# The topic normalized impact factor

Pablo Dorta-González[1] and María-Isabel Dorta-González[2]

[1] *pablo.dorta@ulpgc.es*
Universidad de Las Palmas de Gran Canaria (Spain)

[2] *isadorta@ull.es*
Universidad de La Laguna (Spain)

## Introduction

Traditionally, normalization for field differences has usually been done based on a field classification system. In said approach, each publication belongs to one or more fields and the citation impact of a publication is calculated relative to the other publications in the same field. An example of a field classification system is the JCR subject category list. For these subject categories, Egghe & Rousseau (2002) propose the aggregate impact factor in a similar way as the JIF, taking all journals in a category as one meta-journal. However, the position of individual journals of merging specialties remains difficult to determine with precision and some journals are assigned to more than one category. In this sense, Dorta-González & Dorta-González (2013a) propose the categories normalized impact factor considering all the indexing categories of each journal.

Recently, the idea of source normalization was introduced, which offers an alternative approach to normalizing field differences. In this approach, normalization is achieved by looking at the referencing behaviour of citing journals. Many indices, such as the fractionally counted impact factor (Leydesdorff & Bornmann, 2011), dividing each citation by the number of references, and the 2-year maximum journal impact factor (Dorta-González & Dorta-González, 2013b), considering the 2-year citation time window of maximum impact instead of the previous 2-year time window, have been proposed.

However, all these metrics do not include any great degree of normalization in relation to the specific topic of each journal. In this sense, we use the aggregate impact factor of the citing journals as a measure of the citation potential in the journal topic, and we employ this citation potential in the normalization of the journal impact factor to make it comparable between scientific fields.

## Citation potential and journal topic

The editorial policy of a journal determines its explicit topic. However, the implicit topic can be determined by its scientific impact. In this sense, we can define the topic of the citation impact of a journal, hereafter journal topic, through all the citing journals. For example, if a journal *j* is cited by journals in *n* different fields, then the journal topic can be characterized by all these *n* fields in a proportional form to the number of citations to journal *j*.

We define the citation potential in the topic of journal *j* in a year *y* as the weighted average of the impact factors of all citing journals to *j* in the year *y* with respect to the previous two years. This average is weighted by the number of citations to *j*, excluding self-citations of *j* to *j*.

Consider the example in Table 1. Let *j* be a journal with *JIF* = 2.000 and the citing journals indicated in the table (excluding *j*). The citation potential in the topic of journal *j* is 0.5×1.000 + 0.3×2.500 + 0.15×0.800 + 0.05×1.400 = 1.440. The journal impact factor (2.000) is 39% greater than the citation potential in the topic (2.000 / 1.440 = 1.39) and, therefore, in the comparison with other journals the *JIF* should be proportionally increased in a way that will be illustrated below.

**Table 1: One example**

| Citing journal | Total citations | Weight | *JIF* |
|---|---|---|---|
| 1 | 100 | 100 / 200 = 0.5 | 1.000 |
| 2 | 60 | 60 / 200 = 0.3 | 2.500 |
| 3 | 30 | 30 / 200 = 0.15 | 0.800 |
| 4 | 10 | 10 / 200 = 0.05 | 1.400 |
| Total | 200 | 1 | |

We define the Topic Normalized Impact Factor of journal *j* in year *y* as:

$$TNIF_y^j = \frac{CP_y^J}{CP_y^{T_j}} \times JIF_y^j$$

where $JIF_y^j$ is the journal impact factor of *j* in year *y*, $CP_y^J = \sum_{j \in J} \left( v_y^j \times JIF_y^j \right)$ is the citation potential of database *J* in year *y* (aggregate impact factor in

$J$), and $v_y^j$ is the weight of journal $j$ in database $J$ in the target window of year $y$.

Similarly, $CP_y^{T_j} = \sum_{i \in T_j} \left( w_y^{ij} \times JIF_y^i \right)$ is the citation potential of topic $j$ in year $y$ (aggregate impact factor in topic $j$), and $w_y^{ij}$ is the weight of journal $i$ in the topic of $j$ in year $y$.

