

New Indicators of the Technological Impact of Scientific Production

Vicente P. Guerrero-Bote^{1†}, Henk F. Moed², Félix Moya-Anegón³

¹SCImago Group, Dept. Information and Communication, University of Extremadura, Badajoz, Spain

²Sapienza University of Rome, Italy

³SCImago Group, Madrid, Spain

Citation: Guerrero-Bote, V.P., Moed, H.F., & Moya-Anegón, F. (2021). New indicators of the technological impact of scientific production. *Journal of Data and Information Science*, 6(4), 36–61. <https://doi.org/10.2478/jdis-2021-0028>

Received: Mar. 5, 2021

Revised: May 6, 2021

Accepted: Jun. 3, 2021

Abstract

Purpose: Building upon pioneering work by Francis Narin and others, a new methodological approach to assessing the technological impact of scientific research is presented.

Design/methodology/approach: It is based on the analysis of citations made in patent families included in the PATSTAT database that is to scientific papers indexed in Scopus.

Findings: An advanced citation matching procedure is applied to the data in order to construct two indicators of technological impact: on the citing (patent) side, the country/region in which protection is sought and a patent family's propensity to cite scientific papers are taken into account, and on the cited (paper) side, a relative citation rate is defined for patent citations to papers that is similar to the scientific paper-to-paper citation rate in classical bibliometrics.

Research limitations: The results are limited by the available data, in our case Scopus and PATSTAT, and especially by the lack of standardization of references in patents. This required a matching procedure that is neither trivial nor exact.

Practical implications: Results at the country/region, document type, and publication age levels are presented. The country/region-level results in particular reveal features that have remained hidden in analyses of straight counts. Especially notable is that the rankings of some Asian countries/regions move upwards when the proposed normalized indicator of technological impact is applied as against the case with straight counts of patent citations to those countries/regions' published papers.

Originality/value: In our opinion, the level of sophistication of the indicators proposed in the current paper is unparalleled in the scientific literature, and provides a solid basis for the assessment of the technological impact of scientific research in countries/regions and institutions.

Keywords Technological Impact; Patents; Scientometric; PATSTAT; Scopus



[†] Corresponding author: Vicente P. Guerrero-Bote (E-mail: guerrero@unex.es).

1 Introduction

One of the most important topics in quantitative science and technology studies is the relationship between the two fields. Citation analysis is a key to gaining insight into this relationship. Citations in scientific papers to other scientific papers have long been used to assess the contribution made by individual scholars and research institutions to *scientific* progress, as well as to analyse the cognitive structure of science and the collaboration and flows of knowledge among scientific workers (Price, 1963). Citations made in patents to the scientific literature may similarly be used to study the influence of scientific and scholarly work on *technological* development. Pioneering work on patent citation analysis was conducted by Francis Narin and co-workers, who started exploring the use of patent citations to measure technological impact in the 1980s (Narin & Olivastro, 1992; Narin, Hamilton, & Olivastro, 1997). This is the perspective taken in the present communication. Just as a citation from the scientific literature is used to determine the *scientific* impact of a published scientific article, a citation from a patent can be used to determine an article's *technological* impact. The relationship between scientific citation and patent citation is an ongoing subject of study (Ahmadpoor & Jones, 2017; Poege et al., 2019; Veugelers & Wang, 2019), although some concerns have been raised about its real relevance (Bryan, Ozcan, & Sampat, 2020; Callaert, Pellens, & Van Looy, 2014). Nevertheless, patent analysis to construct science and technology indicators has become an active and complex discipline in its own right. Comprehensive overviews of the main methodological issues involved are given in Hinze and Schmoch (2004), Van Raan (2017), and more recently in Schmoch and Kahn (2019) and Veugelers and Wang (2019).

Technological research and development come at important costs. For this reason, when the development of a product has been completed successfully, it needs to be patented to obtain the protection that can allow its economic exploitation. A patent is a document describing a product, with the pertinent bibliography and citations to other patents and other documents. One part of the bibliography is provided by the authors and another part by the patent office's examiners. The evaluation procedure varies significantly from one national patent office to another, and consequently so may the bibliography, both the part required from the authors and the part incorporated by the examiners.

A patent family can be defined as a set of patents filed in various countries to protect a single invention. The protection in the country in which the first application is made—the priority country—is then extended to other countries. In other words, a patent family relates to the same invention disclosed by one or more common inventors and patented in more than one country. A *first* major problem in patent



citation analysis addressed in the present communication is that there may be substantial differences between members of the same patent family regarding the non-patent literature (NPL) references they may contain. This is especially the case for the examiner-incorporated references since the patents of a family tend to pass different examination processes in the different patent offices. The process may be slower or faster, and may lead to more or fewer NPL references being incorporated.

The process of registering a patent to protect an invention is neither easy nor cheap. A *second* problem addressed in this communication is that the protection of an invention in a technologically advanced country will give it a greater competitive edge and a greater potential market than in a developing country. Seeking protection in advanced countries with larger markets is harder and more expensive than it is in developing countries with small markets because greater benefits are expected from the invention. Therefore, a triadic patent—i.e. a set of corresponding patents filed in the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO), and the Japan Patent Office (JPO) for the same invention with the same applicant or inventor—tends to have more economic and technological significance than a patent family covering one or more developing countries/regions.

A *third* problem is that there appear to exist large differences in the number of citations to scientific papers among patents and patent families. In a citation analysis of the scientific literature, advanced citation impact indicators correct for differences in the length of reference lists in scientific papers in a subject field. A similar “citing or source” normalization would be needed in an analysis of patent citations to take different patent families’ propensity to cite scientific literature into account.

The aim of this study was to explore new indicators of the technological impact of recent research papers—indicators based on citations to the scientific literature that are found in patents, and taking into account the three problems outlined above:

1. To avoid differences in references to the scientific literature among patents from the same patent family, the patent family rather than an individual patent is taken as the basic unit on the “citing side” of the citation analysis. The SCImago Research Group has retrospectively assigned *all* non-patent references in the various members of a patent family to *each* patent in that family.
2. To differentiate among patent families according to the country/region in which patent protection was sought, a weight is assigned to each country/region based on that country/region’s state of economic development. A citation from a triadic patent then has greater weight than a patent citation from an economically developing country/region. In a country/region with greater economic and technological activity, it is more difficult and expensive



to obtain a patent. When an applicant selects a country/region of this type it is because the patent involves state-of-the-art technology and is expected to yield significant economic benefits that will compensate for the investment.

3. The reciprocal of the number of references to scientific papers in a patent family is used as a weighting factor normalizing the number of patent family citations received by a scientific paper, thus correcting for differences among patent families in their propensity to cite the scientific literature. In this way, each patent family distributes its recognition (which is already weighted by the state of economic development of the countries/regions in which its patents are applying for protection) equitably among all of its references to scientific papers.

The newly proposed indicators of technological impact are applicable to the scientific production of countries, institutions, research groups, or even individual authors.

2 Data

Elsevier created and maintains Scopus, a bibliographic database of scientific literature (Hane, 2004; Pickering, 2004). It indexes more than 31,000 scientific journals, conference proceedings, and books. Several works have described its characteristics (Archambault et al., 2009; Leydesdorff, Moya Anegón, & Guerrero-Bote, 2010; Moya-Anegón et al., 2007), and it has been used in numerous scientometric studies (Gorraiz, Gumpenberger, & Wieland, 2011; Guerrero-Bote & Moya-Anegón, 2015; Jacsó, 2011; Moya-Anegón et al., 2018). Scopus classifies publications by Subject Area, and then by Specific Subject Area or Category. There are more than 300 categories grouped under 26 subject areas. In addition, there is a Multidisciplinary Area that contains such journals as Nature or Science. The present study covers documents indexed in Scopus up to 2018 and published after 2002.

