Science Progress (2018), 101(3), 293–310
Paper 1800279 https://doi.org/10.3184/003685018X15337564592469

Bibliometrics in a digital age: help or hindrance

KAREN BLAKEMAN

Karen Blakeman graduated from the University of Birmingham with a degree in Microbiology and later obtained a diploma in Information Science from London City University. Prior to setting up her company, RBA Information Services, she worked as a microbiologist and then spent ten years in the pharmaceutical and health care industry as an information scientist. She moved to the international management consultancy group Strategic Planning Associates before becoming a freelance consultant in 1989. She currently provides training and advice on search, social media and collaborative tools and has written many articles and books on information search techniques. She also runs workshops and has spoken at many international seminars and conferences. She is an Honorary Fellow of CILIP (the Chartered Institute of Library and Information Professionals) and in 2002 she received the Information World Review Information Professional of the Year award. E-mail: karen.blakeman@rba.co.uk

ABSTRACT

Bibliometrics are a range of techniques and quantitative measures that provide an analysis of written publications such as books and articles, and which assess the impact of research outputs. They are commonly applied to individual authors in the form of citation metrics but can also be used to assess the influence of research groups or even entire institutions. With the increased importance of social media as a means of communicating and publicising research findings, additional alternative measures of impact (altmetrics) are now being used. In addition to analysing the reach of a research output, bibliometrics can also be used as search tools to identify related and updated research, author networks and connections between institutions. This review summarises the range of tools and services that are available, their advantages and disadvantages, and some of the challenges and issues presented by the existence of multiple digital versions of research outputs.

Keywords: *bibliometrics, altmetrics, citation indexes,* h-*index, impact factors, Scopus, Web of Science, Google Scholar*

1. Introduction

The term bibliométrie was used by Paul Otlet in 1934 in his book *Traité de Documentation*[1,2] and defined by him as the measurement of all aspects related to the publication and reading of books and documents[3]. The English form bibliometrics was proposed by Alan Pritchard in a paper published in 1969 entitled 'Statistical bibliography or bibliometrics?' and defined by him as 'the application of mathematics and statistical methods to books and other media of communication'[4]. In his conclusion he says, 'it is to be hoped that this term BIBLIOMETRICS will be used explicitly in all studies which seek to quantify the processes of written communication and will quickly gain acceptance in the field of information science'.

Originally, bibliometric methods were used to trace relationships between journal articles, the most well-known tool being citation analysis. This involves looking at the list of references appended to a document that provide supporting evidence and background to the research, and which can be used to assess the merit of the document. It also allows the reader to search forward in time from a known article to more recent publications that cite the known article and which may have additional and updated information. Historically, there have been many citation indexes covering various subjects. Legal citation indexes, in particular, were known in the 18th and 19th centuries and provided information on whether a case had been overturned, reaffirmed, questioned or cited in subsequent cases[5]. They continue to be published mainly in electronic form to this day.

In the 1960s, the Science Citation Index (SCI) began publication. The principles and applications of the index had been outlined by Eugene Garfield in 1955[6,7], and in 1965 Derek J. de Solla Price described how the network of citations could be used to describe the nature of the total world network of scientific papers by 'linking each published paper to the other papers directly associated with it'[8]. It was suggested that this technique could be used as an alternative to the practice of separately counting the numbers of papers published and citations as a quantitative measure of the impact or influence of an author, research group or journal. The SCI was produced and officially launched by the Institute for Scientific Information (ISI) in 1964. It is a component of the Web of Science, which was previously owned by Thomson Reuters but is now part of Clarivate Analytics (https://clarivate.com/). It has become one of the key tools for searching and analysing the peer-reviewed research literature along with Scopus (https://www.scopus.com/), owned by Elsevier, and Google Scholar (https://scholar.google.com/).

Moving forward to 1998, a new search engine that concentrated on freely available information on the web appeared on the scene[9]. It was called Google and quickly became one of the more popular web search tools because the results it returned to the user were more relevant than those of other similar services at that time. In common with other search engines, Google uses, in part, the frequency with which terms are mentioned in documents together with the terms' location within those documents to determine the relevance of articles with respect to the user's query. Google was unique though, in using what it called PageRank as a major component of its sorting algorithm[10]. PageRank looks at the number and the 'importance' of links that point to a page in order to determine that page's place in the results list[11]. It is, in essence, a form of citation analysis. It is still part of Google's main sorting algorithm and is also used to a certain degree in its academic search engine Google Scholar. Google Scholar was developed further to include information on author, article and journal citations as part of its database.

More changes came with the advent of social media, bringing together traditional peer-reviewed literature and a means of publicising research to a much wider audience. This, combined with the rise of free and open access to some papers, has given researchers new opportunities for increasing awareness of their work not only amongst fellow researchers but also to the general public. Blogs, Twitter, Facebook and news outlets are just some of the platforms that are now commonly used for promoting publications and research, and as a result a new branch of bibliometrics called altmetrics has arisen[12].

Bibliometrics in its various forms has become established as an important part of the assessment of academic research. Originally put forward as a quantitative measure to complement qualitative assessment of research output, many in academia are concerned that too much emphasis is now placed on it. Some of these concerns are reviewed later in this article. Nevertheless, bibliometrics are here to stay so it is important to understand how they are calculated and by whom, and the problems that can arise if they are used indiscriminately.

2. Scopus, Web of Science and Google Scholar

The three major databases that offer bibliometric analyses and citation searching are Scopus, Web of Science and Google Scholar. Scopus and Web of Science are subscription services. Google Scholar is free to search and provides free citation analyses, but the searcher may have to pay to view articles. All three have different coverage in terms of the type of research outputs and journal titles. They have their own specific, bibliometric measures as well as the standard metrics such as the author level h-index, and they sometimes calculate individual metrics in different ways.