In the example of Table 1, considering that $CP_y^J = 1.800$ then the normalized score of journal $j$ is $CP_y^J / CP_y^{T_j} = 1.800 / 1.440 = 1.25$ and the $TNIF_y^j = 1.25 \times 2.000 = 2.500$. This amount is greater than the *JIF* because the citation potential of the database is greater than the citation potential in the topic of the journal.

**Materials and Methods**

We used four impact indicators: 2-year journal impact factor (*2-JIF*), 5-year journal impact factor (*5-JIF*), fractionally counted impact factor (*FCIF*), and our topic normalized impact factor (*TNIF*).

We designed a cluster sample. Cluster sampling is a two-stage sampling design in which, firstly, one single cluster is randomly selected from a set of clusters and, secondly, all observations in the selected cluster are included in the sample. Four fields (journal categories), each one from a different cluster obtained by Dorta-González & Dorta-González (2013a), were considered. This was motivated in order to obtain journals with systematic differences in publication and citation behaviour. A total of 224 journals were considered. The journal categories and the number of journals are: Astronomy & Astrophysics (56); Biology (85); Engineering, Aerospace (27); and History & Philosophy of Science (56).

**Results and discussion**

In the empirical application we studied which impact indicator produces a closer data distribution among scientific fields in relation to its centrality and variability measures.

The fields considered are very different in relation to the citation behavior and some of them are penalized by the *JIF* (Table 2). Note that the medians of the *JIF* in Astronomy & Astrophysics and Biology are very much higher than those in Aerospace Engineering and History & Philosophy of Science; in general, more than three times higher. However, the central-tendency measures of the *TNIF* are closer in all the fields considered.

Moreover, the *TNIF* reduces the between-group variance in relation to the within-group variance more than the rest of the indicators analyzed in this paper. Within- and between-group variability are both components of the total variability in the combined distributions. So: within variability + between variability = total variability. A more detailed analysis of the results can be found in Dorta-González et al. (2014).

**Table 2: Central-tendency and variability**

| | ***2-JIF*** | ***5-JIF*** | ***FCIF*** | ***TNIF*** |
|---|---|---|---|---|
| *Astronomy & Astrophysics* | | | | |
| Median | 1.683 | 1.757 | 0.31919 | 1.723 |
| Sd | 4.292 | 4.548 | 0.39276 | 2.457 |
| *Biology* | | | | |
| Median | 1.540 | 1.719 | 0.20534 | 2.993 |
| Sd | 2.115 | 2.375 | 0.26257 | 3.086 |
| *Engineering, Aerospace* | | | | |
| Median | 0.549 | 0.654 | 0.13113 | 1.507 |
| Sd | 0.605 | 0.734 | 0.12627 | 2.390 |
| *History & Philosophy of Science* | | | | |
| Median | 0.442 | 0.553 | 0.09312 | 1.810 |
| Sd | 0.603 | 0.636 | 0.07779 | 5.274 |

**Conclusions**

We provide a source normalization approach based on the journal topic and we compare it with some popular impact indicators.

The fields considered are very different in relation to the citation behavior. For this reason, the *JIF* in Astronomy & Astrophysics and Biology are very much higher than the *JIF* in Aerospace Engineering and History & Philosophy of Science. However, the *TNIF* is very close in all the fields considered.

**References**


Dorta-González, P., & Dorta-González, M. I. (2013a). Comparing journals from different fields of science and social science through a JCR subject categories normalized impact factor. *Scientometrics*, 95(2), 645–672.

Dorta-González, P., & Dorta-González, M. I. (2013b). Impact maturity times and citation time windows: The 2-year maximum journal impact factor. *Journal of Informetrics*, 7(3), 593–602.

Dorta-González, P., Dorta-González, M.I., Santos-Peñate, D.R., & Suárez-Vega, R. (2014). Journal topic citation potential and between-field comparisons: The topic normalized impact factor. *Journal of Informetrics* 8(2), 406–418.

Egghe, L., & Rousseau, R. (2002). A general framework for relative impact indicators. *Canadian Journal of Information and Library Science*, 27(1), 29–48.

Leydesdorff, L., & Bornmann (2011). How fractional counting of citations affects the Impact Factor: Normalization in terms of differences in citation potentials among fields of science. *Journal of the American Society for Information Science & Technology*, 62(2), 217–229.