In 2008, at the behest of the patent statistics working group led by the Organization for Economic Co-operation and Development (OECD), the EPO released PATSTAT (“EPO Worldwide PATent STATistical Database”), designed to assist in statistical research on patents. There are other patent databases, such as NBER (the United States) and IIP (Japan). But PATSTAT has worldwide coverage, includes more information, and has some auxiliary products that can resolve various problems. For these reasons, PATSTAT has become a de facto standard (Kang & Tarasconi, 2016). Nevertheless, it has two biases. One is towards European data since national patent offices exchange information with the EPO based on formal agreements (although these agreements change over time and leave some gaps). The other is towards the



patent examination process since data that are not actually required in this process may be of lower quality. For the present study, we used the 2019 Spring edition.

PATSTAT collects references both to NPL and to patents. Due to its patent application and examination orientation, its data require debugging and standardizing to allow linkage with data taken from other databases. For example, the lack of standardization has long been a cause of problems in relating applicants and inventors with the data available in company databases. The first attempts to normalize names were made with the standardization tables of Thomson Scientific's Derwent World Patent Index (2007) and the USPTO's CONAME file, and further development and improvements are continuing to be made (Coffano & Tarasconi, 2014; Lissoni, 2012; Lotti & Marin, 2013; Magerman, Van Looy, & Song, 2006; Maraut & Martínez, 2014; Raffo & Lhuillery, 2009; Thoma & Torrisi, 2007; Schoen, Heinisch, & Buenstorf, 2014).

In order to link the NPL citations in patents with scientific publications, one faces the same problem as in the case of the names of the applicants or inventors—namely, the absence of standardization. We only know of one publicly available publication that has taken on this issue. It was part of the development of *Lens Influence Mapping* (Jefferson et al., 2018), and used PubMed and Crossref as the scientific literature databases. However, this study indicated neither how the cases in which a patent citation could be linked to more than one DOI (Digital Object Identifier) were resolved, nor how the method ensured that the retrieved documents actually corresponded to the patent-cited references they were linked with.

The SCImago Research Group has developed a procedure for linking the NPL references of the patents indexed in the PATSTAT database to the scientific papers indexed in the Scopus bibliographic database. This procedure has been implemented with reasonable results and assumable costs (Guerrero-Bote, Sánchez-Jiménez, & Moya-Anegón, 2019; Moya-Anegón et al., 2020). It can be summarized as follows:

1. Data pre-processing: Preparation of the data to facilitate and streamline the subsequent processes.
2. Pre-selection of candidate pairs: Some coincidences of the elements of NPL references in PATSTAT and publication records in Scopus are used to pre-select candidates for a matching pair.
3. Automated evaluation of candidate pairs: The matching elements of each candidate pair are evaluated, and the pair is assigned a score. Thus, for each NPL reference, an ordered list is obtained of the Scopus references that could match it.
4. Human validation: For each NPL reference, the top-ranked Scopus reference is checked manually, ensuring that they really do match.



Since human validation plays a key role in the matching process, the precision was clearly going to be very high, but it still cannot be claimed to be 100%. In the spring 2019 version of PATSTAT (the one used for this study), patent applications from 2003 or later include 54 million NPL references, of which 20 million are to documents dated 2003 or later, and, of these, 7.4 million have been linked following the above procedure. Few of the remaining 12.6 million references are scientific papers included in Scopus: from a manual check of a sample stratified by year of 600 of these remaining references, we estimated that only 0.37 million of them are actually documents included in Scopus, which would represent a recall of about 95%.

3 Methods

In general terms, a patent family can be said to be a group of related patents. But there are different specific definitions of the concept (Martínez, 2011). The EPO (European Patent Office, 2019), for instance, uses two definitions:

- Simple family, DOCDB family, or Espacenet patent family: All applications of the family have the same priorities.
- Extended family or INPADOC family: All applications of the family are linked to the same root priority application. There may be applications in the same family that have no priority in common, but do have one in common with a third application which is the one that unites them.

In the present work, for convenience our focus is on the concept of a Simple family, which groups together all the applications corresponding to the same invention (European Patent Office, 2019).

As our goal was to develop indicators of the technological impact of recent scientific literature, we decided to take all patent applications into account, regardless of whether or not they had been granted. The reason was that if an applicant had taken the necessary steps to start the procedure then it was because the inventor and/or applicant considered the invention to be commercially exploitable, regardless of whether for legal (e.g. the application was late) or other reasons the patent would ultimately be granted. It should also be noted that many patents are not exploited at all or are exploited only very briefly, which makes it hard to determine their value.

Just as not all scientific papers are equally important, not all patent applications have the same weight. Although there have been many studies on indicators of patent value, such as those relating to patent citations or renewals, these indicators require time to manifest themselves, i.e. time to be cited or time to be renewed. On top of this, there is the time necessary for a patent application to be published. This



is why we elected to weight patent applications according to the countries/regions in which protection was being sought.

In order to differentiate among patent families according to this criterion, we decided to assign a weight to each country/region. We explored three candidate indicators with which to generate this weight: the number of patent applications in the country/region, the country/region's GDP, and its investment in R&D. The first (the number of patent applications in a country/region) was discarded because of the existence of very small countries/regions with a small market but a disproportionate number of patents requesting protection, this being a consequence of regional offices and treaties such as the Eurasian Patent Organization (EAPO), the European Patent Office (EPO), and the European Patent Convention (EPC). The third (investment in R&D) was ruled out due to the difficulty of getting reliable annual data and because there is no direct relationship between the amount of R&D investment and patent applications. We therefore chose as the best option a weighting factor based on a country/region's GDP as an indicator of the commercial potential that was to be expected for that country/region. Specifically, a weight was defined for each country/region and for each particular year that was equal to its share of World GDP. The weight associated with a patent application is then determined by the country/region in which protection is requested in the filing year.

To weight citations from a patent family, an alternative approach would have been to use the absolute number of patents in the family or the number of designated countries/regions as the weighting factor. But such an approach would not take into account the state of economic development of the designated countries/regions. These criteria were therefore not used in the definition of the indicators that we finally explored.

While the NPL references may have different origins (applicant, examiner, opponent, etc.), no such differentiation was made in the present work. As will be outlined further below, both the number and the origin of NPL references are influenced by each national office's evaluation procedure, and this varies significantly from one country/region to another. Since patent families constitute the source or citing side of the citation process, duplicate references to scientific papers within each patent family's applications were merged into one. Also, to allow for differences among patents in the number of references to scientific papers, the aforementioned country/region weighting factor was divided by the total number of papers cited in each patent family.

For instance, for a patent family which makes r citations to scientific papers, and which designates countries/regions c_1 and c_2 which have shares of World GDP equal to $GDP.c_1$ and $GDP.c_2$, respectively, the weighting factors w_1 and w_2 for those countries/regions c_1 and c_2 are calculated as follows:



$$w_1 = \frac{GDP.c_1}{r}; w_2 = \frac{GDP.c_2}{r}.$$

In the tables below, a weighting factor based solely on a country/region's share of World GDP will be denoted as a GDP-related weight, and a weighting factor resulting from dividing this share by the total number of a patent family's unique citations to papers as a fractional GDP-related weight.

In Table 1, patent families constitute the source or citing side of the citation process, and their citations to scientific papers are referred to as *cites* in the column headings. Articles published in the scientific literature constitute the target or cited side of the process, and are referred to as *papers* in the column headings. The countries/regions in the 1st column are the 28 designated nations of the patent applications with the greatest weighted number of citations to scientific papers (see the last column). These 28 countries/regions together account for 98% of the total weighted number of citations to scientific papers.

One observes in Table 1 that 5,400,857 patent families have the United States as a designated country, although of course they may designate other countries/regions as well. These families contain 3,533,076 citations to scientific papers indexed in Scopus, so that the average number of citations to papers per patent family is 0.65. Only a small percentage of patent families contain citations to papers, so that when only these families (those with at least one citation to papers) are considered then the average number of citations to papers per patent family rises to 33.29.

As mentioned above, for each patent family designating a given country/region, a weighting factor is calculated which is defined as the ratio between the designated country/region's share of World GDP and the patent family's total number of citations to papers. For example, calculating this ratio for all families designating the United States and having at least one citation to papers for each year considered, and computing the average over those years, one obtains a value of 0.234. This number can be interpreted as the "value" of a citation in a patent family designating the United States to an arbitrary scientific paper. It therefore relates to the *citing* side of the citation process, not to the *cited* side. It will be used below as a type of *source normalization factor* in calculating a scientific paper's technological impact.

