Towards new scientific development models and research assessment support tools

Henk F. Moed

Independent researcher and scientific advisor, Amsterdam, The Netherlands. Email: hf.moed@gmail.com

Paper submitted to OECD Blue Sky Conference, Ghent, 19-21 September 2016

Version 25 July 2016

Summary

This contribution proposes new bibliometric support tools for research assessment in two major application domains: monitoring scientifically developing countries, and assessing the research performance of individual researchers, research groups and institutions. A model is introduced to measure the state of a country's scientific development. To reduce the role of journal impact factors in journal and research assessment it is proposed to develop indicators of the quality of journals' manuscript peer review process, based on an analysis of this process itself. A bibliometric, flexible author self-assessment tool is described that may provide authors with information that they can use against inaccurate calculation or inappropriate use of indicators. The need is underlined to conduct a meta-analysis of the numerous world university ranking systems that are currently available.

1. Introduction

Based on the notion of a multi-dimensional research assessment matrix (AUBR, 2010), the choice of a research assessment methodology depends upon a series of factors: What is the unit of assessment? What is the aspect to be assessed? What is the objective of the process? And what is the state of the system under assessment in its wider environment? Indicators suitable in one context may be inappropriate in another. When setting up a research assessment process, a crucial element is a "meta-analysis" of the units under assessment in which metrics are not so much used as tools to evaluate individual units, but to inform the process' objectives and general design (Moed & Halevi, 2015). The present contribution focuses on two broad application domains which receive a high priority in national and supra-national science policies. Both are founded in the notion that research and development is a key driver to innovation, growth and economic prosperity.

The *first* focuses on its implications for policies within and towards developing countries, and relates to the reinforcement of their scientific-technological infrastructures. These countries need tools to monitor the state of their scientific development. They need a framework that enables them to roughly categorize countries in terms of scientific development and compare "like with like" rather than directly compare them on the basis of absolute numbers. Such numbers are often of limited value and can even be discouraging as they are extracted from global ranking systems which tend to position scientifically developed countries rank in a ranking's upper part. The framework should also acknowledge that the

scientific and economic conditions in which a less-developed country finds itself at a certain moment are not static, but can be viewed as a phase in a process that more scientifically developed countries have already started previously with great success. Section 2 presents a model of a country's scientific development using bibliometric indicators based on publications in international, peer-reviewed journals.

The *second* application domain is that of assessing research performance of individuals, research groups, and research institutions. Securing a political basis for academic research is a principal concern and a joint responsibility of (supra-) national political domain and the academic research community. National research assessment exercises, performance-based funding, and assessments within institutions are important elements in this process. They increasingly need validated information and valid assessment methodologies in which indicators play a key role. Despite the critique in the DORA Manifesto (DORA, 2009) published about 6 years ago, journal impact factors are still heavily used by research managers, evaluators and publishers in the assessment of individual researchers and scientific journals. Moreover, an indicator of the performance of individual researchers, h-index (Hirsch, 2005) has become a very popular bibliometric measure, despite serious critique on its validity and limited scope of application (e.g., Adler, Ewing & Taylor, 2008).

Section 3 proposes the development of three new, alternative research assessment tools, namely indicators of the quality of a journal's manuscript peer review system, a flexible bibliometric author self-assessment tool, and a meta-analysis of the numerous world university ranking systems that are being published nowadays.

2. A bibliometric model that describes the state of scientific development

A simplified and experimental bibliometric model for different phases of development of a national research system distinguishes four phases: i) a pre-development phase; ii) building up; iii) consolidation and expansion; and iv) internationalization, and is described in Table 1.

Table 1. A bibliometric model for capturing the state of a country's scientific development

Phase	Description	Trend in # published articles (*)	Trend in % internationally co-authored publications
Pre-development	Low research activity without clear policy of structural funding of research	**~**	**~**
Building up	Collaborations with developed countries are established; national researchers enter international scientific networks	**+**	**++**
Consolidation and expansion	The country develops its own infrastructure; the amount of funds available for research increases	**++**	**-**
Internationalization	Research institutions in the country start as fully-fledged partners, increasingly take the lead in international collaboration	**+**	**+**

(*). ~ denotes: no clear trend; +: increase; - : decline; ++: strong increase. Source: UNESCO (2014).

