Reproducibility, consensus and reliability in bibliometrics

Raul I. Mendez-Vasquez[1] and Eduard Suñen-Pinyol[2]

[1]raul.mendez@fundaciorecerca.cat, [2]eduard.sunen@fundaciorecerca.cat
Fundació Catalana per a la Recerca i la Innovació (FCRI), Bibliometrics. Pg. Lluís Companys, 23 E-08010 Barcelona (Spain)

Introduction

Bibliometrics, and scientometrics in general, have been enjoying what seems to be an endless party. Far from stopping, the demand for bibliometric indicators from governmental bodies, administrators and researchers, is continuously growing. During this "give me the indicators" phase several solutions have been provided by the community, let say new and more sophisticated indicators, which in turn geared the transition to the present "give me the indicators, but really?" phase. The impressive penetration of bibliometric indicators in decision making processes, some of which are crucial in the development of researchers' careers, has also brought the necessity for credibility on bibliometrics, and more specifically, on how it is practiced. Examples of improper use of bibliometric indicators have raised skepticism among users of bibliometric reports[1].

As a scientific discipline, bibliometrics is subject to the principle of replication and corroboration of results, just like any other discipline. Precisely, the credibility of scientists goes hand in hand with the reproducibility of their results.

The objective of this contribution is to bring attention to the importance of the reproducibility of the number of publications as an indicator of the quality of bibliometric reports.

Methods

We compared the numbers of publications estimated by three units following this schema: CTWS vs. BAC (us) and SCIMAGO vs. BAC. Sixteen universities reported in the CTWS Leiden Ranking 2011/2012, and 20 universities reported in the Iberoamerican Ranking SIR 2012 produced by SCIMAGO were selected for the study. Source, type of document, language and period were matched in each comparison. The numbers of publications produced by the BAC were sourced with the National Citation Report for Spain (NCR), an *ad hoc* database built in July 2012 as a live extraction from the Web of Science that compiles all the publications between 1970 and 2011, with at least one address in Spain. The unification was performed by hand based solely on the information contained in the address field of the NCR. Hierarchy relationships such as university campuses and institutes, affiliated hospitals, etc, were reconstructed in the system. All the addresses were also located to a specific administrative unit (a city in the majority of cases). Both, the information on the organizational hierarchy and location of the addresses were used to unify the name variants of subunits whenever mother organizations were not present in the addresses. Changes in the structure of the organizations within the analyzed period were recorded in the system. The unification terminated when a precision higher than 97% was achieved.

Results

A simple examination of the number of publications of a small set of universities revealed important reproducibility issues, even when controlling for source dataset, period of time and the document type (Table 1. several rows and columns were removed). A positive and statistically significant correlation ($p<0.01$) was observed between the numbers of publications produced by the three units (CTWS & BAC, *rho* 0.785; SCIMAGO & BAC, *rho* 0.860). The dispersion around the regression line was smaller in the comparison between SCIMAGO & BAC, than between CTWS & BAC, suggesting the presence of an outlier observation, whose removal increased the correlation between CTWS and BAC (*rho* 0.975, $p<0,001$). The concordance between the rankings produced by the three units was also positive and high, (CTWS & BAC, *tau* 0.733, $p<0.001$; SCIMAGO & BAC, *tau* 0.705, $p<0.001$). Removing the mentioned outlier observation increased the concordance between the CTWS and BAC (*tau* 0.905, $p<0.001$)

Discussion

These technical issues may explain the observed variability in the number of publications.

1) Completeness of the unification. The CTWS unit selected the universities with at least 500 publications per year and extended the unification to the name variants occurring at least five times in the source dataset. The BAC unit aims at attributing all variants to corresponding universities. However, mistakenly attributed name variants and non-identified variants were allowed to a maximum of 3%. The CTWS unit attributed the publications

[1]The title of a number of articles published in Nature in 2010 reflect this position: "Assessing assessment", "Do metrics matter?", "How to improve the use of metrics", "Let's make science metrics more scientific". Available at: http://www.nature.com/news/specials/metrics/index.html.

based on author names, a procedure not performed by the BAC. SCIMAGO provides no information on the unification in the website of the report.

Table 1. Differences in the number of publications produced by three units.