The "percentage weighted cites to papers" in the last column of Table 1 is calculated in the following manner. First, for each designation country/region the weighted number of citations to papers is calculated by multiplying the number of citations in patent families designating that country/region by its normalization factor. Next, these weighted numbers of citations are summed over all designation countries/regions. Finally, for each country/region, the share of weighted citations is computed relative to the sum over all countries/regions, expressed as a percentage.

Although it is unsurprising that the United States is the country with the largest share of weighted citations, it has fewer patent families than China and only slightly





Table 1. Patent citation based indicators by designated country/region.

Designated country/ region	# Patent Families	Families with cites to papers (%)	# Pat Fam cites to papers	Avg cites to papers per Pat Fam	Weighted avg GDP share	Avg cites to papers per Pat Fam with ≥ 1 cites	Pat Fam with filing year ≥ 2003 (%)	Pat Fam cites to papers (%)	Weighted cites to papers (%)
United States	5400857	11.65	3533076	0.65	0.234	33.29	22.47	81.15	45.91
China	8571905	4.24	1463274	0.17	0.108	33.12	35.66	33.61	12.50
Germany	2408775	15.79	2158516	0.90	0.052	35.02	10.02	49.58	6.18
Japan	5096385	4.65	1410294	0.28	0.079	38.61	21.20	32.39	5.88
United Kingdom	2037468	18.02	2136675	1.05	0.041	35.47	8.48	49.08	4.72
France	1887911	19.15	2105069	1.12	0.040	35.66	7.85	48.35	4.47
Italy	1832796	19.28	2085714	1.14	0.031	35.94	7.62	47.91	3.49
Spain	1762192	20.08	2086719	1.18	0.021	35.92	7.33	47.93	2.28
Netherlands	1753318	20.02	2078575	1.19	0.013	36.03	7.29	47.74	1.37
Turkey	1742792	20.05	2074734	1.19	0.011	36.08	7.25	47.66	1.22
Canada	513166	25.89	1111076	2.17	0.024	47.31	2.13	25.52	0.96
Switzerland	1745305	20.05	2076016	1.19	0.009	36.06	7.26	47.69	0.95
Sweden	1748215	20.01	2075724	1.19	0.008	36.07	7.27	47.68	0.82
Belgium	1732148	20.18	2075083	1.20	0.007	36.08	7.21	47.66	0.78
South Korea	2348083	5.96	815412	0.35	0.018	41.86	9.77	18.73	0.78
Poland	1707430	20.20	2059367	1.21	0.007	36.21	7.10	47.30	0.74
Austria	1742992	20.07	2075891	1.19	0.006	36.07	7.25	47.68	0.65
Norway	1264904	22.10	1772202	1.40	0.006	38.36	5.26	40.71	0.55
Denmark	1740152	20.09	2076218	1.19	0.005	36.08	7.24	47.69	0.52
Brazil	296790	20.85	478260	1.61	0.029	47.00	1.23	10.99	0.52
Australia	446278	21.88	857725	1.92	0.018	49.69	1.86	19.70	0.52
Greece	1731130	20.19	2074796	1.20	0.004	36.08	7.20	47.66	0.45
Finland	1741247	20.10	2075883	1.19	0.004	36.07	7.24	47.66	0.41
Ireland	1730187	20.20	2074865	1.20	0.004	36.08	7.20	47.66	0.41
Portugal	1728495	20.23	2076124	1.20	0.003	36.08	7.19	47.69	0.37
Czech Republic	1734030	20.18	2075923	1.20	0.003	36.06	7.21	47.68	0.32
Russian Federation	534368	7.92	275369	0.52	0.025	43.57	2.22	6.33	0.31
Romania	1733753	20.15	2074321	1.20	0.003	36.09	7.21	47.65	0.27

Table 1. Continued

Legend to Table 1:

Designated country/region: The country/region designated in a patent application.

Patent Families: The number of PATSTAT patent families applying for protection in the country/region.

Families with cites to papers (%): Percentage of patent families that include citations to scientific papers indexed in Scopus.

Pat Fam cites to papers: The number of citations in patent families to scientific papers indexed in Scopus.

Avg cites to papers per Pat Fam: The average number of citations to scientific papers indexed in Scopus per patent family (the ratio between the previous two numbers).

Weighted avg GDP share: The country/region's time-averaged share of World GDP calculated by weighting the share in each year by its number of patent applications in that year.

Avg cites to papers per Pat Fam with ≥1 cites: The average number of citations to scientific papers indexed in Scopus in patent families that make at least one such citation.

Pat Fam with filing year ≥2003 (%): Percentage of patent families with filing year 2003 or later relative to the total number of patent families.

Pat Fam cites to papers (%): Percentage of patent family citations to scientific papers indexed in Scopus relative to the total number of patent family citations to papers.

Weighted cites to papers (%): Weighted citations to papers (expressed as a percentage of the total over all countries/regions).



Research Paper

more than Japan. The main difference stems from the absolute number of patent citations to scientific papers, as can be seen in the 4th column. Specifically, in China and in Japan there must be many patents that do not cite any scientific paper indexed in Scopus.

In the 8th and 9th columns, one can clearly see the differences between a country/region's percentage of patent families and its percentage of patent family citations. Keep in mind that there are many patent families that apply for protection in multiple countries/regions, which explains why these percentages sum to more than 100%.

One also observes that there are differences between countries/regions in the average number of patent citations to papers. These reflect the differences among the patent authorities regarding their requirements and review processes. It can be seen that while the United States accounts for almost 45% of the weighted cites to papers, only three other countries/regions exceed 5%.

Table 2. Patent citation according to document type of the cited paper.

Document type	# Patent families	Avg cites to papers	Avg refs in patents	Avg cites to papers (GDP-based weighting)	Avg cites to papers (fractional GDP-based weighting)
Article	670831	0.008	38.9	0.020	0.0023
Conference Paper	340577	0.018	17.4	0.025	0.0048
Review	177172	0.012	47.1	0.018	0.0019
Short Survey	24580	0.020	48.1	0.018	0.0013

Legend to Table 2:

Patent families: Number of citing patent families.

Avg cites to papers: Average number of patent families citing any given document of that type.

Avg refs in patents: Average number of citations to papers in patent families which cite that document type.

Avg cites to papers (GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP.

Avg cites to papers (fractional GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP divided by the total number of a patent family's unique citations to papers.

Table 2 presents the data listed by the four "citable" document types: Article, Conference Paper, Review, and Short Survey. It shows that the most cited document type is Article followed by Conference Paper. However, although the Article is the most frequently occurring document type, if one considers the average number of citing patent families per paper (3rd column), the Short Survey ranks first, followed by the Conference Paper, the Review, and the Article. The 4th column lists the average number of citations to scientific papers indexed in Scopus that are included in patents that cite a particular document type. The Conference Paper receives on average less than half the number of citations received by the other types. The 5th column gives the average number of citations weighted with the share of global GDP of designated countries/regions in patent families that cite a particular type of document, and the 6th column the average number of citations weighted with the



fractional weight, as described above, defined as the ratio between a country/region's share of global GDP and the total number of citations to papers in a citing patent family.

It is striking that Conference Papers have the greatest average fractional weight. This can be explained by the fact that in fields like computer science conference papers have greater relevance in combination with the observation (obtained from a secondary analysis not presented in the present communication) that the United States is the country with the greatest percentage of citations to conference papers as well as the greatest GDP-related weight of its patent families.

Table 3 is similar to Table 2, but relates to the cited papers' publication year rather than document type. It shows declines in the number of patent families, the average number of citations weighted with both the GDP share and the fractional weight. The more recent the paper, the less time it has had to be cited, as is true also for scientific paper-to-paper citations. The average number of citations to papers in the citing patents grew until 2013. This was to be expected since old patents have a smaller time frame from which to pick papers to cite (only post-2003 papers were considered) than that of more recently published patents.

