Library Management

"Sustainable management" of commercial electronic research resources and of its use in bibliometrics
Amanda Regolini, Emilie Gentilini, Marie-Pascale Baligand, Emmanuelle Jannès-Ober,

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

"Sustainable management" of commercial electronic research resources and of its use in bibliometrics

Amanda Regolini and Emilie Gentilini

DPV-IST, Irstea, Grenoble, France

Marie-Pascale Baligand

DPV-IST, Irstea, Lyon, France, and

Emmanuelle Jannès-Ober

DPV-IST, Irstea, Antony, France

Abstract

Purpose – This paper aims to analyze the pertinence of Bradford's law applied to a set of electronic periodical collections in a given discipline.

Design/methodology/approach – This study applied Bradford's law using citations rather than articles from a theme, thus ensuring the pertinence of the periodicals cited. Concentrating on citations in articles focuses on the stage following research rather than prior stages. The sources are actually used because they have been cited.

Findings – The study demonstrates the data centrality phenomenon based on the leading periodicals in a theme. There was a relationship between the electronic periodical collection and the number of citations in the relevant journals in a theme.

Originality/value – Although today's use of periodical titles is defined first and foremost by readership statistics, a more detailed analysis on the qualitative aspect can be an additional benefit in the negotiations conducted on a macro scale. Using Bradford's law could be envisaged as a complementary tool for a sustainable management of commercial electronic resources.

Keywords Resource management, Electronic periodicals, Bradford's law, Citations analysis, Decision aids, Electronic resources

Paper type Research paper

1. Introduction

Any scientific periodical collection manager needs to know whether subscriptions respond to the users' expectations without exceeding the budget allotted, or even reducing it. In 1934, Samuel C. Bradford in Great Britain set out a law, now eponymous, making it possible to identify periodicals within a given theme. The fundamentals of this law have been condensed in few words by Donald Urquhart: "For any subject there is a concentration of references in a few periodicals and other references are scattered more and more thinly in an increasing number of periodicals" (Shenton and Hay-Gibson, 2009)

Traditionally, Bradford-distributions have been a tool for selection of the most central information sources in a field. The law has been very useful in documentation to organize bibliographical work and scientific documentation. Bradford's law is

applicable to build paper periodical collections and it is still employed to this end nowadays, i.e. (Goossens, 2011).

At the scale of a library, the question can be raised of whether using the Bradford law for managing electronic periodical collections accurately estimates the correlation between the subscriptions and the periodicals actually within a given theme.

In addition, citation analysis can also be used to determine which journals treat similar subjects. This is the method used by Eugene Garfield to create the Science Citations Index (SCI) database in 1955, in which the major journals within a given theme are determined by their citation rate (Garfield, 1955).

The relationship between the demand for articles and the citations in relevant journals for a given theme has been discussed in terms of usefulness by The British Library Lending Division (Scales, 1976) and after that by the INIST (French Institute for Scientific and Technical Information (Salaun *et al.*, 2000). The first study carried out showed a low correlation between citations and frequency of use. In the second comparative analysis it was shown a high correlation between most frequently requested articles and most frequently cited papers according to SCI. Nevertheless, the two studies are not comparable as the context as well as the time scales and geographical standards had changed in 25 years.

At an information service level, in a research centre it is easy to exploit "local faculty publications data" (Nisonger, 2007) and compile a list of journal's titles where their papers are published. In the same context this data can be examine to explore the journal's titles of the references they have cited. The present study applies Bradford's law using references citations rather than articles from a theme, thus ensuring the pertinence of the periodicals cited, concentrating on citations in articles focuses on the stage following research rather than prior stages. The sources are actually used because they have been cited.

However, publishers make electronic periodicals available according to two major models: subscription on a title-by-title basis and subscription by bundle, otherwise known as "the big deal" (all of a publisher's titles) (Shubber, 2011). Bundling entails publishers offering libraries packages of titles for a single fee, which lacks in flexibility because their contents may be poorly adapted to the true needs of libraries in that they also include journals in themes that are not suited to its users' interests.