	(A)	(B)	A-B	(A-B)/A	(C)	(D)	C-D	(C-D)/C
UB	7,672	11,804	-4,132	-53,86	15,290	16,222	-932	-6,10
UAB	5,992	9,319	-3,327	-55,52	13,262	13,200	62	0,47
UCM	6,616	8,863	-2,247	-33,96	13,240	12,160	1,080	8,16
UPM	2,323	8,813	-6,490	-189,2	7,458	11,096	-3,638	-48,78
UAM	5,236	8,034	-2,798	-53,44	10,591	10,873	-282	-2,66
UV	5,077	7,892	-2,815	-55,45	11,191	10,458	733	6,55
UGR	3,966	5,918	-1,952	-49,22	9,128	8,117	1,011	11,08
USC	3,589	5,181	-1,592	-44,36	7,132	6,854	278	3,90
US	3,848	4,909	-1,061	-27,57	7,933	6,366	1,567	19,75
UPC	3,067	4,900	-1,833	-59,77	11,068	6,502	4,566	41,25
UZAR	3,394	4,612	-1,218	-35,89	7,607	6,102	1,505	19,78
EHU	3,047	4,536	-1,489	-48,87	7,520	6,535	985	13,10
n			**16**	**16**			**20**	**20**
Avg[1]			**-2,165**	**-51,40**			**659**	**7,30**
SDev.[2]			**1,508**	**-39,37**			**1,722**	**19,56**
CI[3]			**-739**	**-19,29**			**755**	**8,57**

A, data reported in the Leiden Ranking 2011/2012; **B**, number of publications estimated by BAC; **A-B**, magnitude of the difference between CTWS and BAC; (**A-B)/A**, percentage of change between CTWS and BAC; **C**, data reported in the Iberoamerican Ranking SIR 2012; **D**, number of publications estimated by BAC applying SCIMAGO criteria, but sourcing the analysis with the WOS; **C-D**; magnitude of the difference between SCIMAGO and BAC; (**C-D)/C**, percentage of change between SCIMAGO and BAC. 1; average; 2, standard deviation; 3, 95% confidence interval of the average. Acronyms: UB, Univ. de Barcelona; UAB, Univ. Autònoma de Barcelona; UCM, Univ. Complutense de Madrid; UPM, Univ. Politécnica de Madrid; UAM, Univ. Autónoma de Madrid; UV), Univ. de València; UGR, Univ. de Granada; USC, Univ. de Santiago de Compostela; US, Univ. de Sevilla; UPC, Univ. Politècnica de Catalunya; UZAR, Univ. de Zaragoza; EHU, Univ. del País Vasco.

2) Exactness of the unification. The CTWS unit estimated a 5% of false negative cases, while the BAC ensures a maximum percentage of error of 3%. SCIMAGO provides no information on this regard.

3) Proximity to the units under analysis. Two observations support the notion that local knowledge may explain a substantial part of the observed discrepancies: 1) the difference between SCIMAGO & BAC was smaller than between CTWS & BAC, and 2), SCIMAGO attributed more publications to their neighboring universities (UGR & US) than BAC, and vice versa in the case of the UB & UAB). A comparison of the number of publications of the Dutch universities between CTWS and BAC may shed some light on the effect that local knowledge or "regional peculiarities" (Moed, 1996) have on this indicator.

4) Delineation of the universities. The CTWS unit took into account "important university institutes" and changes in the structure of universities, while BAC took into account institutes, but also faculties, technical schools, locations, and structural changes. Failing to aggregate the publications of subunits could also explain the observed differences (de Mesnard, 2012).

5) Completeness and accuracy of the database (location of addresses). There is a difference between the sources used by the CTWS unit and BAC. The NCR may compile fewer records than the WOS, as addresses have to be located to Spain and errors are likely to happen during this process. This inconsistency may also play a lesser role in the comparison between CTWS and BAC.

Final considerations

Discrepancies in the number of publications of universities in the order of 10^2 or 10^3 are irrelevant when comparing the figures produced by different units. However, the magnitude of the difference might represent half of the output in some cases. Fortunately, the numbers of publications produced by the three units correlated pretty well, and the rankings were concordant. Technical issues can no longer be used as arguments to explain divergences of this magnitude, as none of the factors presented here are completely dependent on the technical capacity of a unit, rather than on procedural decisions: 1) completeness and 2) exactness of the unification, 3) knowledge of the surrounding environment, 4) completeness and accuracy of the source or 5) the type of document and period of time. The findings suggest that a consensus addressing these factors would do more in reaching a methodological "greatest common denominator" between the different units enabling improving the reproducibility of the indicators.

References

de Bruin R.E. & Moed H.F. (1990). The unification of addresses in scientific publications. In: Egghe L., Rousseau R. (editors.), *Informetrics*, 65-78.

Butler L. (1999). Who 'Owns' this Publication?, *Proc. ISSI*, 87-96.

de Mesnard L. (2012). On some flaws of university rankings: The example of the SCImago report. *The Journal of Socio-Economics 41*(5), 495–9

Moed H.F. (1996). Differences in the construction of SCI based bibliometric indicators among various producers: A first overview. *Scientometrics, 35*(2):177-191 DOI: 10.1007/BF02018476

Van Raan A. F. J. (2005a). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, *62*(1), 133-143.