Table 3. Patent citation according to publication year of the cited paper.

Year	# Patent families	Avg cites to papers	Avg refs in patents	Avg cites to papers (GDP-based weighting)	Avg cites to papers (fractional GDP-based weighting)
2003	230308	0.019	31.49	0.0218	0.00400
2004	232905	0.019	33.77	0.0212	0.00350
2005	229644	0.019	35.01	0.0208	0.00318
2006	214822	0.019	35.43	0.0203	0.00292
2007	205747	0.019	37.53	0.0202	0.00278
2008	192682	0.019	39.03	0.0201	0.00261
2009	177198	0.019	39.74	0.0195	0.00255
2010	160483	0.019	40.16	0.0193	0.00249
2011	140461	0.019	41.07	0.0191	0.00244
2012	116753	0.018	42.90	0.0189	0.00233
2013	87978	0.016	46.07	0.0188	0.00207
2014	61586	0.014	42.73	0.0186	0.00178
2015	38225	0.011	39.68	0.0185	0.00134
2016	18082	0.006	42.64	0.0189	0.00064
2017	6046	0.002	42.45	0.0202	0.00022
2018	806	0.000	38.90	0.0191	0.00003

Legend to Table 3:

Patent families: Number of citing patent families.

Avg cites to papers: Average number of patent families citing a document type.

Avg refs in patents: Average number of citations to papers in patent families citing a document type.

Avg cites to papers (GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP.

Avg cites to papers (fractional GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP divided by the total number of a patent family's unique citations to papers.



Research Paper

Table 4 is similar to Table 3, but gives results by Scientific Area. As expected, there are large differences between areas.

Table 4. Patent citation according to Scientific Area of the cited paper.

Scientific Area	# Patent families	Avg cites to papers	Avg refs in patents	Avg cites to papers (GDP-based weighting)	Avg cites to papers (fractional GDP-based weighting)
Computer Science	251649	0.019	16.10	0.025	0.00525
Engineering	394749	0.018	21.55	0.024	0.00466
Materials Science	228670	0.018	23.88	0.024	0.00428
Physics and Astronomy	204802	0.017	24.03	0.025	0.00401
Energy	56034	0.020	15.61	0.022	0.00395
Mathematics	111652	0.021	18.10	0.024	0.00386
Chemistry	241128	0.015	27.28	0.020	0.00350
Chemical Engineering	154459	0.017	33.27	0.020	0.00309
Decision Sciences	11859	0.021	13.79	0.028	0.00280
Health Professions	21004	0.017	29.71	0.021	0.00247
Environmental Science	51047	0.015	24.88	0.020	0.00230
Pharmacology, Toxicology and Pharmaceutics	120068	0.011	40.29	0.018	0.00227
Agricultural and Biological Sciences	78973	0.012	34.36	0.018	0.00193
Dentistry	4080	0.011	28.90	0.019	0.00192
Biochemistry, Genetics and Molecular Biology	271492	0.007	45.62	0.018	0.00187
Earth and Planetary Sciences	19357	0.009	15.67	0.022	0.00172
Medicine	270703	0.006	46.07	0.018	0.00164
Multidisciplinary	67184	0.017	55.33	0.019	0.00161
Immunology and Microbiology	94414	0.009	46.12	0.018	0.00160
Business, Management and Accounting	10229	0.010	11.89	0.030	0.00154
Veterinary	10619	0.012	32.96	0.018	0.00153
Neuroscience	43846	0.011	55.51	0.019	0.00141
Nursing	13688	0.011	54.48	0.018	0.00132
Arts and Humanities	37896	0.015	50.43	0.021	0.00128
Social Sciences	22800	0.007	19.65	0.025	0.00100
Economics, Econometrics and Finance	2044	0.004	11.76	0.034	0.00062
Psychology	6026	0.006	63.15	0.021	0.00059

Legend to Table 4:

Patent families: Number of citing patent families.

Avg cites to papers: Average number of patent families citing a document type.

Avg refs in patents: Average number of citations to papers in patent families citing a document type.

Avg cites to papers (GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP.

Avg cites to papers (fractional GDP-based weighting): Average number of citations weighted by a factor defined as a country/region's share of global GDP divided by the total number of a patent family's unique citations to papers.

A general conclusion to be drawn from Tables 1–4 is that a large part of the scientific papers receive no citations from patents, and that patent citations are very unevenly distributed among the scientific areas, document types, and cited publication years.

4 Proposed indicator for countries/regions and other aggregates

In developing a proposal for an indicator, we were inspired by the item-oriented, field-normalized, citation-score average indicator introduced by Lundberg (2007) which formed the base for the CWTS' mean normalized citation score (Waltman et al., 2011) and SCImago's normalized impact (SCImago Research Group, 2010). In this case however, we did not assume a priori that scientific contributions in all subject fields have the same technological value. It is clearly plausible that a contribution in Engineering or Pharmacology would have a greater technological impact than a contribution in the Humanities. We therefore decided not to implement any subject field normalization. Similar reasoning led us to also reject normalization according to document type.

Hence, we defined the Technological Impact of an aggregate as

$$TI = \frac{1}{n} \sum_{i=1}^n \frac{a_i}{e_i}; a_i = \sum_{j=1}^p w_{ij}$$

where n is the number of (cited or not) papers published by the aggregate (e.g. an institution, a journal, a country/region), a_i is the sum of the weights of the patent citations received by paper i , e_i is the expected value of the sum of weights for papers of the same age, p is the number of patent families citing the i -th paper (p depends on i), w_{ij} is the weight of the citation from patent family j received by the i -th paper as defined in Section 3:

$$w_{ij} = \frac{GDP \cdot c}{r}$$

in which GDP represents the World GDP, c a designated country/region's share of the World GDP, and r the number of citations in a patent family to scientific papers indexed in Scopus. The indicator TI is size-independent and, as a measure, corresponds to the average citation obtained by the aggregate's papers.

As an analogue of the total normalized citation score indicator known as "Brute Force", we define Technological Force as:

$$TFO = n \cdot TI = \frac{\sum_{i=1}^n a_i}{e_i}.$$

This indicator is size-dependent, and is a measure of the total technological contribution of the aggregate.



5 Results

In analysing the behaviour of the indicators, only results for the main countries/regions will be presented. We shall begin by focusing on *TFO* as a size-dependent indicator, and compare it with other size-dependent measures of a country/region's performance such as the number of patent families that have applied for protection in the country/region, its scientific publication output, a "brute force" indicator defined as the product of the publication output and the field-normalized citation impact, its scientific excellence output (Top 10% and Top 1%), and total citations from patent families to its scientific output. To make these size-dependent indicators more readily comparable, they will be expressed as percentages with respect to the global total. These percentages may sum to more than 100% because there are families that apply for protection in more than one country/region, and papers may be published in collaboration among several countries/regions.

Table 5 presents the scores of the 28 countries/regions with the highest value of *TFO*. In all indicators, the United States and China rank top, although not in the same proportions. In the 2nd column, which gives the percentage of families applying for protection in a country/region, one notes that China ranks higher than the United States. Also visible is the effect of the EPC. While in the past a designation fee was paid that depended on the designated countries/regions, since 2007 all European patent applications pay a single fee and designate all countries/regions. This causes many patents to seek protection in a large number of European Union countries, with the result being that around 7% of all patent families apply for protection in almost all countries of the EU.

One observes in the table that China has a smaller percentage of patent family citations than would have been expected given its total publication output. This is because patents that apply for protection in China are characterized by making few citations to papers, and those that cite papers from China itself even fewer. Nonetheless, these patents have an above-average weight precisely due to this low number of references and because China's GDP is the world's second largest, meaning that its cited references also contribute more, leading to a *TFO* within the expected range.

Figure 1 shows the same data as presented in Table 5 for the 12 countries/regions with the highest values of Technological Force (We plot just 12 countries/regions so that the profiles can be clearly seen.). The figure shows a large number of families applying for protection in China, Japan, and South Korea. Also, the EPC effect can be seen, with the other EU countries following in the wake of Germany and the United Kingdom in the number of patent families applying for protection. It is notable that India has a different profile, with relatively few patent families applying for protection in its territory.



