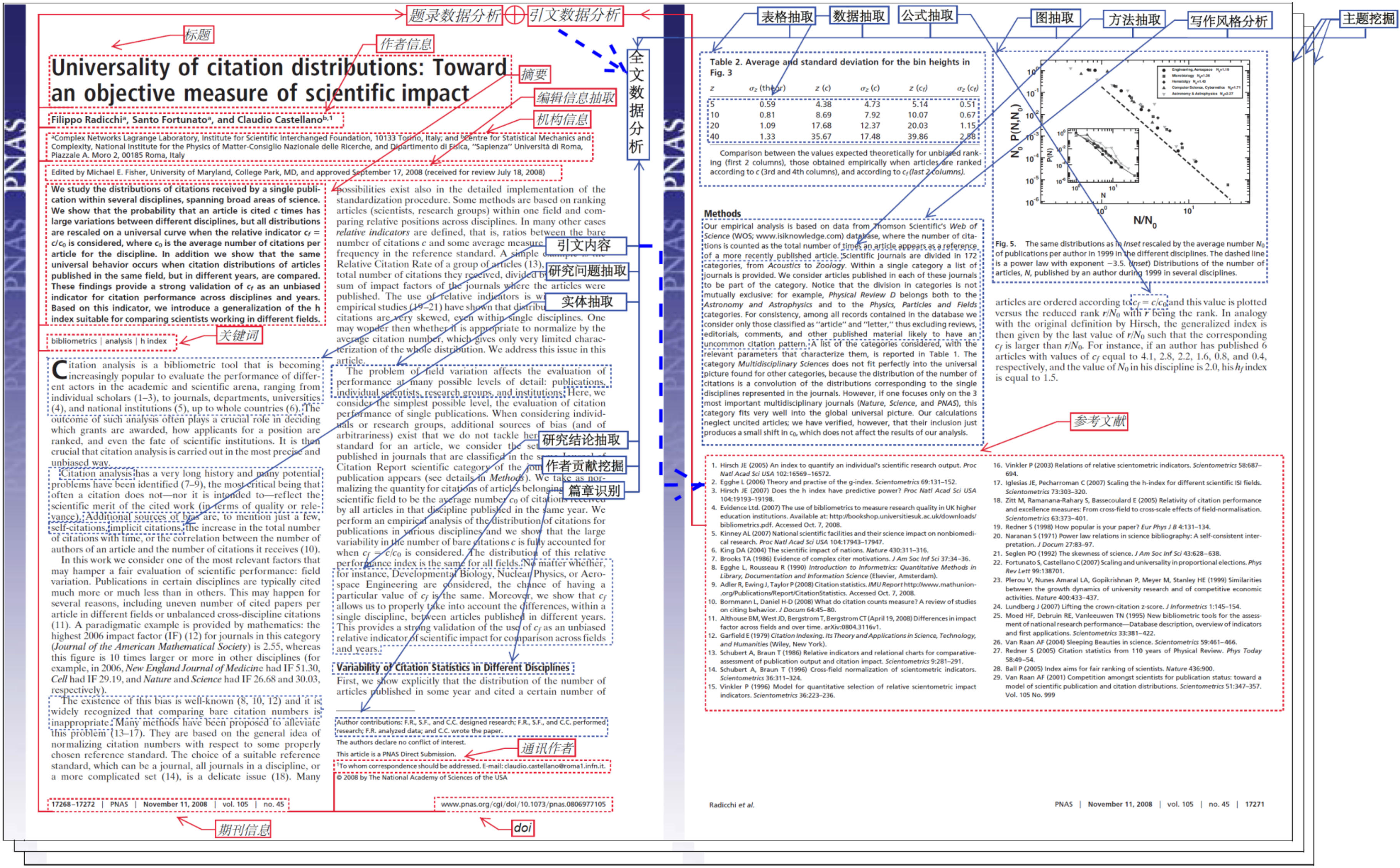

Universality of citation distributions: Toward an objective measure of scientific impact