Scopus is owned by Elsevier and claims to have the largest abstract and citation database of the world's research output in the fields of science, technology, medicine, social sciences, and arts and humanities[13]. It covers books, journals, conference proceedings and trade journals. In addition to author level metrics, Scopus provides journal impact factors, SCImago Journal Rank (SJR) and Source Normalized Impact per Paper (SNIP).

Web of Science, owned by Clarivate Analytics and previously known as Web of Knowledge, incorporates the SCI that was mentioned in the introduction. It provides citation search facilities and gives access to multiple databases that reference cross-disciplinary research.

Google Scholar is a free search engine that is separate from the main Google web search tools, and indexes the full text of scholarly literature across a wide range of publishing formats and disciplines. It is worth looking at Google Scholar in more detail as it indexes more than just the peer-reviewed literature, often includes multiple versions of a paper (for example author manuscript, preprint, institutional repository copies, final published version, pirated copies), and has a high level of usage because it is free to search. It was released as a test database in November 2004 and includes: most peer-reviewed online academic journals; books; conference papers; theses and dissertations; preprints; abstracts; technical reports, including court opinions (USA only); and patents[14].

Most scholarly publishers allow Google free access to their publications for indexing purposes, including subscription only content. This enables anyone to freely search everything within the Google Scholar database. However, when it comes to viewing individual articles the user may find that they have to pay for access. In some cases, the paper may be free as an open access article or Google may offer a free institutional repository copy.

The level of coverage and number of titles in Google Scholar has frequently been questioned. It does not publish a list of journals that it crawls nor of the publishers that are included in its database, and the update frequency is not disclosed. There have however been numerous studies that suggest Google Scholar's coverage of the sciences and social sciences is comparable with other academic databases.

One study estimated that Google Scholar can find almost 90% of all scholarly documents on the Web written in English[15]. Another study looking at the coverage of highly-cited documents in Google Scholar, Web of Science and Scopus found significant differences between the databases in some disciplines[16], suggesting that all three might need to be searched to ensure that all relevant literature on a topic is found. It analysed 2,515 highly-cited documents published in 2006 in the Google Classic Papers collection[17] and looked at whether or not the documents in the collection were also covered in Web of Science and Scopus, and if the citation counts were similar. It found that a significant percentage of highly-cited documents in the Social Sciences and Humanities (8.6–28.2%) were not included in Web of Science and Scopus. In the Natural, Life, and Health Sciences they found that the proportion of missing highly-cited documents in Web of Science and Scopus was much lower. It is important to be aware of the difference in the level and the nature of the coverage of Scopus, Web of Science and Google Scholar as it affects the bibliometrics that they each calculate, and in particular the author level *h*-index metric.

As well as the peer-reviewed literature, Google Scholar includes web pages, blog postings, and self-published articles in its index. If it finds a document on the web that is structured and organised in a way similar to that of a journal article it is added to the database. Google Scholar has been criticised for not vetting journals and that it includes predatory and hijacked journals in its index. Predatory open-access publishing is an exploitative open-access publishing business model that involves charging publication fees to authors without carrying out the editorial and publishing services that they claim to provide and which are associated with legitimate journals[18]. A hijacked journal is a legitimate journal, the title of which has been taken and used on a fake website. The purpose is to fool authors into submitting papers to the fake site rather than the original, legitimate publication. Although Google Scholar does indeed have such questionable papers on its system that can affect the calculation of some bibliometrics, recent studies have shown that these types of publications are occasionally to be found in Web of Science and Scopus as well[19,20].

3. Journal bibliometrics and impact factors

Journal level bibliometrics – one of the more well known being the impact factor – are important in that they influence decisions on where to submit articles for publication. Authors are encouraged to publish in a journal with as high an impact factor as possible as this is increasingly being used as a factor in assessing the overall performance of a research group or institution and, ultimately, can affect the level of funding that they receive.

A journal's impact factor is based on how often articles published in that journal are cited by articles in other journals. The higher a journal's impact factor, the more often articles in that journal are cited by other articles. The impact factor can therefore give an approximate indication of how prestigious a journal is in its field. Not all journals have impact factors, and the importance of impact factors varies between disciplines.

Web of Science's Impact Factor is based on what it calls 'citable documents', which are articles and reviews. It can be found in the Web of Science Journal Citation Reports (JCR) (https://clarivate.com/products/journal-citation-reports/) and is the average number of times articles from a journal published in the previous two years have been cited in the JCR year.

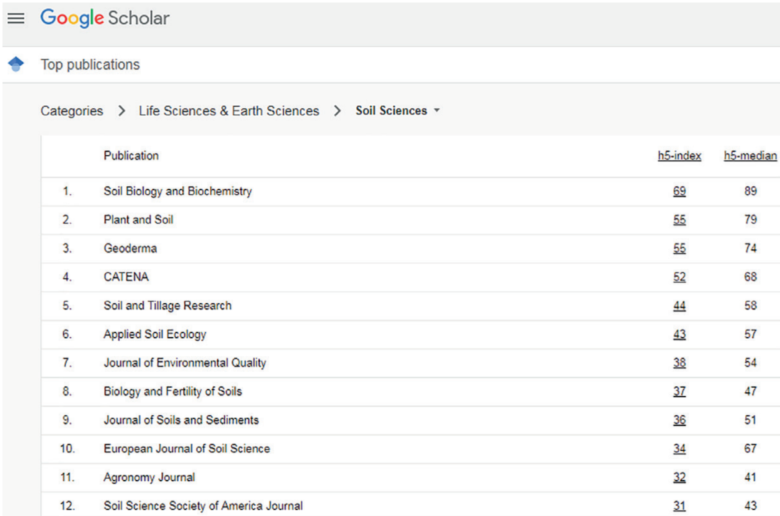
Scopus has a similar metric called CiteScore[21]. It is calculated for the current year based on the number of citations received by a journal in that year for the documents published over the past three years, and divided by the documents indexed in Scopus published in those three years. CiteScore includes all document types indexed by Scopus, including articles, reviews, letters, notes, editorials and conference papers.