Table 5. Size-dependent indicators of the 28 top-ranked %TFO countries/regions.

Country/Region	%Families	%Output	%BF	%Exc10	%Exc1	%Fam.Cit.	%TFO
United States	22.47	24.25	36.37	37.76	46.45	43.74	36.88
China	35.66	14.05	11.48	13.09	13.58	6.68	14.32
Germany	10.02	6.13	8.32	8.71	10.05	7.93	8.31
Japan	21.20	5.21	4.80	4.59	4.10	7.64	7.83
United Kingdom	8.48	7.05	10.85	11.23	14.03	7.19	6.50
France	7.85	4.30	5.57	5.86	6.56	4.59	4.74
South Korea	9.77	2.50	2.53	2.71	2.65	3.51	4.56
Canada	2.13	3.66	5.50	5.76	7.16	4.25	4.40
Italy	7.62	3.65	5.01	5.26	5.82	3.27	3.29
India	0.23	3.80	3.06	2.89	2.52	1.53	2.74
Netherlands	7.29	2.02	3.55	3.90	5.22	2.64	2.56
Spain	7.33	2.97	3.71	3.98	4.41	2.19	2.49
Switzerland	7.26	1.50	2.70	2.91	4.11	2.48	2.48
Australia	1.86	2.97	4.52	4.83	6.05	2.22	2.36
Chinese Taiwan	2.74	1.44	1.43	1.58	1.34	1.59	2.19
Sweden	7.27	1.33	2.16	2.31	2.93	1.70	1.65
Belgium	7.21	1.12	1.81	1.94	2.58	1.47	1.48
Singapore	0.31	0.66	1.11	1.25	1.79	0.97	1.33
Denmark	7.24	0.83	1.49	1.59	2.18	1.03	1.12
Austria	7.25	0.82	1.23	1.27	1.63	1.05	1.10
Israel	0.32	0.75	1.10	1.15	1.41	1.18	1.03
Finland	7.24	0.69	1.06	1.10	1.33	0.83	1.03
Chinese Hong Kong	0.30	0.69	1.13	1.31	1.71	0.79	1.00
Brazil	1.23	2.14	1.84	1.66	1.55	0.75	0.92
Russian Federation	2.22	2.12	1.36	1.06	1.04	0.60	0.78
Poland	7.10	1.41	1.23	1.07	1.15	0.59	0.76
Norway	5.26	0.67	1.08	1.08	1.37	0.51	0.72
Iran	0.00	1.28	1.16	1.28	1.17	0.33	0.65

Legend to Table 5:

All the percentages are relative to the global total.

%Families: Percentage of PATSTAT patent families with filing year 2003 or later applying for protection in the country/region.

%Output: Percentage of scientific papers indexed in Scopus.

%BF: Brute Force percentage, i.e., %Output \times Normalized Impact (NI).

%Exc10: Percentage of scientific papers of excellence (Top 10%).

%Exc1: Percentage of scientific papers of excellence (Top 1%).

%Fam.Cit.: Percentage of patent family citations received.

%TFO: Percentage Technological Force.

Table 5 reveals Australia's low position in terms of the percentage of patent families citing its papers, and the EPC effect in Sweden and Belgium in this sense. Above all, one notes the high position of South Korea in the table (7th), while it ranks just 14th in terms of scientific output.

Table 6 gives the Pearson correlation coefficients between the size-dependent indicators of the 40 countries/regions with the greatest scientific outputs. The indicator that correlates least well with the others is *Families*, that corresponding to patent families applying for protection. There is a strong correlation between the indicators that combine quantity and quality of scientific production (*BF*, *Exc10*,



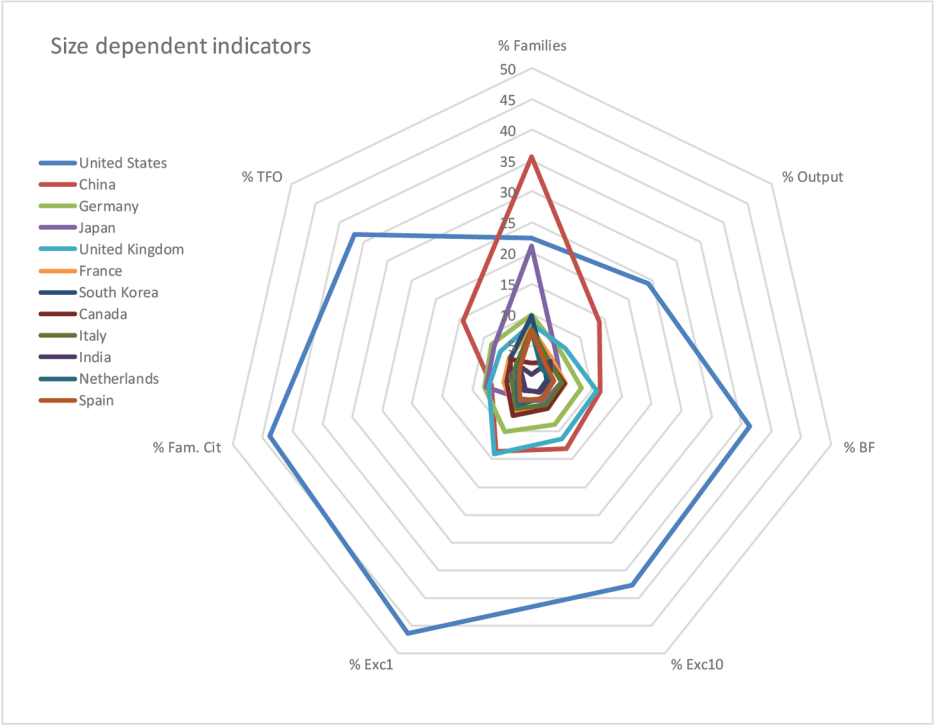


Figure 1. Size-dependent indicators of the 12 countries with greatest Technical Force. Percentages relative to the global total.
% Families: Percentage of PATSTAT patent families with filing year 2003 or later applying for protection in the country/region.
% Output: Percentage of scientific papers indexed in Scopus.
% BF: Brute Force percentage, i.e., %Output \times Normalized Impact (NI).
% Exc10: Percentage of scientific papers of excellence (Top 10%).
% Exc1: Percentage of scientific papers of excellence (Top 1%).
% Fam. Cit: Percentage of patent family citations received.
% TFO: Percentage of Technological Force.

Table 6. Pearson correlation matrix between the size-dependent indicators for the 40 countries/regions with the greatest scientific output.

	%Families	%Output	%BF	%Exc10	%Exc1	%Fam.Cit.	%TFO
%Families	1	0.72	0.60	0.61	0.58	0.55	0.68
%Output	0.72	1	0.97	0.97	0.95	0.92	0.987
%BF	0.60	0.97	1	0.99	0.99	0.98	0.98
%Exc10	0.61	0.97	0.99	1	0.99	0.97	0.98
%Exc1	0.58	0.95	0.99	0.99	1	0.97	0.97
%Fam.Cit.	0.54	0.92	0.98	0.97	0.97	1	0.97
%TFO	0.68	0.98	0.98	0.98	0.97	0.97	1



and *ExcI*). Technological Force has a stronger correlation with all the other indicators than the direct values of patent family citations.

The analyses presented in Tables 7 and 8 are analogous to those of Tables 5 and 6, but based on size-independent rather than size-dependent indicators. The scores in these tables are relative to the global average, not the actual values of the indicators. The same set of 28 countries/regions are considered, i.e., those with the largest Technological Force (*TFO*), and they listed in Table 7 in the same order as in Table 5. But the actual ranking based on Technological Impact (*TI*) differs strongly from that based on *TFO*. Indeed, the top 10 countries/regions ranked by *TI* would be Singapore (1.97), South Korea (1.67), Switzerland (1.62), United States (1.55), Finland (1.44), Chinese Hong Kong (1.41), Chinese Taiwan (1.40), Japan (1.40) Israel (1.35), and Denmark (1.33).

