Journal of Informetrics 5 (2011) 64–74

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Structured evaluation of the scientific output of academic research groups by recent *h*-based indicators

Fiorenzo Franceschini*, Domenico Maisano

Politecnico di Torino, Dipartimento di Sistemi di Produzione ed Economia dell'Azienda (DISPEA), Corso Duca degli Abruzzi 24, 10129 - Torino, Italy

ARTICLE INFO

Article history:
Received 29 June 2010
Received in revised form 25 August 2010
Accepted 25 August 2010

Keywords:
Bibliometric positioning
h-Index
h-Spectrum
Successive *h*-indices
Ch-index
Research evaluation
Academic research group
Scientific production

ABSTRACT

Evaluating the scientific output of researchers, research institutions, academic departments and even universities is a challenging issue. To do this, bibliometric indicators are helpful tools, more and more familiar to research and governmental institutions.

This paper proposes a structured method to compare academic research groups within the same discipline, by means of some Hirsch (*h*) based bibliometric indicators. Precisely, five different typologies of indicators are used so as to depict groups' bibliometric positioning within the scientific community. A specific analysis concerning the Italian researchers in the scientific sector of *Production Technology and Manufacturing Systems* is developed. The analysis is supported by empirical data and can be extended to research groups associated to other scientific sectors.

1. Introduction

Evaluating the scientific production of research groups, departments and even universities and research organizations, is not a new issue. For many years, different rankings have been periodically published by many institutions. For example, the doctoral programs' ranking by the National Research Council or several annual universities' rankings, such as those by the US News and World Report, the Shanghai Jiao Tong University, the Center for Science and Technology Studies at the University of Leiden, the German Center for Higher Education Development, and many others [Buela-Casal, Gutiérrez-Martínez, Bermúdez-Sánchez, & Vadillo-Muñoz, 2007; Bach & Llerena, 2007; Kellner & Ponciano, 2008; Lindsey, 1991; Lombardi, Craig, Capaldi, & Gater, 2002; Van den Berghe et al., 1998]. While the inherent limitations of these rankings have been richly discussed [Billaut, Bouyssou, & Vincke, 2010; Bornmann, 2010; Cai Liu, Cheng, & Liu, 2005, 2006; Opthof & Laydesdorff, 2010; Van Raan, 2005], it is interesting to notice that they are becoming more and more demanded.

In general, rankings are based on a number of evaluation criteria, one of which is the scientific production. Traditional approaches for evaluating and comparing the scientific production of researchers are peer-review and bibliometric indicators [Da Luz et al., 2008]. In particular, when large-scale evaluations are performed (e.g. over hundreds or even thousands of scientists), bibliometric indicators seem to be the only practicable instrument [Van Raan, 2000]. This is one of the reasons why they are more and more familiar to the scientific community and governmental organizations of several world countries [Orr, 2004].

The purpose of this paper is proposing a structured comparison of several academic research groups within the same discipline, by some Hirsch (*h*) based bibliometric indicators. We precise that the *h*-index is a relatively recent but very popular

* Corresponding author. Tel.: +39 011 5647225; fax: +39 011 5647299.
E-mail address: fiorenzo.franceschini@polito.it (F. Franceschini).

1751-1577/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2010.08.003

indicator that synthetically aggregates two important aspects of the output of a scientist: diffusion/impact – represented by the number of citations per paper – and productivity – represented by the number of papers. A scientist has index h if h of his or her papers have at least h citations each and the other papers have $\leq h$ citations each [Hirsch, 2005]. For more on the advantages/disadvantages of h and the large number of proposals for new variants and improvements, we refer the reader to the vast literature and extensive reviews [Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Burrell, 2007; Bornmann, Mutz, & Daniel, 2008; Egghe, 2010; Franceschini & Maisano, 2010a; Franceschini & Maisano, 2010c; Glänzel, 2006; Van Raan, 2006a; Rousseau, 2008].

For this comparison to be as much exhaustive as possible, five different typologies of h-based indicators are used. They are respectively: the h-spectrum [Franceschini & Maisano, 2010b], the h-index of a research group (h_{GROUP}) [Moed, 2005], the *successive* h-index of a research group (h_2) [Schubert, 2007], the ch-index [Ajiferuke & Wolfram, 2010; Franceschini et al., 2010], and the h-index of single publications (h_{SINGLE}) [Schubert, 2009].