Scopus also provides journal SJR and SNIP. SJR ranks publications by weighted citations per document. Citations are weighted more or weighted less depending on the source they come from. Subject field, quality and reputation of the journal all have a direct effect on the value of a citation[22,23]. SNIP measures contextual citation impact by weighting citations based on the total number of citations in a subject field[24]. It was created by Professor Henk Moed at the Centre for Science and Technology Studies (CTWS), University of Leiden[25]. The impact of a single citation is given a higher value in subject areas where citations are less likely, and *vice versa*. For example, journals in Mathematics, Engineering and Social Sciences tend to have higher values than titles in Life Sciences. This enables direct comparison of sources in different subject fields, the difference in journals' SNIP being due to the quality of the journals and not the different citation behaviour between subject fields.

Google has its own list of metrics (Google Scholar Metrics or Scholar Metrics) based on articles published in the last five years. For example, Scholar metrics currently covers articles published between 2013 and 2017 inclusive, and the metrics are based on citations from all articles that were indexed in Google Scholar in July 2018 (https://scholar.google.com/scholar/metrics.html#overview). Specifically excluded from the metrics are:

- court opinions, patents, books, and dissertations
- publications with fewer than 100 articles published between years being covered
- publications that received no citations to articles published between the years being covered

There is a top 100 list of journals across all disciplines ordered by their five-year *h*-index (*h*5-index) and *h*5-median (https://scholar.google.com/citations?view_op=top_venues&hl=en). The *h*5-index is the *h*-index for articles published in the last five complete years. It is the largest number *h* such that *h* articles published in, for



Google Scholar

Top publications

Categories > Life Sciences & Earth Sciences > Soil Sciences

	Publication	h5-index	h5-median
1.	Soil Biology and Biochemistry	69	89
2.	Plant and Soil	55	79
3.	Geoderma	55	74
4.	CATENA	52	68
5.	Soil and Tillage Research	44	58
6.	Applied Soil Ecology	43	57
7.	Journal of Environmental Quality	38	54
8.	Biology and Fertility of Soils	37	47
9.	Journal of Soils and Sediments	36	51
10.	European Journal of Soil Science	34	67
11.	Agronomy Journal	32	41
12.	Soil Science Society of America Journal	31	43

Figure 1 *Screenshot of Google Scholar Metrics showing the top publications for the category Life Sciences & Earth Sciences, subcategory Soil Sciences.*

example, 2013–2017 have at least *h* citations each. If the *h*5-index of a journal is 185 then 185 articles published in the journal between 2013 and 2017 have received at least 185 citations for each article. The *h*5-median is the median number of citations for the articles that make up its *h*5-index.

To see which articles in a publication were cited the most and who cited them, the reader clicks on the *h*5-index number to view the articles and the citations underlying the metrics. It is also possible to explore publications in a broad category of research by selecting it from the drop-down Categories menu at the top of the publications list, and to drill down to more specific research areas by selecting a subcategory. Figure 1 shows the list of publications in the category of Life Sciences & Earth Sciences, subcategory Soil Sciences.

4. Author level bibliometrics

4.1 h-*index*

The *h*-index (Hirsch index or Hirsch number) is the most well-known author-level metric and attempts to measure both the productivity and the citation impact of the publications of a researcher. It was first proposed in 2005 as a measure of characterising the scientific output of a researcher by Jorge Eduardo Hirsch, a physicist at the University of California, San Diego[26]. The index is based on the set of a scientist's most cited papers and the number of citations that they have received in other publications. Hirsch defines the index as follows: 'A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each.' For example, if a researcher has 20 papers, each of which has at least 20 citations, their *h*-index is 20. The index can also be applied to a research group of scientists, a department, university, country, or even used to assess the whole content of a journal[27].

Scopus, Web of Science and Google Scholar can all calculate an *h*-index but the values usually differ because each database covers different publications, different types of output, and different ranges of years. Scopus and Web of Science automatically generate an *h*-index for any author on request. Google Scholar will only show an *h*-index for an author who has created a profile for themselves on Scholar and 'claimed' their research papers[28]. The name of an author who has a Google Scholar profile is underlined in green in the article snippet shown on the search results page. Figure 2 shows how authors with and without Google Scholar profiles are displayed.

The *h*-index is useful for: comparing researchers with similar career lengths; comparing researchers in a similar field, subject, or department; and those who publish in the same journal categories. The *h*-index is not useful for: comparing researchers from different fields, disciplines, or subjects; comparing researchers at an early stage in their career with established scientists; assessing departments and subjects where research output is typically books or conference proceedings.

Hirsch, himself, warns that a single number can never give more than a rough approximation to an individual's profile and that many other factors should be considered:

'There will be differences in typical *h* values in different fields, determined in part by the average number of references in a paper in the field, the average number of papers produced by each scientist in the field, and the size (number of scientists) of the field. Scientists working in nonmainstream areas will not achieve the same very high *h* values as the top echelon of those working in highly topical areas. Although I argue that a high *h* is a reliable indicator of high accomplishment, the converse is not necessarily always true. There is considerable variation in the skewness of citation distributions even within a given subfield, and for an author with a relatively low *h* that has a few seminal papers with extraordinarily high citation counts, the *h* index will not fully reflect that scientist's accomplishments. Conversely, a scientist with a high *h* achieved mostly through papers with many co-authors would be treated overly kindly by his or her *h*.'