Table 7. Size-independent indicators of the 28 top-ranked %TFO countries/regions.

Country/Region	Fam.Rel.	%Q1	NI	%Exc10	%Exc1	Avg.Fam.Cit.	TI
United States	2.99	1.38	1.50	1.68	2.24	1.98	1.55
China	2.97	0.83	0.82	1.00	1.13	0.46	0.92
Germany	7.14	1.24	1.36	1.53	1.92	1.35	1.32
Japan	5.23	1.09	0.92	0.95	0.92	1.47	1.40
United Kingdom	6.00	1.37	1.54	1.72	2.33	1.17	0.98
France	9.69	1.26	1.30	1.47	1.79	1.11	1.06
South Korea	6.44	1.13	1.01	1.17	1.24	1.38	1.67
Canada	4.17	1.39	1.50	1.69	2.29	1.25	1.20
Italy	11.14	1.25	1.37	1.55	1.87	0.95	0.88
India	0.18	0.70	0.81	0.82	0.78	0.41	0.69
Netherlands	20.03	1.55	1.76	2.08	3.04	1.40	1.26
Spain	13.70	1.24	1.25	1.44	1.74	0.77	0.81
Switzerland	26.89	1.49	1.81	2.09	3.21	1.75	1.62
Australia	3.78	1.38	1.52	1.75	2.39	0.81	0.80
Chinese Taiwan	5.22	1.20	0.99	1.18	1.08	1.09	1.40
Sweden	30.30	1.52	1.63	1.87	2.58	1.34	1.20
Belgium	36.04	1.42	1.62	1.87	2.71	1.39	1.30
Singapore	4.02	1.34	1.69	2.06	3.20	1.54	1.97
Denmark	48.57	1.54	1.80	2.07	3.09	1.32	1.33
Austria	48.83	1.27	1.50	1.66	2.32	1.35	1.30
Israel	4.36	1.46	1.47	1.66	2.20	1.67	1.35
Finland	58.69	1.38	1.54	1.73	2.27	1.25	1.44
Chinese Hong Kong	3.53	1.39	1.64	2.06	2.92	1.20	1.41
Brazil	3.33	0.88	0.86	0.84	0.85	0.36	0.40
Russian Fed.	2.30	0.49	0.64	0.54	0.58	0.28	0.33
Poland	28.16	0.79	0.87	0.82	0.95	0.42	0.50
Norway	47.68	1.37	1.60	1.73	2.37	0.81	1.07
Iran	0.00	0.74	0.91	1.08	1.07	0.25	0.47

Legend to Table 7.

All indicators are relative to the global average, so that a value of unity would correspond to that global average.

Fam.Rel.: Ratio between the number of patent families applying for protection in the country/region and its number of scientific papers. **%Q1:** Percentage of scientific papers in the first quartile.

NI: Normalized Impact.

%Exc10: % Excellence 10 (Top 10%).

%Exc1: % Excellence 1 (Top 1%).

Avg.Fam.Cit.: Average number of patent family citations per paper.

TI: Technological Impact.



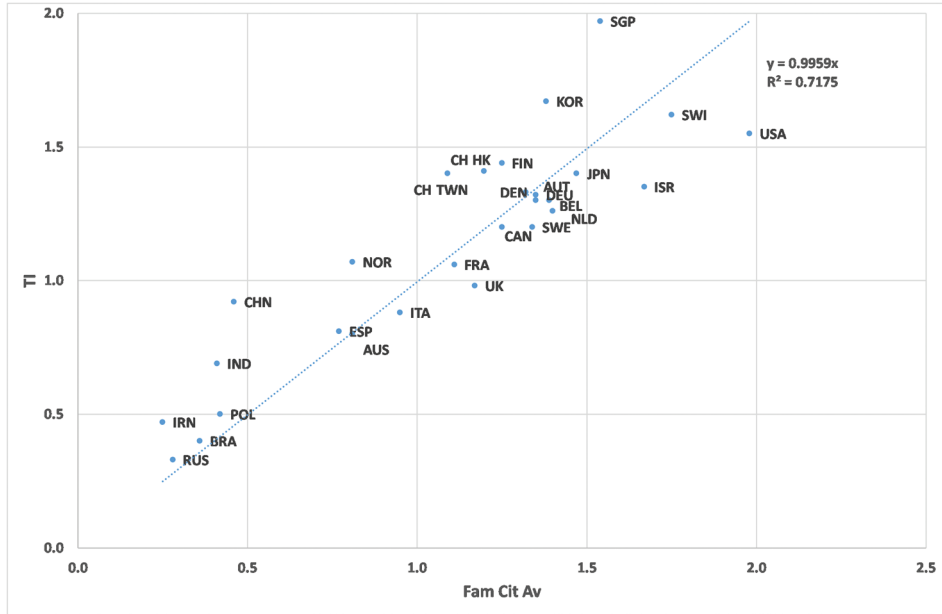


Figure 2. Technological Impact (TI) vs average number of patent family citations received per paper (Avg. Fam.Cit.) of the 28 countries/regions with the highest values of Technical Force.

Figure 2 is a scatter plot of Technological Impact (*TI*) vs the average number of patent family citations received (*Avg.Fam.Cit.*) for the 28 countries/regions with the greatest values of Technical Force (*TFO*). While *TI* is an advanced, normalized indicator, *Avg.Fam.Cit.* is a simpler, raw indicator. The plot clearly shows that Asian countries/regions in particular have relatively higher positions on the *TI* scale than on the *Avg.Fam.Cit.* scale. This means that *TI* indeed does make a difference in revealing features that remain hidden in analyses of straight patent citation counts.

Figure 3 is the size-independent analogue of Figure 1. The profiles it shows are more homogeneous than those shown in Figure 1 with there only being more variations in the average number of patent family citations.

The correlation matrix in Table 8 shows that Fam.Rel. (the ratio between a country/region's global share of patent families applying for protection and its global share of scientific papers) is the least correlated with the other indicators. The quality indicators of scientific production (% Q1, NI, % Exc10, and % Exc1) have greater correlations with each other, and Technological Impact logically correlates more strongly with the average number of patent citations. Technological Impact correlates a little more strongly with % Q1 and % Exc10 than with the other two indicators of scientific quality. In any case, the correlations of *TI* are weaker (around 0.7) than those of *TFO* with the rest of the size-dependent indicators (around 0.9).



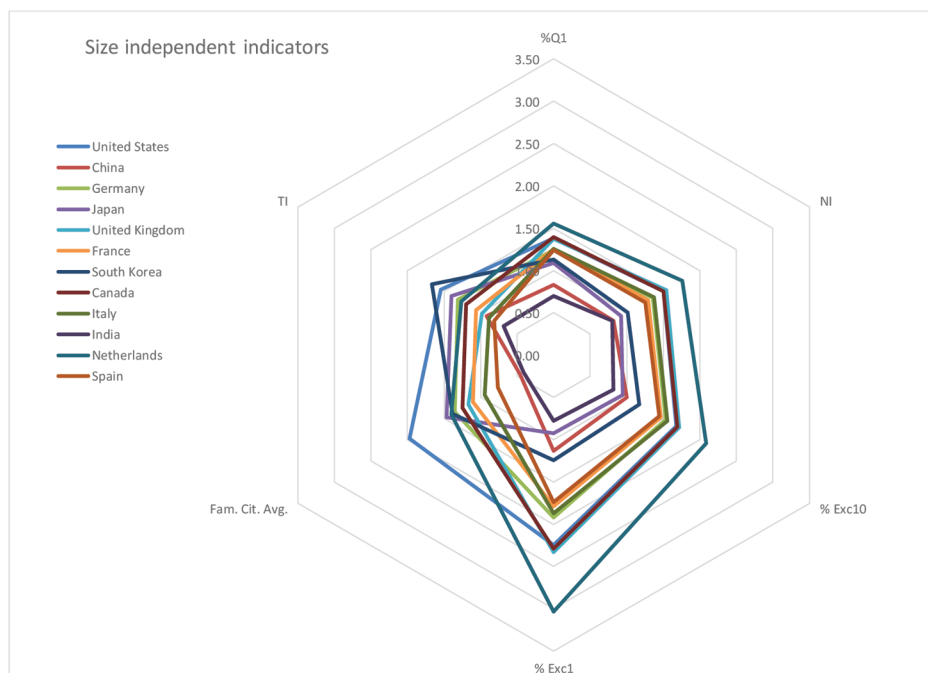


Figure 3. Size-independent indicators of the 12 countries with greatest Technical Force. All indicators are relative to the global average.

