-80.
*https://doi.org/10.1108/00220410810844150*

**Callaert, Julie**; **Van-Looy, Bart**; **Verbeek, Arnold**; **Debackere, Koenraad**; **Thijs, Bart** (2006). "Traces of prior art: An analysis of non-patent references found in patent documents". *Scientometrics*, v. 69, n. 1, pp. 3-20.
*https://doi.org/10.1007/s11192-006-0135-8*

**Callon, Michel**; **Courtial, Jean-Pierre**; **Penan, Hervé** (1993). *La scientométrie*. Paris: PUF. ISBN: 2130452493

**De-Moya-Anegón, Félix**; **Chinchilla-Rodríguez, Zaida**; **Vargas-Quesada, Benjamín**; **Corera-Álvarez, Elena**; **Muñoz-Fernández, Francisco-José**; **González-Molina, Antonio**; **Herrero-Solana, Víctor** (2007). "Coverage analysis of Scopus: A journal metric approach". *Scientometrics*, v. 73, n. 1, pp. 53-78.
*https://doi.org/10.1007/s11192-007-1681-4*

**De-Moya-Anegón, Félix**; **Guerrero-Bote, Vicente P.**; **López-Illescas, Carmen**; **Moed, Henk F.** (2018). "Statistical relationships between corresponding authorship, international co-authorship and citation impact of national research systems". *Journal of informetrics*, v. 12, n. 4, pp. 1251-1262.
*https://doi.org/10.1016/j.joi.2018.10.004*

**De-Moya-Anegón, Félix**; **Guerrero-Bote, Vicente P.**; **López-Illescas, Carmen**; **Moed, Henk F.** (2020). "The citation impact of social sciences and humanities upon patentable technology". *Scientometrics* (in press).

*European Patent Office* (2018). *Data catalog Patstat global. Version 5.11. EPO Patstat customers*.
*https://www.epo.org*

**Garfield, Eugene** (1955). "Citation indexes for science: A new dimension in documentation through association of ideas". *Science*, v. 122, n. 3159, pp. 108-111.
*https://doi.org/10.1126/science.122.3159.108*

**Garfield, Eugene** (1972). "Citation analysis as a tool in journal evaluation". *Science*, v. 178, pp. 471-479.
*https://doi.org/10.1126/science.178.4060.471*

**González-Pereira, Borja**; **Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. Journal of informetrics, v. 4, n. 3, pp. 379-391.
*https://doi.org/10.1016/j.joi.2010.03.002*

**Gorraiz, Juan**; **Gumpenberger, Christian**; **Wieland, Martin** (2011). "Galton 2011 revisited: a bibliometric journey in the footprints of a universal genius". *Scientometrics*, v. 88, n. 2, pp. 627-652.
*https://doi.org/10.1007/s11192-011-0393-y*

**Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2012). "A further step forward in measuring journals' scientific prestige: The SJR2 indicator". *Journal of Informetrics*, v. 6, n. 4, pp. 674-688.
*https://doi.org/10.1016/j.joi.2012.07.001*

**Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2015). "Analysis of scientific production in food science from 2003 to 2013". *Journal of food science*, v. 80, n. 12, pp. R2619-R2626.
*https://doi.org/10.1111/1750-3841.13108*

**Guerrero-Bote, Vicente P.**; **Moed, Henk F.**; **De-Moya-Anegón, Félix** (2021). "New indicators of technological impact of scientific production". *Journal of data and information science* (accepted).

**Guerrero-Bote, Vicente P.**; **Sánchez-Jiménez, Rodrigo**; **De-Moya-Anegón, Félix** (2019). "The citation from patents to scientific output revisited: a new approach to Patstat / Scopus matching". *El profesional de la información*, v. 28, n. 4, e280401.
*https://doi.org/10.3145/epi.2019.jul.01*

**Halevi, Gali**; **Moed, Henk F.** (2012). "The technological impact of library science research: A patent analysis". In: Archambault, E.; Gingras, Y.; Larivière, V. (eds.). *Proceedings of 17$^{th}$ International conference on science and technology indicators* (v. 1, pp. 371-380), Montréal: Science-Metrix and OST.
*https://www.ost.uqam.ca/en/publications/proceedings-of-17th-international-conference-on-science-and-technology-indicators*

**Hane, Paula J.** (2004). "Elsevier announces Scopus service". *Information today*.
*http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=16494*

**Huang, Mu-Hsuan**; **Huang, Wei-Tzu**; **Chen, Dar-Zen** (2014). "Technological impact factor: An indicator to measure the impact of academic publications on practical innovation". *Journal of informetrics*, v. 8, n. 1, pp. 241-251.
*https://doi.org/10.1016/j.joi.2013.12.004*

**Jacsó, Péter** (2011). "The h-index, h-core citation rate and the bibliometric profile of the Scopus database". *Online information review*, v. 35, n. 3, pp. 492-501.
*https://doi.org/10.1108/14684521111151487*