The model assumes that during the various phases of a country's scientific development, the number of published articles in peer-reviewed journals shows a more or less continuous increase, although the rate of increase may vary substantially over the years. The share of a country's internationally co-authored articles, however, discriminates between the various phases in the development. Table 1 reflects the discontinuities that the model assumes that take place in the indicator values over time when moving from one phase into another. The model is experimental and needs to be further validated and empirically tested. From a bibliometric point of view, citation impact indicators should be added.

International scientific collaboration is not only a matter of development; cultural and historical factors play a role as well. The development process that the model seeks to capture takes place during a time span that is much longer than the time period analysed in this section. Rather than following one particular country through all stages, this study tracks a series of countries during the most recent ten-year period and draws hypotheses on the current phase of a country's development on the basis of the assertions made in the development model. Figure 1 applies the model to the scientific output of a group of major countries in the Persian Gulf Region (Moed, 2016a). For a similar analysis of 25 Asian countries, the reader is referred to UNESCO (2014, Ch 4, pp. 79-92).

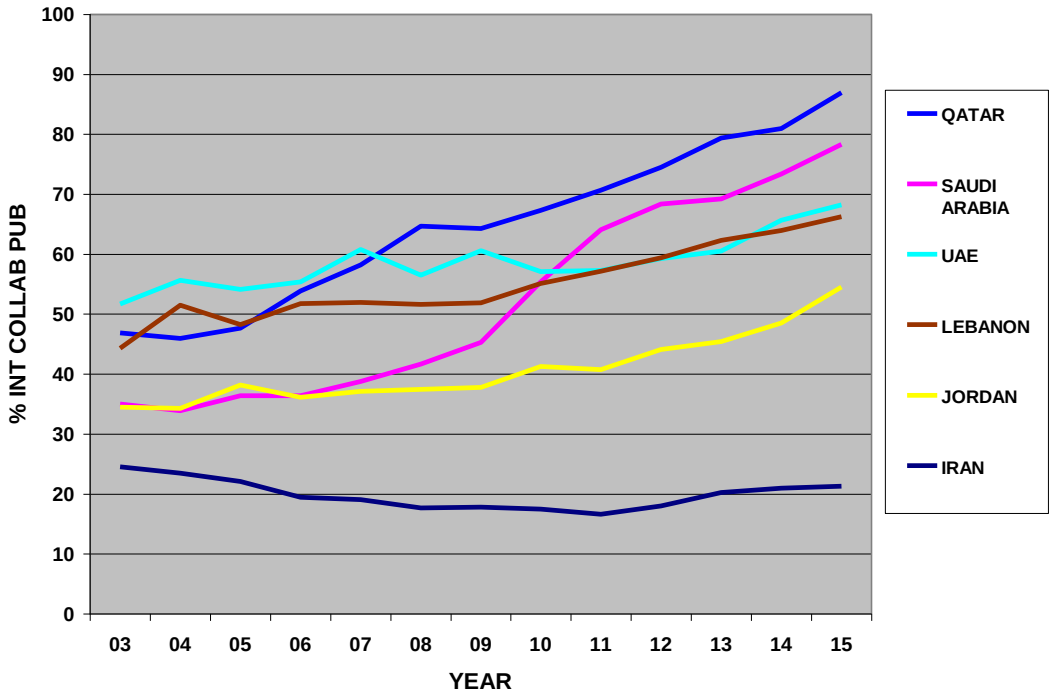

Figure 1. Percentage of internationally co-authored publications (relative to total publication output) of major Gulf States. Data from Elsevier's Scopus. Figure copied from Moed (2016a).