% Q1: Percentage of scientific papers in Q1.

NI: Normalized Impact.

% Exc10: Percentage of scientific papers of excellence (Top 10%).

% Exc1: Percentage of scientific papers of excellence (Top 1%).

Fam. Cit. Avg.: Average number of patent family citations per paper.

TI: Technological Impact.

Table 8. Pearson correlation matrix between size-independent indicators for the 40 countries/regions with the greatest scientific output.

	Fam.Rel.	%Q1	NI	%Exc10	%Exc1	Avg.Fam.Cit.	TI
Fam.Rel.	1	0.20	0.25	0.18	0.21	0.08	-0.00
%Q1	0.20	1	0.92	0.92	0.86	0.83	0.74
NI	0.25	0.92	1	0.99	0.98	0.74	0.69
%Exc10	0.18	0.92	0.99	1	0.98	0.75	0.73
%Exc1	0.21	0.86	0.98	0.98	1	0.70	0.69
Avg.Fam.Cit.	0.08	0.83	0.74	0.75	0.70	1	0.90
TI	-0.00	0.74	0.69	0.73	0.69	0.90	1



Research Paper

Figure 4 is a scatter plot of the countries/regions according to technological impact and scientific impact. One observes that Singapore has a high scientific impact, but above all stands out for its technological impact. Australia and the United Kingdom have significant scientific impact, but technological impact below the global average. Germany and Canada have lower scientific impact but greater technological impact. Japan, Chinese Taiwan, and South Korea stand out for their technological impact but not for their scientific impact.

The relative sizes of the circumferences also mark different profiles. While in Japan the thick circumference (Exc10) is the smallest, smaller than the two thin ones (TFO and patent family citations), in China the smallest is that of patent family citations. And in Australia and the United Kingdom, the largest circumference (TFO) is proportionally closer to the other two than in other countries/regions.

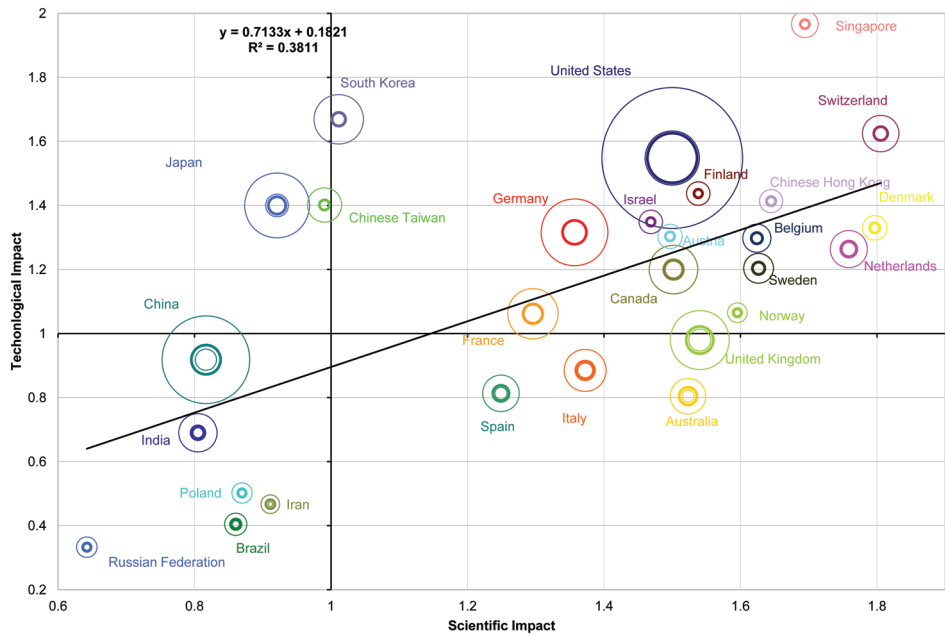


Figure 4. Scatter plot of Technological Impact vs Scientific Impact of the 28 countries/regions with the highest values of TFO. The circumferences correspond to: Technological Force (the outer thin circumference), Excellence 10 (thick circumference), the number of patent family citations (the inner thin circumference).

Figure 5 shows the temporal evolution of Technological Impact, Normalized Impact, and the average number of patent family citations for the two countries with the greatest scientific output and the countries/regions of the present authors. While the average citation rate from patent families has been in steady decline, TI has been stable except in the last two years. This loss of stability is due to the time it takes for the authorities to publish patent applications.





Figure 5. Annual evolution of three indicators—Normalized Impact (NI), average number of patent family citations (Fam.Cit. Avg.), and Technological Impact (TI)—in four countries.

6 Conclusions

The proposed indicators, Technological Impact (*TI*) and Technological Force (*TFO*), are designed to complement each other—the first being size-independent, and the second size-dependent. Their construction takes into account the countries/regions in which patent protection was sought, a patent family's propensity to cite scientific papers, merging and deduplication of the reference lists of applications corresponding to the same invention, and the time separating the publication date of a scientific paper and its subsequent citation in a patent.

Similar to the Normalized Scientific Impact indicator which normalizes the number of citations in scientific papers-to other papers with the average citation rate of articles in the same subject field and with the same publication year and document type, these advanced indicators of technological impact, being normalized with the scores corresponding to papers of the same age, are capable of partially eliminating the inevitable decline in patent citations to papers published in more recent years. It has to be noted, however, that the indicator *TI* was found to lose stability in the last two years due to the time that it takes for the various authorities to publish patent applications.

The indicators explored in the present communication show strong linear correlations with the indicators of scientific production, but are also able to reveal features that remain hidden in analyses of straight patent citation counts. An example was seen in showing the uniqueness of such technologically advanced countries as Singapore, South Korea, or Japan whose R&D systems are better integrated with innovation.



Research Paper

The proposed indicators are useful tools with which to characterize the technological orientation of research institutions.

Acknowledgments

This work was 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.

Author contributions

Vicente P. Guerrero-Bote (guerrero@unex.es): Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Funding acquisition (Lead), Investigation (Equal), Methodology (Lead), Project administration (Lead), Resources (Equal), Software (Lead), Supervision (Equal), Validation (Equal), Visualization (Equal), Writing-original draft (Lead), Writing-review & editing (Equal); Henk F. Moed (henk.moed@uniroma1.it): Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Funding acquisition (Equal), Investigation (Equal), Methodology (Equal), Project administration (Equal), Resources (Equal), Software (Equal), Supervision (Equal), Validation (Lead), Visualization (Equal), Writing-original draft (Equal), Writing-review & editing (Lead); Félix Moya-Anegón (felix.moya@scimago.es): Conceptualization (Equal), Data curation (Supporting), Formal analysis (Supporting), Funding acquisition (Equal), Investigation (Equal), Methodology (Equal), Project administration (Equal), Resources (Equal), Software (Supporting), Supervision (Supporting), Validation (Supporting), Visualization (Supporting), Writing-original draft (Supporting), Writing-review & editing (Supporting).