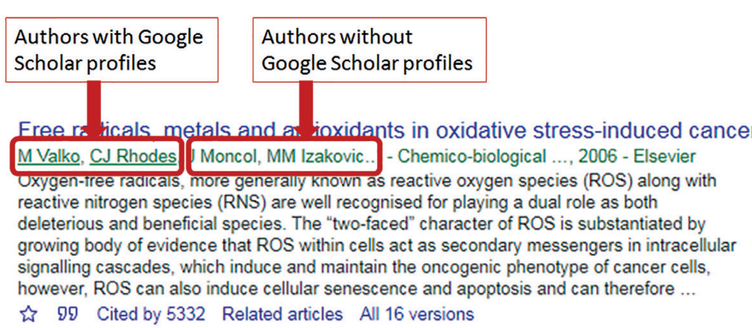


***Figure 2** Screenshot of Google Scholar search result showing how authors with and without Google Scholar profiles are displayed.*

4.2 *i10-index*

The *i*10-Index was created by Google Scholar as an alternative to the *h*-index and is displayed on an author's My Citations page (see Figure 3). It is the number of publications with at least 10 citations. It is a very simple measure, easy to use and calculate, but is only used by Google Scholar.

4.3 g-*index*

The *g*-index is another index for quantifying productivity in science, based on an author's publications record. This was put forward by Leo Egghe in 2006[29]. The index is calculated based on the distribution of citations received by a given researcher's publications, such that given a set of articles ranked in decreasing order of the number of citations that they received, the *g*-index is the unique largest number such that the top *g* articles received together at least g^2 citations.

The following description and example is taken from Harzing's online Publish or Perish tutorial[30]:

'A *g*-index of 20 means that an academic has published at least 20 articles that combined have received at least 400 citations. However, unlike the *h*-index these citations could be generated by only a small number of articles. For instance an academic with 20 papers, 15 of which have no citations with the remaining five having respectively 350, 35, 10, 3 and 2 citations would have a *g*-index of 20, but a *h*-index of 3 (three papers with at least 3 citations each).'

Harzing points out the *h*-index and *g*-index are limited by the number of papers one publishes. Both indices – and especially the *g*-index – favour academics that publish more papers, provided they are cited at least moderately well. They are not a reliable way to assess the impact of those who have published only one or two ground-breaking papers and not published any further highly cited work. For these academics, the total number of citations might be a more appropriate metric.

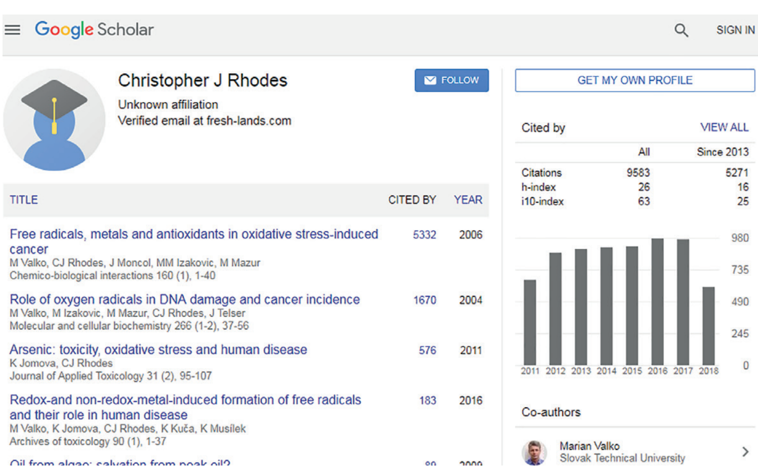


***Figure 3** Screenshot of a Google Scholar author profile showing number of citations,* h*-index and* i*10-index.*

5. Alternative metrics

Several developments over the last 10–15 years have led to the adoption of quite a different approach to measuring the impact of research outputs. One is the increase in the number of articles that are published as open access, enabling free access to content that was previously only available to readers *via* subscription or on payment of a per-article fee. A second development is that open access articles are more likely to be read and discussed in the news and on social media platforms. Social media also provides an author with a range of options for sharing their research and promoting it to a much wider audience than was previously possible. The number of document views and downloads together with social media 'mentions', 'likes' and 'shares' as a means of measuring the reach or impact of a piece of work, has led to the creation of a whole new branch of bibliometrics called altmetrics.

Altmetrics are non-traditional bibliometrics and complement the more traditional citation metrics such as impact factors and *h*-index. The term altmetrics was proposed by Priem *et al*. in 2010[12]. Although altmetrics are often thought of as metrics about articles, they can be applied to people, journals, books, data sets, presentations, videos, source code repositories and web pages.

Some document collections, such as PLOS (Public Library of Science) (https://www.plos.org/), automatically record and make public the number of times an article has been viewed or downloaded, the number of citations across different platforms, and the level of discussion on social media (https://www.plos.org/article-level-metrics). Figure 4 is an example of PLOS article metrics showing the number of views and downloads of an article. Figure 5 shows the number of citations for the same article, where it has been discussed and tweets that refer to it.

This approach has been developed further by a company called Altmetric (https://www.altmetric.com/)[31]. Many journals and articles now have an Altmetric 'badge' embedded in their web pages that show the level of interest that has been generated by an article in the news and social media. Figure 6 shows some of the Altmetric data relating to an article in Nature on perovskite-sensitised cells and lists the number of mentions on Twitter, blogs, Facebook, Google+ and coverage by news outlets.

For articles that do not already have data associated with them, Altmetric provides a browser bookmarklet that can be used to bring up the data that is available (Figure 7). Clicking on any of the elements in the list takes the reader to an Altmetric page for that article showing detailed information on the 'citations'.

