Discipline Impact Factors: A Method for Determining Core Journal Lists

Graeme Hirst*
Information Science Group, Department of Engineering Physics, Australian National University, Canberra, ACT 2600, Australia

A method of determining core journals for a discipline, using data from the *Journal Citation Reports* to generate *discipline impact factors*, is described.

One problem continually faced by libraries is which serials in a given discipline they should buy to meet the needs of their users, without subscribing to any unnecessary journals. The construction of *core lists* of journals for specific disciplines by various bibliometric methods [1] is one way of overcoming this problem. A new method of core-list construction using the *Journal Citation Reports* (*JCR*) [2, 3] has been developed. It is believed to be more accurate and less time-consuming than previous techniques, and can be performed manually by librarians themselves, whereas some previous studies, such as [4] and [5], have required large computerized data bases. It is also applicable to much smaller fields than these methods. For example, the study on the journals most cited by agricultural science [4] started off with an "arbitrary" [4, p. 272] core list of 347 journals! Errors introduced by the arbitrariness were minimized by the size of the field and data involved. The method to be described below can be used with a field whose total literature—core and fringe—is perhaps thirty journals or less. It involves the computation of an *impact factor* for the journals in a given discipline.

The *discipline impact factor* (DIF) is similar to the impact factor used in the *JCR* [1, 6, 7], which measures the average number of times a paper in a given journal is cited, except that the DIF measures the number of times a paper in a journal is cited in the core literature of the given discipline. This definition is, of course, circular: A knowledge of the core journals is required to determine the core journals. The procedure to be described overcomes this problem by iteration.

A good approximation to a core list may be determined as follows: A starting journal or set of journals relevant to the field concerned is first selected; in most disciplines, there are at least one or two journals whose importance to the field is obvious, if only from their titles, and these will suffice as a starting point. Call these journals the citing set C. For each journal J cited by a journal in C, compute its DIF according to the formula

$$\text{DIF} = nc/ns,$$

where nc is the number of citations of J by journals in C over a time period tc, and ns is the number of citable items published by J over a time period ts. These data are readily available in the *JCR*. The choice of the time periods tc and ts is fairly arbitrary, and may be made to suit the recent data available. For example, in a study done in 1977 [8], using the 1975 and 1976 *JCR*, tc was 1974-1975 and ts was 1972-1975. This completes the first iteration.

Subsequent iterations are then made as follows: The results of the previous iteration are examined, and journals which have high DIFs are added to C if they are not already there. Members of C with very low DIFs may be deleted from the set. This requires careful subjective judgment. The DIFs of journals cited by members of the updated C are then calculated. The procedure is terminated when C has stabilized. Usually, at most three iterations are required. Then those journals whose DIF exceeds a certain threshold are deemed to be core journals for the discipline. The threshold chosen is dependent on the citation characteristics of the discipline, and would be adjusted for an individual library according to its resources and the needs of its users. Any general scientific journals or journals from related fields that are frequently used by the discipline will also appear in the results, and should be included in the core list along with journals native to the field.

A detailed example of the use of the DIF to construct a core list for computer science is given in [8].

Because the data required are so easy to obtain, it is possible to consider much more data in a shorter time than one can with methods requiring tedious and error-prone examination of the journals themselves, as some previous methods [1, 9] have required, and the iterations serve to reduce the possibility of error in the choice of the original set C. The core list for computer science [8] is therefore more accurate than previous noniterated lists [9], and required much less time to produce.

The DIF method has a mild bias towards longer established journals, as they have more citable items before ts

*Current address: Department of Computer Science, University of British Columbia, Vancouver, BC, Canada V6T 1W5.

0002-8231/78/0029-0171$01.00

to be cited. But if only citations made to articles published during *ts* are considered, then the bias swings to those journals with a high *immediacy index* [2, 3, 6]. This bias may be preferred; but in either case, one should beware of journals which were once central to the discipline, yet still have high DIFs due to continued citation of old articles. For example, in our computer science study [8], we found that citations to *Proceedings of the Cambridge Philosophical Society*, highly placed on the previous list [9], were never made to articles published after 1958.