References

- Ahmadpoor, M., & Jones, B.F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, 357(6351), 583–587.
- Archambault, E., Campbell, D., Gingras, Y., & Larivière, V. (2009). “Comparing bibliometric statistics obtained from the Web of Science and Scopus”. *Journal of the American Society for Information Science and Technology*, 60 (7), 1320–1326. <https://doi.org/10.1002/asi.21062>
- Bryan, K.A., Ozcan, Y., & Sampat, B. (2020). In-text patent citations: A user’s guide. *Research Policy*, 49(4), 103946.
- Callaert, J., Pellens, M., & Van Looy, B. (2014). Sources of inspiration? Making sense of scientific references in patents. *Scientometrics*, 98(3), 1617–1629.
- Coffano, M., & Tarasconi, G. (2014). Crios—Patstat database: Sources, contents and access rules. Center for Research on Innovation, Organization and Strategy, Crios Working Paper n. 1. <https://ssrn.com/abstract=2404344> (accessed: 01/06/2020). <https://doi.org/10.2139/ssrn.2404344>
- Derwent. (2007). *World Patents Index—Derwent patentee codes*, Revised edition 8. Thomson Corporation. Leuven Manual. ISBN: 0 901157 38 4. <http://ips.clarivate.com/m/pdfs/mgr/patenteeCodes.pdf>
- European Patent Office. (2019). *Data catalog Patstat global. Versión 5.13. EPO Patstat customers*. <https://www.epo.org>



- Gorraiz, J., Gumpenberger, C., & Wieland, M. (2011). "Galton 2011 revisited: A bibliometric journey in the footprints of a universal genius". *Scientometrics*, 88(2), 627–652. <https://doi.org/10.1007/s11192-011-0393-y>
- Guerrero-Bote, V.P., & Moya-Anegón, F. (2015). "Analysis of scientific production in food science from 2003 to 2013". *Journal of food science*, 80(12), R2619–R2626. <https://doi.org/10.1111/1750-3841.13108>
- Guerrero-Bote, V.P., Sánchez-Jiménez, R., & Moya-Anegón, F. (2019). "The citation from patents to scientific output revisited: A new approach to Patstat / Scopus matching". *El profesional de la información*, 28(4), e280401. <https://doi.org/10.3145/epi.2019.jul.01>
- Hane, P.J. (2004). "Elsevier announces Scopus service". *Information today*. <http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=16494> (accessed: 01/06/2020).
- Hinze, S., & Schmoch, U. (2004). Opening the black box. In: Moed, H.F., Glänzel, W., and Schmoch, U. (2004) (eds.). *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*. Dordrecht (the Netherlands): Kluwer Academic Publishers, 215–235.
- Jacsó, P. (2011). "The h-index, h-core citation rate and the bibliometric profile of the Scopus database". *Online information review*, 35(3), 492–501. <https://doi.org/10.1108/1468452111151487>
- Jefferson, O.A., Jaffe, A., Ashton, D., Warren, B., Koellhofer, D., Dulleck, U., Bilder, G., Ballagh, A., Moe, J., DiCuccio, M., Ward, K., Bilder, G., Dolby, K., & Jefferson, R.A. (2018). "Mapping the global influence of published research on industry and innovation". *Nature biotechnology*, 36(1), 31–39. <https://doi.org/10.1038/nbt0818-772a>
- Kang, B., & Tarasconi, G. (2016). "Patstat revisited: Suggestions for better usage". *World patent information*, 46, 56–63. <https://doi.org/10.1016/j.wpi.2016.06.001>
- Leydesdorff, L., Moya Anegón, F., & Guerrero-Bote, V.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*, 61(2), 352–369. <https://doi.org/10.1002/asi.21250>
- Lissoni, F. (2012). "Academic patenting in Europe: An overview of recent research and new perspectives". *World patent information*, 34(3), 197–205. <https://doi.org/10.1016/j.wpi.2012.03.002>
- Lotti, F., & Marin, G. (2013). "Matching of Patstat applications to AIDA firms: Discussion of the methodology and results". *Bank of Italy occasional paper*, n. 166. <https://ssrn.com/abstract=2283111> (accessed: 01/06/2020). <https://doi.org/10.2139/ssrn.2283111>
- Lundberg, J. (2007). "Lifting the crown—citation z-score". *Journal of Informetrics*, 1(2), 145–154. <https://doi.org/10.1016/j.joi.2006.09.007>
- Magerman, T., Van-Looy, B., & Song, X.Y. (2006). Data production methods for harmonized patent statistics: Patentee name standardization. Technical report, K.U. Leuven. <https://ec.europa.eu/eurostat/documents/3888793/5836029/KS-AV-06-002-EN.PDF> (accessed: 01/06/2020).
- Maraut, S., & Martínez, C. (2014). "Identifying author–inventors from Spain: Methods and a first insight into results". *Scientometrics*, 101(1), 445–476. <https://doi.org/10.1007/s11192-014-1409-1>
- Martínez, C. (2011). "Patent families: When do different definitions really matter?", *Scientometrics*, v. 86, pp. 39–63. <https://doi.org/10.1007/s11192-010-0251-3>



Research Paper

- Moya-Anegón, F., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Corera-Álvarez, E., Muñoz-Fernández, F.J., González-Molina, A., & Herrero-Solana, V. (2007). "Coverage analysis of Scopus: A journal metric approach". *Scientometrics*, 73(1), 53–78. <https://doi.org/10.1007/s11192-007-1681-4>
- Moya-Anegón, F., Guerrero-Bote, V.P., López-Illescas, C., & Moed, H.F. (2018). "Statistical relationships between corresponding authorship, international co-authorship and citation impact of national research systems". *Journal of informetrics*, 12(4), 1251–1262. <https://doi.org/10.1016/j.joi.2018.10.004>
- Moya-Anegón, F., López-Illescas, C., Guerrero-Bote, V.P., & Moed, H.F. (2020). The citation impact of social sciences and humanities upon patentable technology. *Scientometrics*, to be published.
- Narin, F., & Olivastro, D. (1992). Status report: Linkage between technology and science. *Research Policy*, 21, 237–249.
- Narin, F., Hamilton, K.S., & Olivastro, D. (1997). The increasing linkage between US technology and public science. *Research Policy*, 26, 317–330.
- Organisation for Economic Co-operation and Development. (2009). *OECD Patent Statistics Manual*, OECD, Paris. <https://www.oecd.org/sti/inno/oecdpatentstatisticsmanual.htm>
- Pickering, B. (2004). "Elsevier prepares Scopus to rival ISI Web of science". *Information world review*, n. 8.
- Poege, F., Harhoff, D., Gaessler, F., & Baruffaldi, S. (2019). Science quality and the value of inventions. *Science advances*, 5(12), eay7323.
- Price, D.J.D (1965). Networks of scientific papers. *Science*, 149, 510–515.
- Raffo, J.D., & Lhuillery, S. (2009). "How to play the 'Names game': Patent retrieval comparing different heuristics". *Research policy*, 38(10), 1617–1627. <https://doi.org/10.2139/ssrn.1441172>
- Schmoch, U., & Khan, M. (2019). Methodological challenges for creating patent indicators. In: Glanzel, W., Moed, H.F., Schmoch, U., & Thelwall, M. (eds.), *Springer Handbook of Science and Technology Indicators*. Cham (Switzerland): Springer Nature, 907–927.
- Schoen, A., Heinisch, D., & Buenstorf, G. (2014). "Playing the 'Name game' to identify academic patents in Germany". *Scientometrics*, 101(1), 527–545. <https://doi.org/10.1007/s11192-014-1400-x>
- SCImago Research Group. (2010). *SIR Methodology: General Considerations*. <https://www.scimagoir.com/methodology.php> (accessed: 01/06/2020).
- Thoma, G., & Torrisi, S. (2007). Creating powerful indicators for innovation studies with approximate matching algorithms. A test based on Patstat and Amadeus databases (No. 211). KITEs, Centre for Knowledge, Internationalization and Technology Studies, Università Bocconi, Milano, Italy. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.573.8107&rep=rep1&type=pdf> (accessed: 01/06/2020).
- Van Raan, A.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*, 2, 13–50.
- Veugeliers, R., & Wang, J. (2019). Scientific novelty and technological impact. *Research Policy*, 48(6), 1362–1372.



Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S., & van Raan, A.F. (2011). "Towards a new crown indicator: Some theoretical considerations". *Journal of informetrics*, 5(1), 37–47.
<https://doi.org/10.1016/j.joi.2010.08.001>



This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

