Impact Factors: the grand delusion

Burg W. Flemming

Received: 15 December 2011 / Accepted: 16 December 2011 / Published online: 30 December 2011
© Springer-Verlag 2011

In 2012, Geo-Marine Letters (GML) enters its 32nd year since its inception by Arnold Bouma in 1981. Although GML has remained a ‘small’ journal, it nevertheless has grown from originally four to presently six issues per year. In the course, the design of the cover page has changed twice, the current one dating back to the year 2000. Seeing that the past decade or so has seen numerous technological advances which have revolutionized the marine sciences, notably spectacular advances in digital seabed imaging, GML will appear with a striking new cover background image beginning with issue 1 of volume 32, 2012.

The past few years have also seen a dramatic development in the highly controversial usage of so-called *impact factors* and *citation indices* as parameters which, together with others, are purported to reflect quality differences between journals and performance differences between scientists. The two indices are closely interlinked, being based on the number of citations a paper and, hence, cumulatively a journal has received over two consecutive years. One should note here that we are dealing with numbers, i.e. quantities, and that no amount of sophisticated statistics can per se convert a quantity into a quality, the latter in the present context always being a conjectural interpretation. By some obscure process, this bibliometric number game (Adam 2002) has unfortunately been reduced to the simplistic message that the higher the impact factor, the better the journal, or the higher the citation index (or derivatives thereof such

as the *h*-index; Hirsch 2005), the more prolific the scientist. This rating practice has been highly controversial from the start, and the proliferation of publications on this issue suggests that the dilemma is still far from being resolved. In spite of this, and probably also in ignorance of earlier warnings (e.g. Seglen 1997; Amin and Mabe 2000; Lawrence 2002), science managers, funding organisations and institute heads have regrettably institutionalised these indices, accepting them as the long awaited and supposedly objective rating indicators for journals and particularly scientists.

As editor, it was of considerable interest to me to probe particular parameters possibly determining the impact factor of a small journal like Geo-Marine Letters relative to larger journals in the same general field of research. An obvious line of approach was to assess any influence of journal size or, more specifically, the number of articles or pages published per year. The results covering eleven journals for the years 2009 and 2010, representing the years 08/09 and 09/10 respectively, are illustrated in Fig. 1. If the impact factor were an objective criterion of quality, then it should be independent, or nearly so, of journal size, i.e. there should be no or only a weak correlation between the two parameters. The opposite is true – between 75% (2010) and 85% (2009) of the correlation can be explained by journal size. An even higher correlation is found when using page number in place of article number. Also striking is the interannual variation, which shows that the impact factor of a journal can fluctuate quite strongly from year to year.

This basically means that, in the present case, all journals on or close to the regression line perform equally well, irrespective of their impact factors. Note that Geo-Marine Letters, the smallest journal in this group, lies on the

B. W. Flemming (✉)
Marine Research Department, Senckenberg Institute,
Suedstrand 40,
26382 Wilhelmshaven, Germany
e-mail: bflemming@senckenberg.de

Fig. 1 Impact factor versus number of articles for some well-known sedimentology and marine science journals for the years 2009 and 2010. Note the strong relationship between the two parameters and also the substantial interannual variation of some journals

regression line in both years. One could construe that journals plotting above the regression line perform a little better, those below a little worse. As pointed out above, however, there is considerable interannual variability, a fact that argues against inherent differences in quality and more in favour of papers ‘coincidentally’ receiving higher or lower than average citations in a particular year. Nevertheless, consistently higher performance could arguably be a quality criterion. This rather simple analysis shows that, primarily, impact factors and citation indices have little to do with quality but very much with relationships between purely technical parameters. This applies also to a variety of other, related parameters such as the size of the community dealing with a specific subject matter, or the mean number of authors per article (e.g. Amin and Mabe 2000).

Moreover, I noted other types of journal groupings in which, for similar numbers of papers, the journals had much higher impact factors, e.g. review journals or journals covering a large number of unrelated fields. This suggests that there are distinctly different journal categories, each apparently governed by a different set of control factors. Consequently, the mixing of such categories would be akin to comparing apples with pears. Although there have been attempts to normalise such differences (e.g. Pudovkin and Garfield 2004), an objective solution would require enormous efforts to neutralise confounding effects of, for example, different citation habits among journals, self-citations, research group inter-citations, buddy citations, peer citations, citation bias for or against controversially debated scientific issues, as well as disregard of foreign-language (non-English) publications. Even the citations within a single article fall into

several different categories (e.g. key issue citations, example citations, methods citations, critique citations) which would have to be individually ranked by some ultimately noncontroversial key.

This discourse could be continued by raising questions about how to cope with manipulations by editors aiming to improve their journal rankings and by authors to increase their personal ratings. In addition, there is clear evidence that the citation hype is having a negative impact on science itself. This can be seen in that small (but potentially important) research fields are being marginalised, that young scientists are being driven into popular research fields where they can maximise their citation indices in the shortest possible time, that datasets are being macerated to glean as many papers as possible, and that increasing data duplication is being observed in publications with only marginally different messages. While impact factors and citation indices may be valid subject matters in their own right for scientific research in bibliometric statistics, scientometrics, information theory, sociology, marketing, etc., their use for quality ranking of journals and scientists is quite inappropriate considering the technicalities and uncertainties addressed above. As currently applied, they are nothing more than a grand delusion.

As far as Geo-Marine Letters is concerned, I am quite happy to see that, with an impact factor of currently 1.73, it is on par with most other journals in this category.

Burg W. Flemming
Editor-in-Chief

References

- Adam D (2002) The counting house. *Nature* 415:726–727
- Amin M, Mabe M (2000) Impact factors: use and abuse. *Perspectives in Publishing* no 1, Elsevier, Amsterdam
- Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci* 102:16569–16572
- Lawrence PA (2002) Rank injustice: the misallocation of credit is endemic in science. *Nature* 415:835–836
- Pudovkin AI, Garfield E (2004) Rank-normalized impact factor: a way to compare journal performance across subject categories. In: *Proc 67th Annu Meet American Society for Information Science & Technology*, 13–18 November 2004, Providence, Rhode Island, vol 41, pp 507–515
- Seglen PO (1997) Why the impact of journals should not be used for evaluating research. *British Med J* 314:498–502