The method sometimes loses fine detail because the *JCR*s do not report journals cited by another less than six times in one year (with some minor exceptions). This does not affect major trends in the results. More annoying is the fact that not all journals included in the *Science Citation Index* [10] are analyzed in the *JCR*. Hopefully, this will change as the *JCR* develops.

In an independent study, Cawkell [11] uses the *JCR* to obtain data to present a connectivity map of journals in a given discipline, in this case acoustics. Cawkell's method and the DIF method complement one another nicely, and may be conveniently used together: The mapping method of Cawkell is a useful way of graphically displaying the results of a DIF analysis. Conversely, the calculation of DIFs can eliminate two subjective judgments required by Cawkell. The first of these is deciding eligibility of less central journals for the map, which Cawkell does arbitrarily, taking into account journal size and citedness; this decision can be made more objective by choosing just those journals whose DIF exceeds the set threshold. Secondly, when drawing the connectivity map, it is possible to make it somewhat deceptive (inadvertantly, of course) by not placing the most important journals in the center. If the map is drawn so that journals whose DIFs are highest are in the center, with decreasing DIF contours radiating out, then this possibility is eliminated.

The results of the DIF method were described above as "a good approximation to a core list." This qualification is made because there are inevitably core journals which the method will not pick up: those read for current awareness but not often cited, for example, or those which are new and for which citation data are not yet published. Librarians must rely on their own experience, and on requests made by library patrons, to identify these journals. (Both Cawkell and the present writer [8, 11] surveyed workers in the disciplines whose literature they studied, in order to aid evaluation of their results.)

Despite minor limitations, the speed and the ease of the DIF method make it a useful tool in core-list construction, whether the field of interest be large or small and the calculations manual or computerized.

References

1. Subramanyam, K. "Criteria for journal selection." *Special Libraries.* 66: 367-371; 1975.
2. Garfield, E. "Introducing *Journal Citation Reports.*" *Essays of an information scientist.* Vol. 2. Philadelphia: ISI Press; 556-571. [Reprinted in: *Current Contents,* No. 35, 1976 August 30.]
3. Garfield, E. (compiler and editor). *Science Citation Index: Journal Citation Reports: A bibliometric analysis of references processed for the Science Citation Index.* Philadelphia: Institute for Scientific Information; Published annually.
4. Garfield, E. "Journal citation studies. 20. Agriculture journals and the agricultural literature." *Essays of an information scientist.* Vol. 2. Philadelphia: ISI Press; 1977: 272-278. [Condensed from: Garfield, E.; Weinstock, M. "What is the literature cited by agricultural scientists? A case study." Paper presented at the World Congress of the International Association of Agricultural Librarians and Documentalists, Mexico City, 14-18 April 1975. Reprinted in: *Current Contents,* No. 20, 1975 May 19.]
5. Garfield, E. "Journal citation studies. 21. Engineering journals." *Essays of an information scientist.* Vol. 2. Philadelphia: ISI Press; 1977: 304-309. [Reprinted in: *Current Contents,* No. 27, 1975 July 7.]
6. Garfield, E. "Citation analysis as a tool in journal evaluation." *Science.* 178: 471-479; 1972. [Reprinted in [3] and in: *Essays of an information scientist,* Vol. 1. Philadelphia: ISI Press; 1977: 527-571.]
7. Garfield, E. "What scientific journals can tell us about scientific journals." *IEEE Transactions on Professional Communication.* PC-16 (4): 200-203; 1973.
8. Hirst, G.; Talent, N. "Computer science journals–An iterated citation analysis." *IEEE Transactions on Professional Communication.* PC-20(4): 233-238; 1977.
9. Subramanyam, K. "Core journals in computer science." *IEEE Transactions on Professional Communication.* PC-19(2): 22-25; 1976.
10. *Science Citation Index.* Philadelphia: Institute for Scientific Information; Cumulated annually.
11. Cawkell, A.E. "Evaluating scientific journals with *Journal Citation Reports*–A case study in acoustics." *Journal of the American Society for Information Science.* 29: 41-46; 1978.