Yet another view of citations and 'influence' is provided by Semantic Scholar (https://www.semanticscholar.org/). Semantic Scholar is a free academic search engine created by AI2 Allen Institute for Artificial Intelligence (https://allenai.org/). It uses machine learning to analyse publications and claims that 'the resulting influential citations, images and key phrases allow our engine to "cut through the clutter" and give you results that are more relevant and impactful to your work'[32]. At present, it only covers publications published in English, and research related to computer science and biomedicine that are open access or freely available. There are some articles in their database that are behind publishers' paywalls but the information that can be generated on these is limited.

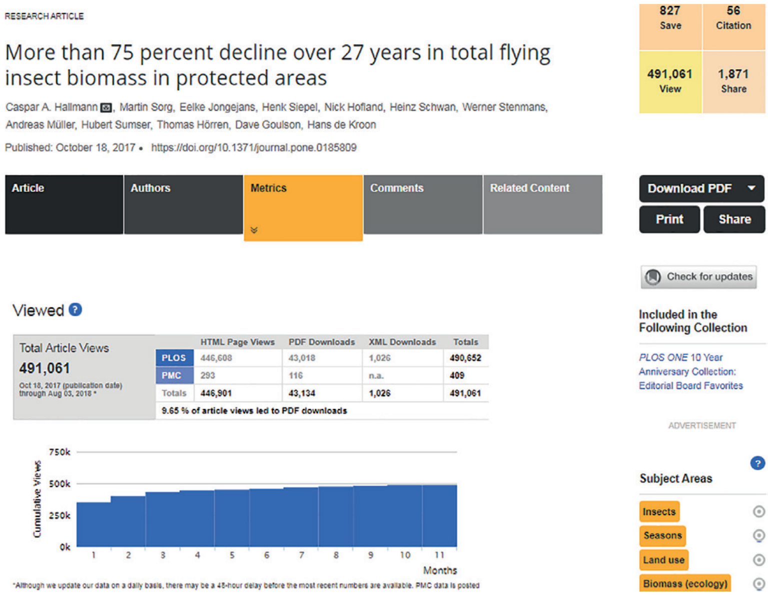


Figure 4 *Screenshot of PLOS article metrics showing number of views and downloads of an article.*

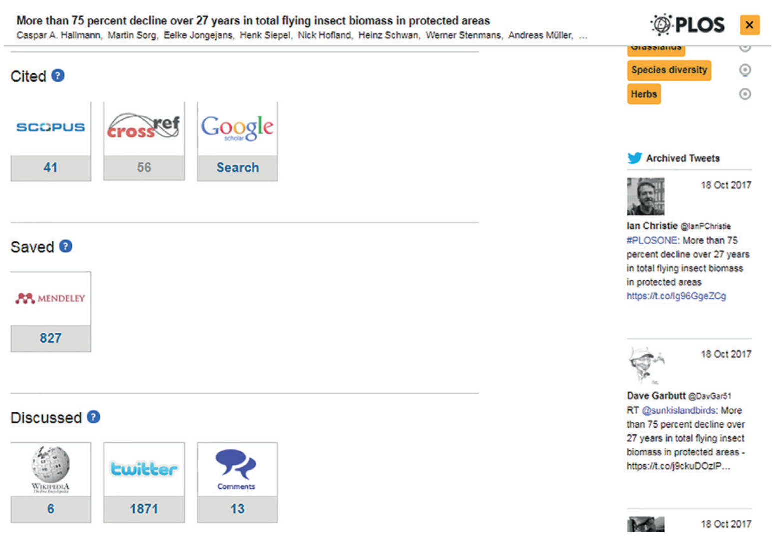


Figure 5 *Screenshot of PLOS article metrics showing the number of citations an article has received, discussions and tweets.*

The content and layout of the results from a Semantic Scholar search is different to that of other free search tools such as Google Scholar. It does not show the *h*-index of authors or journal impact factors but does list the number of citations that it has found in its collection. It also uses a statistical model to generate an estimate of total citations to account for citations outside of their database (Figure 8).

The total citation estimate is often an underestimate. For the paper shown in Figure 8 Semantic Scholar has counted 11 citations in its own database and estimates

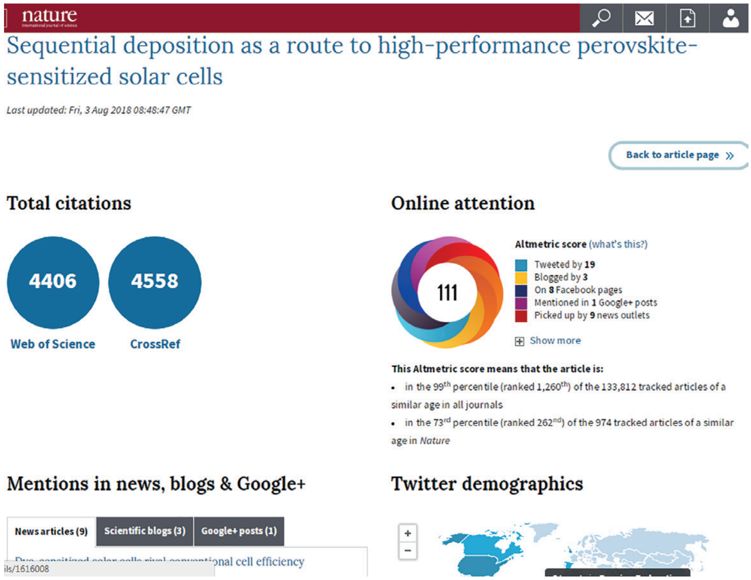


Figure 6 *Screenshot of Altmetric data embedded on an article webpage in Nature.*

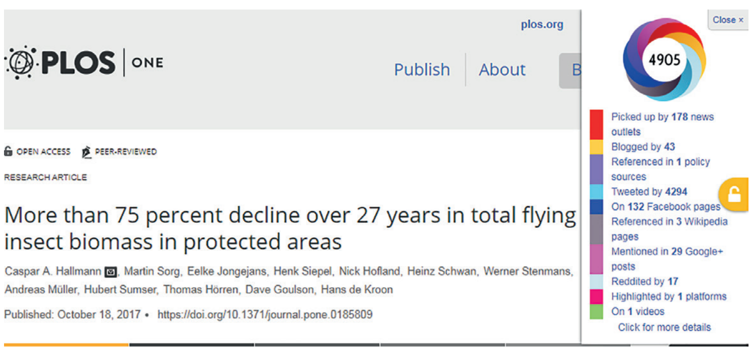


Figure 7 *Screenshot of Altmetric bookmarklet listing social media references to a paper.*

that there are probably a total of 20. At the time of writing this review, Scopus gave a citation count of 41 and Google Scholar 'about 100'.