Thus, this study attempted to determine the scope of the use of one publisher's bundle compared to another and to determine whether or not article citations reflect the pre-eminence of a bundle. In other words, we sought to determine whether Bradford's law could be applied to bundled periodical collections.

2. Data collection method and analysis of the results

To illustrate our development we analyse the sources of citations in articles published within the forest science research theme by researchers at Irstea, a French research institute for engineering in agriculture and the environment.

The search was carried out at the end of May 2011. The sources cited by Irstea researchers working in forest science were studied in order to evaluate the accuracy of Irstea journal's collections in this subject. The sample period (2008-2011) was chosen in agreement with Irstea last research strategic plan.

In any bibliometric analysis, it is fundamental to use a source that is well adapted to the studied field. Even though the coverage of forest science journals is not exhaustive,

particularly for French-language journals, this methodological exercise opted to use the Web of Science (WoS) database of the Web of Knowledge, in an effort to avoid beginning the study with Elsevier's Scopus because Elsevier's publishing activity could bias the results.

Data cleansing and verification is always necessary. For example, journal title abbreviations are occasionally erroneous. Moreover, a journal's title change must be taken into account, such as the "Journal de Sciences Forestières", which became "Annals of Forest Science". Errors were identified using Le Sphinx® software.

A total of 115 articles were obtained from the following search: Irstea affiliation, forest science dating from 2008 to 2011. The articles came from 64 journals and four symposia. These journals were in the following WoS categories: forest science, environmental science, ecology, botanics, and agriculture.

In addition, an analysis of the articles' key words was performed. The indexation used by Thomson Reuters is automatic and not based on a thesaurus. It was decided to recover the indexation in the Irstea Publications database. The terms were manually cleansed, and the Redes 2005[1] software provided the following diagrams (see Figure 1).

The 115 forest science articles specifically covered the following subjects: forest management, biodiversity, plant growth, forest stands, and climate change. The graphs show the key word networks for two examples of concepts: biodiversity and climate change. The thicker the lines linking the key words, the stronger the link between the key words.

These are citations found in the 115 articles of the chosen theme for this analysis. The corpus comprises 6,295 citations, with a mean 54.73 bibliographic references cited per article. It can be observed that 81 per cent of the citations cite other journal titles, as shown in Figure 2.

It is important to note that more than 1,100 citations refer to other types of raw data documents. We found 14 citations to norms, 77 citations to doctoral dissertations, and approximately 1,000 citations to reports, guides, books, conferences, software, statistical data, databases, etc.