Filippo Radicchi[a], Santo Fortunato[a], and Claudio Castellano[b,1]

[a]Complex Networks Lagrange Laboratory, Institute for Scientific Interchange Foundation, 10133 Torino, Italy; and [b]Centre for Statistical Mechanics and Complexity, National Institute for the Physics of Matter-Consiglio Nazionale delle Ricerche, and Dipartimento di Fisica, "Sapienza" Università di Roma, Piazzale A. Moro 2, 00185 Roma, Italy

Edited by Michael E. Fisher, University of Maryland, College Park, MD, and approved September 17, 2008 (received for review July 18, 2008)

We study the distributions of citations received by a single publication within several disciplines, spanning broad areas of science. We show that the probability that an article is cited c times has large variations between different disciplines, but all distributions are rescaled on a universal curve when the relative indicator $c_f = c/c_0$ is considered, where c_0 is the average number of citations per article for the discipline. In addition we show that the same universal behavior occurs when citation distributions of articles published in the same field, but in different years, are compared. These findings provide a strong validation of c_f as an unbiased indicator for citation performance across disciplines and years. Based on this indicator, we introduce a generalization of the h index suitable for comparing scientists working in different fields.

bibliometrics | analysis | h index

Citation analysis is a bibliometric tool that is becoming increasingly popular to evaluate the performance of different actors in the academic and scientific arena, ranging from individual scholars (1–3), to journals, departments, universities (4), and national institutions (5), up to whole countries (6). The outcome of such analysis often plays a crucial role in deciding which grants are awarded, how applicants for a position are ranked, and even the fate of scientific institutions. It is then crucial that citation analysis is carried out in the most precise and unbiased way.

Citation analysis has a very long history and many potential problems have been identified (7–9), the most critical being that often a citation does not—nor is it intended to—reflect the scientific merit of the cited work (in terms of quality or relevance). Additional sources of bias are, to mention just a few, self-citations, implicit citations, the increase in the total number of citations with time, or the correlation between the number of authors of an article and the number of citations it receives (10).

In this work we consider one of the most relevant factors that may hamper a fair evaluation of scientific performance: field variation. Publications in certain disciplines are typically cited much more or much less than in others. This may happen for several reasons, including uneven number of cited papers per article in different fields or unbalanced cross-discipline citations (11). A paradigmatic example is provided by mathematics: the highest 2006 impact factor (IF) (12) for journals in this category (*Journal of the American Mathematical Society*) is 2.55, whereas this figure is 10 times larger or more in other disciplines (for example, in 2006, *New England Journal of Medicine* had IF 51.30, *Cell* had IF 29.19, and *Nature* and *Science* had IF 26.68 and 30.03, respectively).

The existence of this bias is well-known (8, 10, 12) and it is widely recognized that comparing bare citation numbers is inappropriate. Many methods have been proposed to alleviate this problem (13–17). They are based on the general idea of normalizing citation numbers with respect to some properly chosen reference standard. The choice of a suitable reference standard, which can be a journal, all journals in a discipline, or a more complicated set (14), is a delicate issue (18). Many possibilities exist also in the detailed implementation of the standardization procedure. Some methods are based on ranking articles (scientists, research groups) within one field and comparing relative positions across disciplines. In many other cases *relative indicators* are defined, that is, ratios between the bare number of citations c and some average measure of the citation frequency in the reference standard. A simple example is the Relative Citation Rate of a group of articles (13), defined as their total number of citations they received, divided by the weighted sum of impact factors of the journals where the articles were published. The use of relative indicators is widespread, but empirical studies (19–21) have shown that distributions of article citations are very skewed, even within single disciplines. One may wonder then whether it is appropriate to normalize by the average citation number, which gives only very limited characterization of the whole distribution. We address this issue in this article.