More interesting are Semantic Scholar's author pages and maps. These are automatically generated and include a map showing authors who have most influenced an author and authors that have been most influenced by that same author (Figure 9).

This is an easy to use tool that can help expand a search and explore other authors' work in related fields. It is not clear though, how the 'influence score' referred to on the author pages is calculated and Semantic Scholar's explanation is somewhat vague (https://www.semanticscholar.org/faq#influence-score):

'The influence score measures the impact of one author's publications on another author's work. The number is relative to the author whose profile is currently being viewed. The score is based on a weighted combination of citations and Highly Influential Citations.'

Similarly, on 'Highly Influential Citations' it says (https://www.semanticscholar.org/faq#influential-citations):

'Semantic Scholar identifies citations where the cited publication has a significant impact on the citing publication, making it easier to understand how publications build upon and relate to each other. Influential citations are determined utilizing a machine-learned model analyzing a number of factors including the number of citations to a publication, and the surrounding context for each citation.'

Despite the lack of detailed information on how the information is generated, which to a certain degree is to be expected given that machine learning is being used, Semantic Scholar is a useful additional tool for analysing research outputs and identifying links between both articles and authors.

6. Problems with bibliometrics

Journal, article and author level metrics have become an important part of the academic research process showing the reach and popularity of specific articles, authors, and publications. However, the data upon which they are based are often incomplete and dependent on the number and type of publications included in the database. Data that were collected many years ago had to be entered into a database by hand with the inevitable typographical errors and although automation should have made the compilation of metrics considerably easier, the lack of human intervention and checking at the entry level has led to a different problem: that of multiple variants of an author's name and consequently the creation of multiple author IDs all with their own widely differing metrics.

Figure 8 *Screenshot of Semantic Scholar showing article citations and estimated total number of citations for an article.*

Figure 9 *Screenshot of Semantic Scholar author map showing author 'influences' and influential citations.*

This lack of robustness of the *h*-index in particular was examined in detail by Teixeira da Silva and Dobranszki[33]. Barnes in 2016 went even further in his paper 'The construct validity of the *h*-index'[34]. He concludes:

'... that the *h*-index fails any test in terms of construct validity implies that the widespread use of this metric within the higher education sector as a management tool represents poor practice, and almost certainly results in the misallocation of resources.'

And

'The implication is that universities, grant funding bodies and faculty administrators should abandon the use of the *h*-index as a management tool. Such a change would have a significant effect on current hiring, promotion and tenure practices within the sector, as well as current attitudes towards the measurement of academic performance.'

Another problem is that open access, pre-print servers and institutional repositories have resulted in multiple versions of papers being available across the web. These may be legitimate copies of the final version, for example in an author's institutional repository, but it is not uncommon to find copies of a paper at various stages in the publication process: for example, author's original manuscript, 2nd and 3rd drafts, first revision, second revision, preprint, postprint, final published version. When researching the literature it is essential that the final, published work is used and cited rather than an author's manuscript that has not been peer-reviewed or corrected, or a preprint without the final edits. For example, the author of this review found several versions and formats of Pritchard's 1969 paper[4], one of which appeared to be a copy and paste of the text into a document without any bibliographic details apart from the title and author. One wonders how many errors were introduced by the process or if any text was omitted.

The issue is most evident with Google Scholar's 'all versions' link that is associated with most of the entries in its search results. Figure 10 shows the top results of a Google Scholar search on insect biomass decline with the number of 'versions' available for each paper. For the second article in the list (Numerical response of lizards to aquatic insects and short-term consequences for terrestrial prey) Google Scholar has 23 versions. Clicking on the link brings up all versions of the article (see Figure 11). The first one is the final version published within the Wiley Online Library. The next two are identical to the Wiley article and are in the authors' university institutional repositories. The fourth is on a US government website and gives the title, abstract, bibliographical details with a link to the Wiley copy. In fact, in this case all 'versions' are either copies of the final publication or links to it, but that is not always so.

It is suggested by some that preprints and repository copies are being given prominence by Google Scholar and are cited in articles even when the article has been published in a journal. As well as the possibility that earlier versions may contain inaccuracies, there is also the concern that these citations are not included in journal, article and author metrics. This was discussed in some detail in a blog posting and comments on Scholarly Kitchen[35]. A number of reasons were put forward to account for this trend:

- a researcher may have downloaded an early preprint version and continued to cite that
- copying or reusing old references
- Google Scholar may be preferentially sending people to the preprint
- the final published article is on a service behind a paywall which the researcher cannot access

Whatever the reason, Davis[35] points out that 'A citation is much more than a directional link to the source of a document. It is the basis for a system of rewarding those who make significant contributions to public science. Redirecting citations to preprint servers not only harms journals, which lose public recognition for publishing important work, but to the authors themselves, who may find it difficult to aggregate public acknowledgements to their work.'

Apart from the technical issues surrounding the implementation and use of bibliometrics, there is growing concern that they are starting to dominate and control the research process. The *h*-index for example is used to assess a scientist's work, is used in recruitment and is often a factor in assigning funding and grants to individuals, research groups and institutions. Publishing in journals with high impact factors and maintaining a good *h*-index have become a part of academic life.