The citations' date of publication ranged from 1822 to 2011, with 12 citations between 1822 and 1920. The majority of the citations referred to documents published between 1991 and 2010. Nevertheless, as shown in Figure 3, the greatest number of
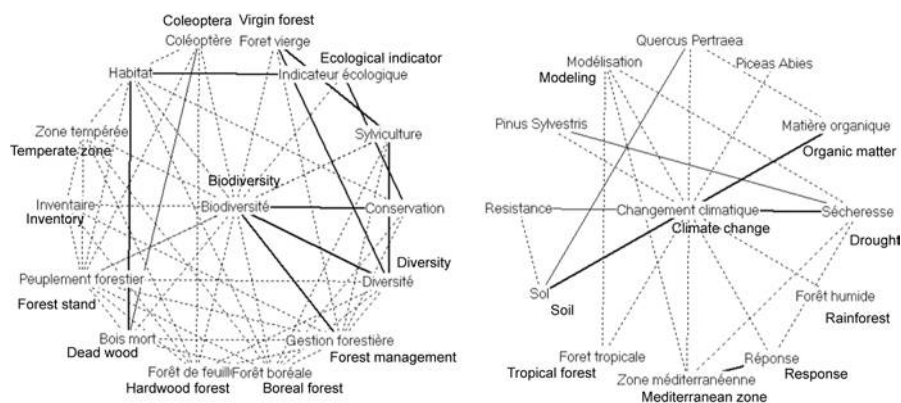
Figure 1.
Representations of the networks of the terms "biodiversity" and "climate change"
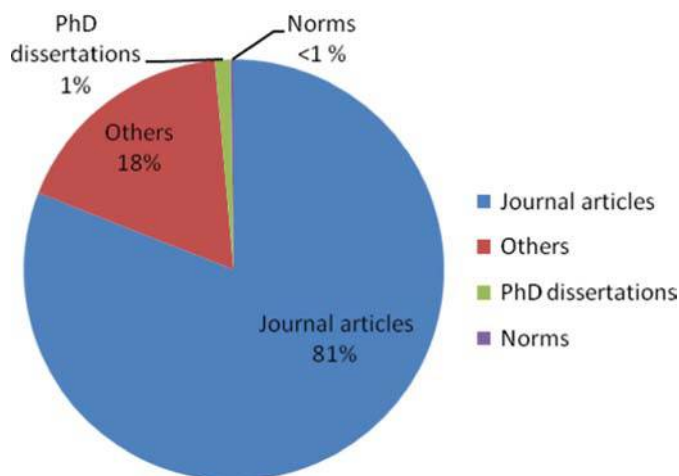
Figure 2.
Scattering of citations by
document type
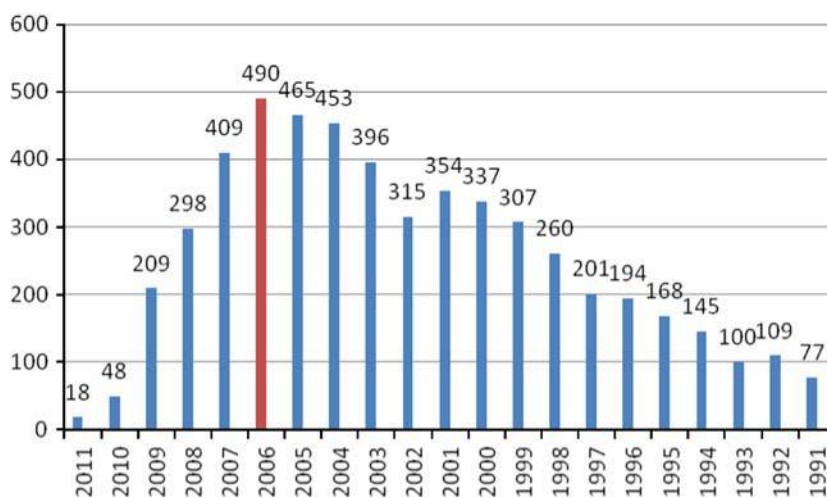
Figure 3.
Scattering of citations by
year between 1991 and
2011

references were dated between 2003 and 2007, with a peak in 2006, between two and
five years before the publication.

To analyse a representative sample, 332 was taken as the upper boundary, i.e. the
maximum number of citations per title, and the titles cited at least three times as the
lower boundary. (The rest of the journals were too peripheral in relation to the subject
to be useful to take them into account). This made up a corpus of 232 titles and 4,069
citations.

Bradford's law
Depending on the sample size, one must consider three, four, five, or more groups or
classes with the same citation content. Here, if our sample of 115 articles is divided into
four classes containing an equivalent number of citations, we observe that the first

class of six journals contains 999 citations, then the following 20 titles have 1,042 citations, the next 44 titles have 1014, and the last 162 journals have 1,014 citations (see Table I).

Knowing the number of journals per class and the number of citations in each class, the mean number of citations per journal title can be calculated. Then, based on these means (R), the empirical q coefficient can be calculated in order to evaluate the Bradford law as $G = G_1*(1 + q + q^2 + q^3)$. Where G is the chosen "corpus", and G_1 is the number of journals in class 1. With an easy "trial and error" we observe that q is close to 3, thus:

$G = G_1*(1 + 3 + 3^2 + 3^3) = G_1*40 = 240$, which is "close" to our corpus of 232 titles.

Figure 4 shows that the number of journals in each class increases like a geometric series from a central node whose factor is constant, as shown by Bradford's law.