**Jefferson, Osmat A.**; **Jaffe, Adam**; **Ashton, Doug**; **Warren, Ben**; **Koellhofer, Deniz**; **Dulleck, Uwe**; **Bilder, G.**; **Ballagh, Aaron**; **Moe, John**; **DiCuccio, Michael**; **Ward, Karl**; **Bilder, Geoff**; **Dolby, Kevin**; **Jefferson, Richard A.** (2018). "Mapping the global influence of published research on industry and innovation". *Nature biotechnology*, v. 36, n. 1, pp. 31-39.
*https://doi.org/10.1038/nbt0818-772a*

**Kang, Byeongwoo**; **Tarasconi, Gianluca** (2016). "*Patstat* revisited: Suggestions for better usage". *World patent information*, v. 46, pp. 56-63.
*https://doi.org/10.1016/j.wpi.2016.06.001*

**Leydesdorff, Loet**; **De-Moya Anegón, Félix**; **Guerrero-Bote, Vicente P.** (2010). "Journal maps on the basis of Scopus data: A comparison with the Journal Citation Reports of the ISI". *Journal of the American Society for Information Science and Technology*, v. 61, n. 2, pp. 352-369.
*https://doi.org/10.1002/asi.21250*

**Liaw, Yi-Ching**; **Chan, Te-Yi**; **Fan, Chin-Yuan**; **Chiang, Cheng-Hsin** (2014). "Can the technological impact of academic journals be evaluated? The practice of non-patent reference (NPR) analysis". *Scientometrics*, v. 101, n. 1, pp. 17-37.
*https://doi.org/10.1007/s11192-014-1337-0*

**Martínez, Catalina** (2011). "Patent families: When do different definitions really matter?". *Scientometrics*, v. 86, pp. 39-63.
*https://doi.org/10.1007/s11192-010-0251-3*

**Merton, Robert K.** (1957). "Priorities in scientific discovery: a chapter in the sociology of science". *American sociological review*, v. 22, n. 6, pp. 635-659.
*https://doi.org/10.2307/2089193*

**Moed, Henk F.** (2005). *Citation analysis in research evaluation*. Springer. ISBN: 978 1 4020 3713 9
*https://doi.org/10.1007/1-4020-3714-7*

**Moed, Henk F.** (2010). "Measuring contextual citation impact of scientific journals". *Journal of informetrics*, v. 4, n. 3, pp. 265-277.
*https://doi.org/10.1016/j.joi.2010.01.002*

**Narin, Francis**; **Hamilton, Kimberley S.**; **Olivastro, Dominic** (1997). "The increasing linkage between US technology and public science". *Research policy*, v. 26, n. 3, pp. 317-330.
*https://doi.org/10.1016/s0048-7333(97)00013-9*

**Narin, Francis**; **Olivastro, Dominic** (1992). "Status report: Linkage between technology and science". *Research policy*, v. 21, n. 3, pp. 237-249.
*https://doi.org/10.1016/0048-7333(92)90018-y*

**Pickering, Bobby** (2004). "Elsevier prepares Scopus to rival ISI Web of science". *Information world review*, n. 8.

**Pinski, Gabriel**; **Narin, Francis** (1976). "Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics". *Information processing and management*, v. 12, n. 5, pp. 297-312.
*https://doi.org/10.1016/0306-4573(76)90048-0*

**Price, Derek J. De-Solla** (1963). *Little science, big science*. New York: Columbia University Press. ISBN: 978 0 231918442
*https://doi.org/10.7312/pric91844*

**Salkind, Neil J.**; **Rainwater, Terese** (2003). *Exploring research*. Upper Saddle River, NJ: Prentice Hall. ISBN: 978 9353068677

**Van-Raan, Anthony F. J.** (2017). "Patent citations analysis and its value in research evaluation: A review and a new approach to map technology-relevant research". *Journal of data and information science*, v. 2, n. 1, pp. 13-50.
*https://doi.org/10.1515/jdis-2017-0002*

**Verbeek, Arnold**; **Debackere, Koenraad**; **Luwel, Marc**; **Andries, Petra**; **Zimmermann, Edwin**; **Deleus, Filip** (2002). "Linking science to technology: Using bibliographic references in patents to build linkage schemes". *Scientometrics*, v. 54, n. 3, pp. 399-420.
*https://doi.org/10.1023/a:1016034516731*

**Waltman, Ludo**; **Van-Eck, Nees-Jan**; **Van Leeuwen, Thed N.**; **Visser, Martijn S.** (2013). "Some modifications to the SNIP journal impact indicator". *Journal of informetrics*, v. 7, n. 2, pp. 272-285.
*https://doi.org/10.1016/j.joi.2012.11.011*

**Zitt, Michel**; **Small, Henry** (2008). "Modifying the journal impact factor by fractional citation weighting: The audience factor". *Journal of The American Society for Information Science and Technology*, v. 59, n. 11, pp. 1856-1860.
*https://doi.org/10.1002/asi.20880*