Bibliometrics are now used as part of research assessment exercises of academic output. In the UK's Research Excellence Framework 2014, citation data from Scopus was used by 11 of the 36 sub-panels[36] as additional information about the academic significance of submitted outputs. In 2015, the findings of an independent review of the role of metrics in research assessment and management was published as a report 'The Metric Tide'[37]. The main findings were summarised in a press release from HEFCE[38]:

- ‘No set of numbers is likely to be able to capture the nuanced judgments that the REF process currently provides, and that it is not currently feasible to assess research outputs or impacts in the REF using quantitative indicators alone.
- Carefully selected indicators can complement decision-making, but a ‘variable geometry’ of expert judgement, quantitative indicators and qualitative measures

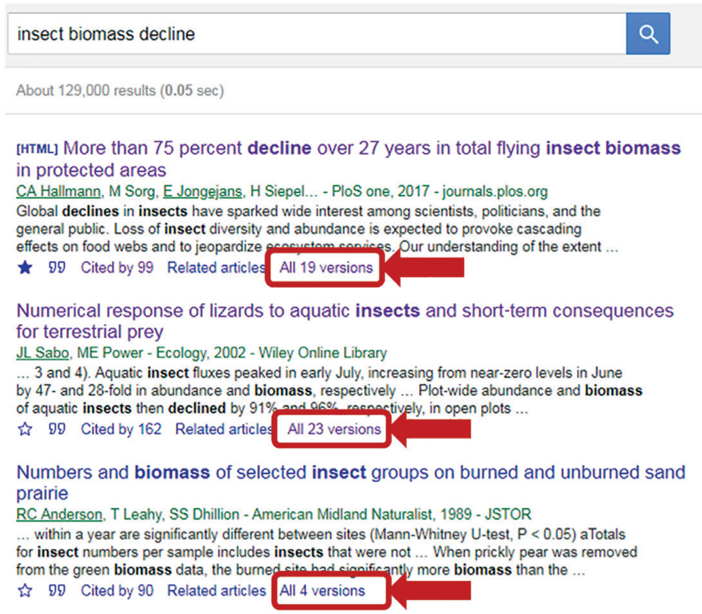


Figure 10 *Screenshot of Google Scholar search results on insect biomass decline and the number of versions available for each paper.*

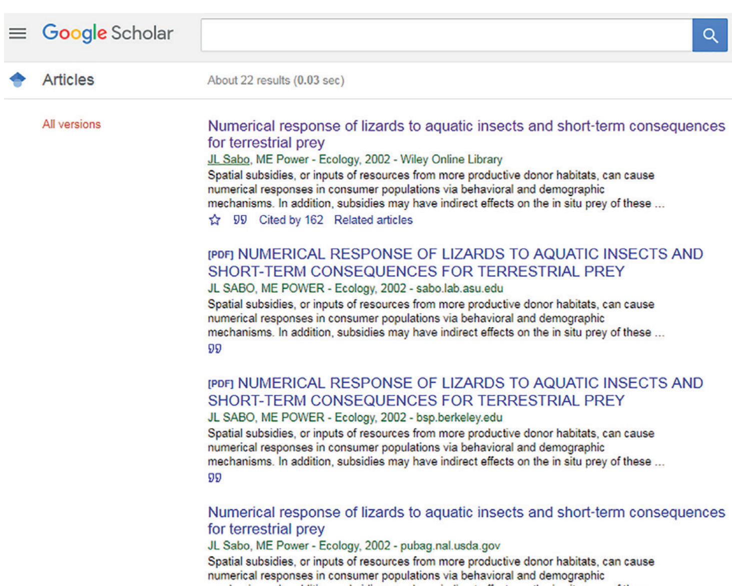


Figure 11 *Screenshot showing links to ‘all versions’ of an article in Google Scholar.*

that respect research diversity will be required. Greater clarity is needed about which indicators are most useful for specific disciplines, and why.

- Inappropriate indicators create perverse incentives. There is legitimate concern that some indicators can be misused or 'gamed': journal impact factors, university rankings and citation counts being three prominent examples.
- The data infrastructure that underpins the use of metrics and information about research remains fragmented, with insufficient interoperability between systems. Common data standards and transparent processes are needed to increase the robustness and trustworthiness of metrics.
- In assessing impact in the REF, as with outputs, it is not currently feasible to use quantitative indicators in place of narrative case studies, as it may narrow the definition of impact in response to the availability of certain indicators. However, there is scope to enhance the use of data in assessing research environments, provided data are sufficiently contextualised.'

Professor James Wilsdon, chairman of the review, commented:

'Metrics touch a raw nerve in the research community. It's right to be excited about the potential of new sources of data, which can give us a more detailed picture of the qualities and impacts of research than ever before. But there are also real concerns about harmful uses of metrics such as journal impact factors, *h*-indices and grant income targets.'

The 2016 Stern Review[39] confirmed many of the findings of 'The Metric Tide' and recommended that citation metrics could indeed continue to be used to support REF sub-panels in their assessment of outputs, but that 'it is not currently feasible to assess research outputs in the REF using quantitative indicators alone'.

7. Concluding remarks

Bibliometrics in all its forms can provide useful measures of the productivity of individuals and groups of researchers, of the activity within a subject area, and the influence of research outputs within both academic and general communities. It is essential though, that one understands how the metrics are calculated, at least in outline if not in detail, and to appreciate that metrics generated from within different collections are not comparable. Furthermore, bibliometrics are a measure of past research and cannot predict the success or otherwise of future activities. Most important of all, it must be remembered that bibliometrics are a quantitative measure of impact and do not necessarily reflect the quality of research. They should, therefore, always be used with caution and to complement qualitative assessment, and not regarded as a replacement or alternative to peer review.