If the percentage of citations is presented as a function of the number of journal titles, it can be seen that the total number of citations will reach 100 per cent following a logarithmic curve, as shown in Figure 5.

This curve demonstrates the data centrality phenomenon based on the leading periodicals in a theme. A small number of titles concentrate many citations.

The six major titles (class 1) are *Forest Ecology and Management* (Elsevier), the *Canadian Journal of Forest Research* (CNRC), *Ecology* (ESA, Ecological Society of

Classes K	(%)	Order of journals (Nos)	Number of journals G	Number of citations GKRK	Mean R
1	25	1-6	6	999	166.50
2	50	7-26	20	1,042	52.10
3	75	27-70	44	1,014	23.04
4	100	71-232	162	1,014	6.25

Table I.
Distribution of journals in each class
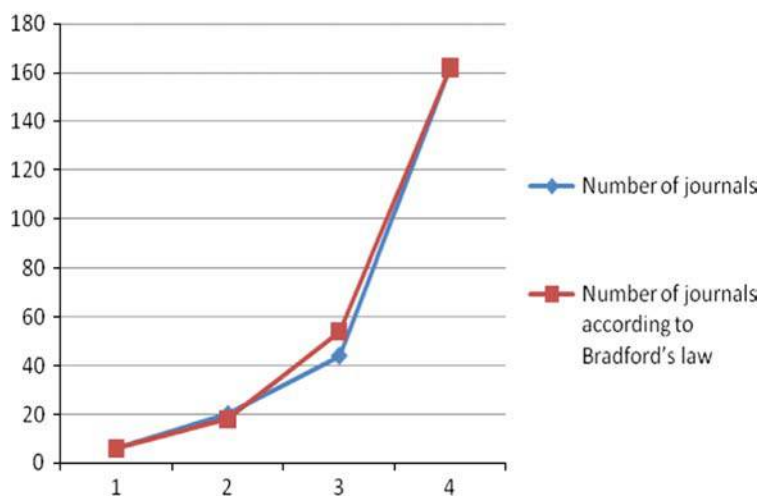
Figure 4.
Number of journals per class
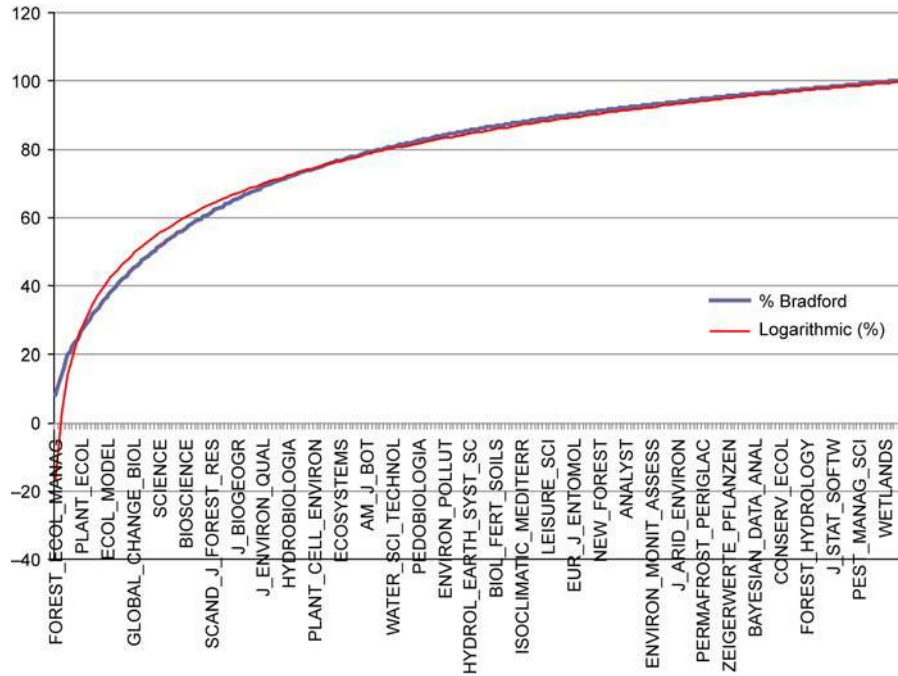
Figure 5.
Percentage of articles per
journal classified in
descending order

America), *Annals of Forest Science (Annales des Sciences Forestières, INRA)* (EDP),
Journal of Ecology (Wiley-Blackwell), and *Plant Ecology* (Springer).