Large differences exist between Iran and Saudi Arabia with respect to the amount of foreign input needed to produce these papers. The percentage of internationally co-authored publications (ICAP) is in 2015 almost 80 per cent for Saudi Arabia, but only around 20 per cent for Iran. The results suggest that the two countries are in different phases of scientific development. While Saudi Arabia and most other Gulf States are still in the building up phase, Iran is currently moving from a consolidation and expansion into the

internationalization phase. This trend can be expected to continue now that the international boycotts are cancelled. The other Gulf States still depend in various degrees upon collaboration with external institutions, increase their ICAP rate and are in a phase of building up a scientific infrastructure. This is especially true for the two countries showing the largest increase of their publication output, namely Qatar and United Arab Emirates.[1]

3. Metrics for individual researchers, research groups and institutions

3.1 Indicators of the quality of the peer review process

The most effective way to reduce the role of citation based journal metrics in journal and research assessment is the development of indicators of the quality of journals' manuscript peer review process, based on an analysis of this process itself, rather than on proxies such as citation-based measures or manuscript rejection rates. To the extent that research evaluation agencies consider journal quality and especially the quality of its review process a relevant criterion in the assessment of individuals or groups, and are interested as to whether a researcher under assessment has submitted his or her articles to a journal with a serious referee procedure and well instructed reviewers, these agencies would profit from more direct measures of journal quality, and indicator developers should make an attempt to develop these.

In a research project aimed to develop indicators of the journal manuscript review process, computational linguistic tools from the domain of 'digital humanities' are useful, but 'classical humanistic' text analysis and a profound knowledge of the manuscript peer review and the publication process are essential as well. The project would have at least two phases. A first phase involves the development of a conceptual model of manuscript peer review, including the construction of referee report profiles, and communication modes between referees, authors and editors. In this phase, a conceptual analysis is conducted of a sample of actual referee reports for a number of sources from different subject fields.

In the second phase of the project, when at least a first version of the model developed in the first phase is available, data mining is carried out of large numbers of electronic submissions. A linguistic analysis is conducted of peer review reports and the communication between authors, referees and editors, using natural language processing and other computational linguistic techniques. It is in this phase that a statistical analysis of large datasets explores the construction of indicators of the peer review process, not only at the level of individual submissions, but also at the level of journals and subject fields. The outcomes of such exploration may lead to adjustments in the model developed in the first phase.

---

[1] Gringas (2014) found that a disproportionally large number of researchers appearing in the Thomson Reuters' list of highly cited researchers indicate Saudi universities as secondary address, thus boosting these institutions up in global university rankings. The articles in which this occurs are counted as internationally co-authored publications in Figure 1 and can be expected to boost up the Saudi percentage of ICAP as well. Even if the influence of this phenomenon is substantial, it underlines the dependence of this country upon the input of foreign researchers and does therefore not violate the conclusions on its state of scientific development.

It must be noted that the quality of the manuscript peer review processes is a research topic in its own right, regardless of whether it aims to contribute to the development of better indicators of journal quality. The outcomes of the proposed study could further enhance the transparency of the manuscript referee process, also for submitting authors. Most importantly in my view, they could help educating and training new reviewers. Next, they could help assessing the effect of peer review upon manuscript quality. This could provide information that can be used to demonstrate the added value of the process.

Finally, it could contribute to further operationalizing the multi-dimensional concept of quality of a journal's peer review process, and develop indicators that can be used to monitor and further improve this process, jointly with advanced online tools editors and reviewers need in their tasks, such as, for instance, plagiary detection tools. In this way, these indicators could potentially be used as an alternative of the journal impact factor as a more direct measurement of the quality of a journal's manuscript peer review process. For a more detailed discussion of this approach the reader is referred to Moed (2016b).

4.2. A bibliometric author self-assessment tool

What is lacking is a widely available bibliometric tool for researchers specially designed for *self-assessment*, enabling users to verify selected publication data, create a set of 'candidate' benchmark authors using bibliometric algorithms similar to those proposed by Eugene Garfield for evaluating faculty (Garfield, 1983), and offer a flexible benchmarking feature as the practical realization of Robert K. Merton's notion of a *reference group*, i.e., the group with which individuals compare themselves, but to which they do not necessarily belong but aspire to (see Holton, 2004). This tool would expectedly make users more acquainted with the ins and outs of bibliometric indicators in general and with the underlying data, and teach them for instance how outcomes of an assessment depend upon the way benchmark sets are being defined.