The problem of field variation affects the evaluation of performance at many possible levels of detail: publications, individual scientists, research groups, and institutions. Here, we consider the simplest possible level, the evaluation of citation performance of single publications. When considering individuals or research groups, additional sources of bias (and of arbitrariness) exist that we do not tackle here. As a reference standard for an article, we consider the set of all articles published in journals that are classified in the same Journal of Citation Report scientific category of the journal where the publication appears (see details in *Methods*). We take as normalizing the quantity for citations of articles belonging to a given scientific field to be the average number c_0 of citations received by all articles in that discipline published in the same year. We perform an empirical analysis of the distribution of citations for publications in various disciplines and we show that the large variability in the number of bare citations c is fully accounted for when $c_f = c/c_0$ is considered. The distribution of this relative performance index is the same for all fields. No matter whether, for instance, Developmental Biology, Nuclear Physics, or Aerospace Engineering are considered, the chance of having a particular value of c_f is the same. Moreover, we show that c_f allows us to properly take into account the differences, within a single discipline, between articles published in different years. This provides a strong validation of the use of c_f as an unbiased relative indicator of scientific impact for comparison across fields and years.

Variability of Citation Statistics in Different Disciplines

First, we show explicitly that the distribution of the number of articles published in some year and cited a certain number of

Author contributions: F.R., S.F., and C.C. designed research; F.R., S.F., and C.C. performed research; F.R. analyzed data; and C.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: claudio.castellano@roma1.infn.it.

www.pnas.org/cgi/doi/10.1073/pnas.0806977105

Table 2. Average and standard deviation for the bin heights in Fig. 3

z	σ_z (theor)	z (c)	σ_z (c)	z (c_f)	σ_z (c_f)
5	0.59	4.38	4.73	5.14	0.51
10	0.81	8.69	7.92	10.07	0.67
20	1.09	17.68	12.37	20.03	1.15
40	1.33	35.67	17.48	39.86	2.58

Comparison between the values expected theoretically for unbiased ranking (first 2 columns), those obtained empirically when articles are ranked according to c (3rd and 4th columns), and according to c_f (last 2 columns).

Fig. 5. The same distributions as in *Inset* rescaled by the average number N_0 of publications per author in 1999 in the different disciplines. The dashed line is a power law with exponent −3.5. (*Inset*) Distributions of the number of articles, N, published by an author during 1999 in several disciplines.

Methods

Our empirical analysis is based on data from Thomson Scientific's *Web of Science* (WOS; www.isiknowledge.com) database, where the number of citations is counted as the total number of times an article appears as a reference of a more recently published article. Scientific journals are divided in 172 categories, from *Acoustics* to *Zoology*. Within a single category a list of journals is provided. We consider articles published in each of these journals to be part of the category. Notice that the division in categories is not mutually exclusive: for example, *Physical Review D* belongs both to the *Astronomy and Astrophysics* and to the *Physics, Particles and Fields* categories. For consistency, among all records contained in the database we consider only those classified as "article" and "letter," thus excluding reviews, editorials, comments, and other published material likely to have an uncommon citation pattern. A list of the categories considered, with the relevant parameters that characterize them, is reported in Table 1. The category *Multidisciplinary Sciences* does not fit perfectly into the universal picture found for other categories, because the distribution of the number of citations is a convolution of the distributions corresponding to the single disciplines represented in the journals. However, if one focuses only on the 3 most important multidisciplinary journals (*Nature*, *Science*, and *PNAS*), this category fits very well into the global universal picture. Our calculations neglect uncited articles; we have verified, however, that their inclusion just produces a small shift in c_0, which does not affect the results of our analysis.

articles are ordered according to $c_f = c/c_0$ and this value is plotted versus the reduced rank r/N_0 with r being the rank. In analogy with the original definition by Hirsch, the generalized index is then given by the last value of r/N_0 such that the corresponding c_f is larger than r/N_0. For instance, if an author has published 6 articles with values of c_f equal to 4.1, 2.8, 2.2, 1.6, 0.8, and 0.4, respectively, and the value of N_0 in his discipline is 2.0, his h_f index is equal to 1.5.

1. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102:16569–16572.
2. Egghe L (2006) Theory and practise of the g-index. *Scientometrics* 69:131–152.
3. Hirsch JE (2007) Does the h index have predictive power? *Proc Natl Acad Sci USA* 104:19193–19198.
4. Evidence Ltd. (2007) The use of bibliometrics to measure research quality in UK higher education institutions. Available at: http://bookshop.universitiesuk.ac.uk/downloads/bibliometrics.pdf. Accessed Oct. 7, 2008.
5. Kinney AL (2007) National scientific facilities and their science impact on nonbiomedical research. *Proc Natl Acad Sci USA* 104:17943–17947.
6. King DA (2004) The scientific impact of nations. *Nature* 430:311–316.
7. Brooks TA (1986) Evidence of complex citer motivations. *J Am Soc Inf Sci* 37:34–36.
8. Egghe L, Rousseau R (1990) *Introduction to Informetrics: Quantitative Methods in Library, Documentation and Information Science* (Elsevier, Amsterdam).
9. Adler R, Ewing J, Taylor P (2008) Citation statistics. *IMU Report* http://www.mathunion.org/Publications/Report/CitationStatistics. Accessed Oct. 7, 2008.
10. Bornmann L, Daniel H-D (2008) What do citation counts measure? A review of studies on citing behavior. *J Docum* 64:45–80.
11. Althouse BM, West JD, Bergstrom T, Bergstrom CT (April 19, 2008) Differences in impact factor across fields and over time. arXiv:0804.3116v1.
12. Garfield E (1979) *Citation Indexing. Its Theory and Applications in Science, Technology, and Humanities* (Wiley, New York).
13. Schubert A, Braun T (1986) Relative indicators and relational charts for comparative-assessment of publication output and citation impact. *Scientometrics* 9:281–291.
14. Schubert A, Braun T (1996) Cross-field normalization of scientometric indicators. *Scientometrics* 36:311–324.
15. Vinkler P (1996) Model for quantitative selection of relative scientometric impact indicators. *Scientometrics* 36:223–236.
16. Vinkler P (2003) Relations of relative scientometric indicators. *Scientometrics* 58:687–694.
17. Iglesias JE, Pecharroman C (2007) Scaling the h-index for different scientific ISI fields. *Scientometrics* 73:303–320.
18. Zitt M, Ramanana-Rahary S, Bassecoulard E (2005) Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics* 63:373–401.
19. Redner S (1998) How popular is your paper? *Eur Phys J B* 4:131–134.
20. Naranan S (1971) Power law relations in science bibliography: A self-consistent interpretation. *J Docum* 27:83–97.
21. Seglen PO (1992) The skewness of science. *J Am Soc Inf Sci* 43:628–638.
22. Fortunato S, Castellano C (2007) Scaling and universality in proportional elections. *Phys Rev Lett* 99:138701.
23. Plerou V, Nunes Amaral LA, Gopikrishnan P, Meyer M, Stanley HE (1999) Similarities between the growth dynamics of university research and of competitive economic activities. *Nature* 400:433–437.
24. Lundberg J (2007) Lifting the crown-citation z-score. *J Informetrics* 1:145–154.
25. Moed HF, Debruin RE, Vanleeuwen TN (1995) New bibliometric tools for the assessment of national research performance—Database description, overview of indicators and first applications. *Scientometrics* 33:381–422.
26. Van Raan AF (2004) Sleeping Beauties in science. *Scientometrics* 59:461–466.
27. Redner S (2005) Citation statistics from 110 years of Physical Review. *Phys Today* 58:49–54.
28. Ball P (2005) Index aims for fair ranking of scientists. *Nature* 436:900.
29. Van Raan AF (2001) Competition amongst scientists for publication status: toward a model of scientific publication and citation distributions. *Scientometrics* 51:347–357.
Vol. 105 No. 999