The 20 titles within the second circle (class 2) are: *Journal of Vegetation Science,
Remote Sensing of Environment, Conservation Biology, Soil Biology and Biochemistry,
Biological Conservation, Ecological Applications, Water Resources Research, Oecologia,
Ecological Modelling, International Journal of Wildland Fire, Journal of Hydrology, Tree
Physiology, Soil Science Society of America Journal, Forestry, Global Change Biology,
Landscape Ecology, Journal of Applied Ecology, Ecological Monographs, Oikos,* and *New
Phytologist.*

All these core titles, slightly more than 80 per cent are accessible at the Irstea.

Of the total 233 titles, 189 can be accessed at the Irstea and nine are accessible from
group collections. A Irstea research group whose specialisation is forest science has
access to all of the titles.

Whilst 44 periodical titles are not within our subscriptions, three in the 20-title
second circle, researcher's demands are fulfilled via the French INIST service or via
interlibrary loan.

3. What position should be adopted in the subscription policy?

The relevance of Bradford's law as a tool for being acquainted with researchers' needs
should be considered in scientific and technical libraries or documentary services.

Moreover taken into account researcher's citations allows to concentrate in the
documents really reads. Therefore the assessment of library collection quality could be
more accurate than using only the ratio of article downloads by journal title.

After applying the Bradford's law, a quarter of the citations corresponded to six periodicals, totaling 999 citations. We obtained 50 per cent citations with 26 periodical titles. A total of 70 periodicals are needed to obtain 75 per cent of the articles. Finally, to reach 100 per cent of the sample's citations, 162 periodicals must be added. If the first 233 journals are subscribed to, all the publications on forest science are obtained. However, only 26 titles are needed to have half of the information, and only six titles are in fact the fundamental journals in the domain.

Irstea subscribes to 81 per cent of the 233 titles. This percentage confirms that the collections to which the institution subscribes are well targeted.

However, the variety of publishers already present in the first circle – Elsevier, Springer, Wiley-Blackwell, among others – makes it impossible to apply Bradford's law in the subscription policy for simple budgetary reasons. The acquisition of several journals on a single-title basis is proportionally more expensive than the bundled subscription.

This problem in our institution is transferable to the national level. If libraries abandoned bundled subscriptions, the risk wold then be to see the price of single-title subscriptions increase considerably to dissuade libraries from this practice (The Economist, May 2011).

Moreover, in France the current organisation is favourable to publishers because access to a collection for a single researcher can be financed by different structures (universities, research centres) (Baligand, 2010). The state must now reduce its deficit: it therefore can no longer pay these redundant subscriptions. Choices must be made in research institutions. Consequently, in-depth knowledge of researchers' needs is becoming an even more vital component that must be considered in the management of future subscriptions.

Although today's use of periodical titles is defined first and foremost by readership statistics, a more detailed analysis on the qualitative aspect can be an additional benefit in the negotiations conduced on a macro scale. Using Bradford's law could be envisaged for a sustainable management of commercial electronic resources.

4. Conclusions

Bradford's law shows the most relevant source titles in a given discipline, application of this law to citations allows one to identify the most widely read sources.

For commercial journals, if substantial restrictions or price increases mandate future budget reductions – and therefore a reduction in the number of subscriptions – one must prepare solutions to Irstea researchers that bypass the problem allowing libraries to offset the reduction in access to paying titles. One solution could be to develop the combined use of open archives and bibliographic tools in research units: for example, returning to the old principle of the post card by diverting the use of open archives and indicating the article's metadata and the author's e-mail address to request an off-print (Guédon, 2011).