The experiences and insights collected in this way would enable researchers subjected to assessment to critically follow assessment processes, and to defend themselves against inaccurate calculation or inappropriate use of indicators. Distributing knowledge and experiences is also a necessary condition for achieving more standardisation of bibliometric concepts and methods. A bibliometric self-assessment tool could become a genuine alternative to using journal impact factors or h-indices in the assessment of research performance of individuals and groups.

4.3. A meta-analysis of world university ranking systems

An Expert Group on the assessment of university-based research noted in 2009 that university rankings have become an increasing influence on the higher education landscape since US News and World Report began providing consumer-type information about US universities in 1983. They "enjoy a high level of acceptance among stakeholders and the wider public because of their simplicity and consumer type information" (AUBR Expert Group, 2009, p 9).

The university ranking systems have been intensely debated. A report from the European University Association concluded that despite their shortcomings, evident biases and flaws, rankings are here to stay. “For this reason it is important that universities are aware of the degree to which they are transparent, from a user’s perspective, of the relationship between what it is stated is being measured and what is in fact being measured, how the scores are calculated and what they mean” (Rauhvargers, 2011, p. 7).

To provide insight to users to the type of issues highlighted by Rauhvargers (2011), a meta-analysis of a series of university ranking systems, – or, more generally, information systems providing indicators of the performance of world universities – would be a useful element. In a first step, the various ranking systems are to be linked with one another at the level of individual institutions, and indicators grouped into families measuring the same or very similar aspects. A second step analyses the degree of overlap of the sets of universities covered between any pair of systems, the consistency of collected data between systems, and statistical correlations between the major indicators, both within and between indicator families. This information could be helpful to assess the uniqueness and the added value of one system compared to all other.

Combined with qualitative insights, a third step identifies characteristics of national academic systems. In addition, it conducts a search towards possible determinants of research performance at the institutional level, and works towards a model of the development of academic institutions.

References

AUBR Expert Group (2010). Expert Group on the Assessment of University-Based Research. Assessing Europe’s University-Based Research. European Commission – DG Research. EUR 24187 EN.

Adler, R., Ewing, J., and Taylor, P. (2008) Citation statistics. A report from the International Mathematical Union. www.mathunion.org/publications/report/citationstatistics0.

DORA (2009). San Fransisco Declaration on Research Assessment. http://am.ascb.org/dora/.

Garfield, E. (1983). How to use citation analysis for faculty evaluation, and when is it relevant. Part 1. Current Contents, 44, 5–13, October 31, 1983. In: Essays of an Information Scientist, 6, 354–362. Philadelphia: ISI Press.

Gringas, Y. (2014). How to boost your university up the rankings. University World News. 18 July 2014 Issue No:329.
http://www.universityworldnews.com/article.php?story=20140715142345754

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569–16572.

Holton, G. (2004). "Robert K. Merton - Biographical Memoirs". Proceedings of the American Philosophical Society 148 (4): 506–517.

Moed, H.F., Halevi, G. (2015). Multidimensional Assessment of Scholarly Research Impact. Journal of the American Society for Information Science and Technology, 66, 1988-2002. Author copy available at: http://arxiv.org/abs/1406.5520.

Moed, H.F. (2016a). Iran's scientific dominance and the emergence of South-East Asian countries as scientific collaborators in the Persian Gulf Region. Scientometrics, 101, 3015-314. Author copy available at http://arxiv.org/ftp/arxiv/papers/1602/1602.04701.pdf.

Moed, H.F. (2016b). Towards new indicators of a journal's manuscript peer review process. Frontiers in Research Metrics and Analytics, to be published.

Rauhvargers, A. (2011). Global rankings and their impact. EUA Report on Rankings 2011. Brussels, Belgium: The European University Association. 79 pp.

UNESCO (2014). Higher Education in Asia: Expanding Out, Expanding Up. ISBN 978-92-9189-147-4 licensed under CC-BY-SA 3.0 IGO. Montreal: UIS. http://www.uis.unesco.org.