Published online: 21 August 2018

8. References

1. Otlet, P. (1934) *Traité de documentatio: le livre sur le livre, théorie et pratique*. Mundaneum Editions, Brussels.
2. Otlet, P. (1934) Traité de documentation: le livre sur le livre, théorie et pratique. https://archive.org/details/OtletTraitDocumentationUgent [accessed 30 July 2018].

3. Rousseau, R. (2014) *Nature*, **512**, 218.
4. Pritchard, A. (1969) *J. Doc.*, **25**, 348–349.
5. Shapiro, F.R. (1992) *J. Assoc. Inf. Sci. Technol.*, **43**, 337–339.
6. Garfield, E. (1955) *Science*, **122**, 108–111.
7. Garfield, E. (2006) *Int. J. Epidemiol*., **35**, 1123–1127.
8. De Solla Price, D.J. (1965) *Science*, **149**, 510–515.
9. Google (n.d.) How we started and where we are today. Google. https://www.google.com/about/our-story/ [accessed 31 July 2018].
10. Brin, S. and Page, L. (1998) *Comput. Networks ISDN*, **30**, 107–117.
11. Strickland, J. (n.d.) Why is the Google algorithm so important? https://computer.howstuffworks.com/google-algorithm1.htm [accessed 11 July 2018].
12. Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2010) Altmetrics: a manifesto. http://altmetrics.org/manifesto/ [accessed 10 July 2018].
13. Elsevier (n.d.) About Scopus. https://www.elsevier.com/en-gb/solutions/scopus [accessed 18 July 2018].
14. Giles, J. (2005) *Nature*, **438**, 554–555.
15. Khabasa, M. and Giles, C.L. (2014) *PLOS One*, **9**, e93949. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0093949 [accessed 12 July 2018].
16. Martín-Martín, A., Orduna-Malea, E. and Delgado López-Cózar, E. (2018) *Scientometrics*, https://doi.org/10.1007/s11192-018-2820-9. https://link.springer.com/article/10.1007%2Fs11192-018-2820-9 [accessed 9 July 2018].
17. Google Scholar (n.d.) Classic papers 2006. https://scholar.google.com/citations?view_op=list_classic_articles&hl=en&by=2006 [accessed 30 July 2018].
18. University of Cambridge (n.d.) Predatory publishers. https://osc.cam.ac.uk/about-scholarly-communication/author-tools/considerations-when-choosing-journal/predatory-publishers [accessed 31 July 2018].
19. Lagzian, M. and Borchardt, G. (2017) *J. Cell Commun. Signal*., **11**, 181–185.
20. Somoza-Fernánde, M., Rodríguez-Gairín, J.M. and Urbano, C. (2016) *EPI*, **25**, 730–737.
21. Boucherie, S. and McCullough, R. (2018) CiteScore metrics updated with 2017 annual values. Elsevier. https://www.elsevier.com/connect/citescore-metrics-updated-with-2017-annual-values [accessed 11 July 2018].
22. Butler, D. (2008) Nature, **451**, 6.
23. Falagas, M.E., Kouranos, V.D., Arencibia-Jorge, R. and Karageorgopoulos, D.E. (2008) *FASEB J.*, **22**, 2623–2628.
24. Beatty, S. (2016) Journal metrics in Scopus: Source Normalized Impact per Paper (SNIP). https://blog.scopus.com/posts/journal-metrics-in-scopus-source-normalized-impact-per-paper-snip [accessed 3 July 2018].
25. Moed, H.F. (2010) *J. Informetr*., **4**, 265–277.
26. Hirsch, J.E. (2005) *Proc. Natl Acad Sci. USA*, **102**, 16569–16572.
27. Jones, T., Huggett, S. and Kamalski, J (2011) *World Neurosurg*., **76**, 36–38.
28. Google Scholar (n.d.) Google Scholar citations. https://scholar.google.com/intl/en/scholar/citations.html [accessed 8 July 2018].
29. Egghe, L. (2006) *Scientometric*s, **69**, 131–152.
30. Harzing, A-W.l. (2016) Metrics: *h* and *g*-index. https://harzing.com/resources/publish-or-perish/tutorial/metrics/h-and-g-index [accessed 30 July 2018].
31. Altmetric (n.d.) Discover the attention surrounding your research. https://www.altmetric.com/ [accessed 31 July 2018].
32. Semantic Scholar (2018) Frequently asked questions. https://www.semanticscholar.org/faq#advantages-of-semantic-scholar [accessed 1 August 2018].
33. da Silva, J.A.T. and Dobranszk, J (2018) *Scientometrics*, **115**, 1107–1113.
34. Barnes, C.S. (2016) *J. Doc*., **72**, 878–895.

35. Davis, P. (2018) Journals lose citations to preprint servers. https://scholarlykitchen.sspnet.org/2018/05/21/journals-lose-citations-preprint-servers-repositories/ [accessed 31 July 2018].
36. REF2014 Research Excellence Framework (n.d.) Citation data. http://www.ref.ac.uk/2014/about/guidance/citationdata/ [accessed 15 July 2018].
37. Wilsdon, J., Allen. L., Belfiore, E., *et al*. (2015) *The metric tide: report of the independent review of the role of metrics in research assessment and management.* Higher Education Funding Council for England, UK. https://responsiblemetrics.org/the-metric-tide/ [accessed 2 August 2018].
38. Higher Education Funding Council for England (2015) Metrics cannot replace peer review in the next REF. http://www.hefce.ac.uk/news/newsarchive/2015/Name,104464,en.html [accessed 6 July 2018].
39. Department for Business & Energy & Industrial Strategy (2016) *Research excellence framework (REF) review: building on success and learning from experience.* Department of Business, Energy and Industrial Strategy, London, UK. https://www.gov.uk/government/publications/research-excellence-framework-review [accessed 5 July 2018].