Finally, we know that for commercial subscriptions, the solution aiming to limit these subscriptions to only the main titles is impossible today because of the publishers' bundled offer, but what will tomorrow hold? If negotiations are successful in making it possible to centralize subscriptions, they may well be based on a detailed needs analysis rooted in thorough knowledge of the most widely read titles. The

emergence of this practice could become a highly useful complementary tool in future negotiations with publishers.

In fact, with the number of article downloads a quantitative selection criterion, the qualitative aspect of Bradford's law applied to citations would usefully enhance the selection criteria. Bradford's law would thus find its place in purchasing policies, particularly if they are carried out on a national scale.

In conclusion, we recommend that this type of analysis based on in-depth knowledge of source periodicals and their use be used in the negotiations of periodical titles and bundles in the future.

Note

1. Available at: www.ugr.es/~rruizb/cognosfera/index.htm

References

Baligand, M.P. (2010), "Le passage au tout-numérique: une 'e-revolution' dans les centres de documentation", *Documentaliste – Sciences de l'information*, Vol. 47 No. 2, pp. 44-6.

The Economist (2011), "Academic publishing. Of goats and headaches", *The Economist*, Vol. 399 No. 8735, pp. 70-1.

Garfield, E. (1955), "Citation indexes for science; a new dimension in documentation through association of ideas", *Science*, Vol. 122, pp. 108-11.

Goossens, B. (2011), "Citation analysis as indicators of collection use and collection management at a research library", *Euraslic 14, Caught in the 'Fishing Net' of Information*, 14th Biennial EURASLIC Meeting, 17-20 May 2011, Lyon.

Guédon, J.C. (2011), "Maîtriser la grande conversation scientifique: le libre accès et la création de la valeur intellectuelle", *L'évaluation des productions scientifiques: des innovations en SHS?*, CNRS, 9-10 juin 2011, Paris.

Nisonger, T.E. (2007), "Journals in the core collection: definition, identification, and applications", *The Serials Librarian*, Vol. 51 Nos 3-4, pp. 51-73.

Salaun, J., Lafouge, T. and Boukacem, C. (2000), "Demand for scientific articles and citations: an example from the Institut de l'information scientifique et technique (France)", *Scientometrics*, Vol. 47 No. 3, pp. 561-88.

Scales, P. (1976), "Citation analyses as indicators of the use of serials: a comparison of ranked title lists produced by citation counting and from use data", *Journal of Documentation*, Vol. 32 No. 1, pp. 17-25.

Shenton, A.K. and Hay-Gibson, N.V. (2009), "Bradford's law and its relevance to researchers", *Education for Information*, Vol. 27, pp. 217-30.

Shubber, K. (2011), "Cash-strapped libraries threaten to end journal subscription deals over rising costs", available at: http://felixonline.co.uk/?article=808 (accessed 10 December 2011).

About the authors

Amanda Regolini is a Bibliometric Specialist Manager. She is in charge of bibliometrics indicators for Irstea, the French National Research Institute of Science and Technology for Environment and Agriculture. Amanda Regolini is the corresponding author and can be contacted at: amanda.regolini@irstea.fr

Emilie Gentilini is an Information Professional and Documentation Center Manager of Irstea Grenoble center. She's also a member of the Irstea's bibliometrics team which produces institutional indicators.

Marie-Pascale Baligand is in charge of electronic resources for Irstea. She works with the French consortium Couperin to negotiate databases and journals.

Emmanuelle Jannès-Ober is Deputy Director of the Direction of Foresight & Scientific Survey at Irstea. She is especially in charge of the service of scientific and technical information.

To purchase reprints of this article please e-mail: **reprints@emeraldinsight.com**
Or visit our web site for further details: **www.emeraldinsight.com/reprints**

This article has been cited by:

1. Mu-Hsuan Huang, Wei-Tzu Huang, Cheng-Ching Chang, Dar-Zen Chen, Chang-Pin Lin. 2014. The greater scattering phenomenon beyond Bradford's law in patent citation. *Journal of the Association for Information Science and Technology* **65**:9, 1917-1928. [CrossRef]