We remark that analysis is focused on research groups, rather than whole universities. In fact, many universities are quite heterogeneous, containing excellent as well as mediocre research groups, and university assessment often fails to give proper credit to those pockets of excellence. Moreover, ranking of individual research groups is a worthwhile but still quite rare endeavour. For example, a comparative study of two academic departments by bibliometric indicators was published almost twenty years ago by Zachos (1991). A more extensive analysis, aimed at investigating statistical properties of bibliometric indicators in general, was conducted by Van Raan (2006b), considering several chemistry and medical research groups. Another study, directed at comparing homologous departments of different universities, has been recently proposed by Lazaridis (2010). In this study, four typologies of department over six Greek universities are considered and ranked using the h-index mean value of academic department members.

Our analysis concerns some Italian academic research groups. In the Italian university system, academic staff is divided in fourteen scientific areas, depending on the discipline (i.e. Mathematical Sciences, Physical Sciences, Biological Sciences, Industrial and Information Engineering, etc.). Scientific areas are in turn divided in a number of sectors concerning more specific subjects of interest [MIUR, 2010]. The attention of this work is focused on the sector of *Production Technology and Manufacturing Systems*. Given a specific university, we define as a (local) research group the whole of the researchers involved in one scientific sector. The five h-based indicators are used for delineating the “bibliometric positioning” of the local research groups investigated. Of course, the number of members of each research group will change depending on the size of each University. The choice of limiting the analysis to a specific scientific sector is due to the fact that it is familiar to the authors and, therefore, it is easier to manage input data (citations statistics) and results. The same procedure could be easily extended to other scientific sectors at national or even international level. We remark that the analysis of two or more scientific sectors would require a normalization to allow comparisons among the results in the different fields, owing to the differences in h-index levels.

The remaining of this paper is organized into three sections. Section 2 illustrates the methodology, with particular attention to the description of bibliometric indicators; Section 3 presents and discusses the analysis results; Section 4 examines pros and cons of the proposed indicators and introduces the concept of bibliometric profile, a graphical tool to support comparison among research groups. Finally, the conclusions are given, summarising the original contribution of the paper.

2. Methodology

Analysis is carried out following a *bottom-up* approach, according to the definition of Noyons and Van Raan (2002). Firstly, researchers associated to the examined research groups are identified by an institutional database [MIUR, 2010]. Then, for each researcher are determined: the number of (1) scientific publications, (2) citations and (3) citers – i.e. different (co-) authors of the citing paper(s) associated to each citation. The information about citers is important, not to establish self-citation patterns, but to construct an indicator (i.e. the ch-index), which evaluates the diffusion/impact of a publication on the basis of the portion of members of the scientific community that are interested in it. Citation statistics are collected using the Google Scholar (GS) search engine. It was decided to use this database because of the greater coverage and since it can be automatically queried through dedicated software applications, such as Publish or Perish or other *ad hoc* applications [Bar-Ilan, 2010; Harzing & van der Wal, 2008]. While determining publications and citations is quite easy and fast, on the other hand, the determination of citers is much more complicated. This procedure was completely automated by an *ad hoc* software application able to query GS automatically. For more information we refer the reader to [Franceschini et al., 2010].

It is worth mentioning that in the Web of Science, the number of citers to a given scientist can be obtained by few clicks, by using the Analyze Results feature. Unfortunately, search is limited to citing papers published on journals/proceedings that are listed by Thomson Scientific. Also, according to many bibliometrists, coverage of Web of Science and Scopus databases – in many fields such as Social Sciences, Computer Science or Engineering Science – is not sufficient [Harzing & van der Wal, 2008]. GS, on the other hand, includes citations to books, book chapters, dissertations, working papers, conference papers, and journal articles published in non-ISI and Open-Access journals. Hence, it probably provides a more comprehensive picture of recent impact. This is one of the reasons why GS has been preferred to other databases for our analysis.

On the other hand, it should not be forgotten that GS is generally less accurate than Web of Science or Scopus. Database mistakes – such as false references, duplications, author ambiguities, etc. – are not so infrequent in GS [Franceschini & Maisano, in pressb; Jacso, 2006]. This is probably due to the automatic generation of the GS data set by scanning and parsing