Fractionally Counted versus Integer-Counted Impact Factors in the ISI Journal Set 2008

Loet Leydesdorff [1] and Lutz Bornmann [2]

[1] *loet@leydesdorff.net*
Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, NL-1012 CX Amsterdam (The Netherlands)

[2] *bornmann@gv.mpg.de*
Max Planck Society, Hofgartenstrasse 8, D-80539 Munich (Germany)

Abstract

The ISI-Impact Factors suffer from a number of drawbacks, among them the incomparability among fields of science because of systematic differences in citation behavior among fields. Can this problem be counteracted by counting citation weights fractionally instead of using whole numbers in the numerators? Fractional citation counts are normalized in terms of the citing sources and thus would take into account differences in citation behavior among fields of science. Differences in the resulting distributions can be tested statistically for their significance at different levels of aggregation. A list of fractionally counted Impact Factors for 2008 is available online at http://www.leydesdorff.net/weighted_if/weighted_if.xls. Using these weighted impact factors, the in-between group variance among the thirteen fields of science identified in the U.S. *Science and Engineering Indicators* is no longer statistically significant. Although citation behavior differs largely between disciplines, the reflection of these differences in citation distributions cannot—*vice versa*—be used as a reliable instrument for the classification (Leydesdorff & Bornmann, 2011).

Introduction

The well-known impact factor (IF) of the Institute of Scientific Information (ISI)—presently owned by Thomson Reuters—is defined as the average number of references to each journal in a current year to "citable items" published in that journal during the two preceding years. Ever since its invention in 1965 (Sher & Garfield, 1965; Garfield, 1972 and 1979a), this ISI-IF has been criticized for a number of seemingly arbitrary decisions involved in its construction. The definition of "citable items"—articles, proceedings papers, reviews, and letters—the choice of the mean (despite the well-known skew in citation distributions; Seglen, 1992), the focus on two preceding years as representation of impact at the research front (Bensman, 2007), etc., have all been discussed in the literature, and many possible modifications and improvements have been suggested (recently, e.g., Althouse *et al.*, 2009).

In this study, fractional counting of citations is used as a means to normalize for differences among fields of science: using fractional counting, a citation in a *citing* paper containing n references counts for only $(1/n)^{th}$ of overall citations instead of a full point (as is the case with integer counting). The ISI-IF is based on integer counting and thus sensitive to differences in citation behavior among fields of science. A fractionally counted IF would correct for these differences in terms of the sources of the citations. Such normalization therefore can also be called "source-normalization" (e.g., Moed, 2010; Van Raan *et al.*, 2010; Waltman & Van Eck, 2010; Zitt, 2010).

The application of the tool of fractional counting of citations to journal evaluation was anticipated by Zitt & Small (2008) and Moed (2010). Zitt & Small (2008) proposed the Audience Factor (AF) as a new indicator, but used the *mean* of the fractionally counted citations to a journal at the journal level (Zitt, 2010). Moed (2010) divided a modified IF (with a window of three years and a somewhat different definition of citable issues) by the *median* of the citation potentials in the Scopus database. He proposed the resulting ratio as the Source Normalized Impact per Paper (SNIP) which is now in use as an alternative to the IF in the Scopus database (Leydesdorff & Opthof, 2010a). However, a quotient between two statistics no longer contains error terms, whereas the IF itself can be considered as a mean (over the last two years) with underlying distributions which can be compared using statistical tests (Bornmann, 2010; Opthof & Leydesdorff, 2010; Plomp, 1992; Pudovkin & Garfield, in print; Stringer *et al*., 2010).

Using fractional counting of citations at the article level, the distributions of citations in the citing documents can be compared in terms of means, medians, variances, and other statistics. Differences among document sets can be tested for their significance independently of whether one uses journals, research groups, or other aggregating variables for the initial delineation of document sets. Although this can be done equally for fractional and integer counting, our hypothesis is that the difference between these two counting methods for citations is caused by the variation in citation behavior among fields.

Methods and materials

Data was harvested from the CD-Rom versions of the *Science Citation Index* (SCI) 2008 and the *Journal Citation Reports* (JCR) 2008. Note that the CD-Rom version of the *SCI* covers fewer journals than the *Science Citation Index-Expanded* (*SCI-E*) that is available at the Web of Science (WoS; cf. Testa, 2010). (This core set is also used for the *Science and Engineering Indicators* of the National Science Board of the USA.[30]) The data on the CD-Rom for 2008 contains 1,030,594 documents published in 3,853 journals.[31] Of these documents, 944,533 (91.6%) contain 24,865,358 cited references. Each record in the ISI set contains conveniently also the total number of references (n) at the document level. Each citation can thus be weighted as $1/n$ in accordance with this number in the citing paper. We used these weights as citation counts in the numerator of a (quasi-)IF.

In a first step, the references to the same journal within a single citing document were aggregated. For example, if the same document cites two articles from *Nature*, the fractional citation count in this case is $2/n$. In this step, citations without a full publication year (e.g., "in press") were no longer included. This aggregation led to a file with 14,367,745 journal citations; 9,702,753 of these (67.5%) contain abbreviated journal names that we were able to match with the abbreviated journal names in the list of 6,598 journals included in the *SCI-E* in 2008.[32]

There was no *a priori* reason to limit our exercise to the smaller list of the CD-Rom version of the *SCI* because all journals can be cited and IFs for all (6,598) journals in the *SCI-E* are available for the comparison. However, only citations provided by the 3,853 journals in the smaller set (of the *SCI*) are counted in this study given the database that is used as source data on the citing side. As one can expect significantly lower numbers of references than those retrievable in the JCR, we name the IFs thus calculated quasi-IFs.

A match in terms of the journal abbreviations in the reference list was obtained in 6,566 (99.5%) of the 6,598 *JCR*-journals. These 6,566 journals contain 19,200,966 (77.2%) of the total of 24,865,358 original references. The citation numbers in this selection are used for computing the total cites for each journal, both fractionally and as integer numbers. When counted fractionally the number of references is 555,510.07 (that is, 2.89% of the total number of references or, in other words, with an average of 34.6 references per citing article).

Notes

[30] Ken Hamilton, communication at the email list sigmetrics@listserv.utk.edu, 3 May 2010.

[31] We found 3,853 journal titles in the download. Ken Hamilton (personal communication, June 1, 2008) reports 3,737 journals used for preparing the *Science and Egineering Indicators 2010* (NSB, 2010) based on the same files (2008).

[32] As an exception, the journal name 'Arthritis and Rheumatism' is abbreviated with 'Arth Rheum/Ar C Res' in the journal list, but with 'Arth Rheum' when used in cited references.

Table 1: Descriptive statistics of the citation data 2008 and the various steps in processing

SCI 2008	*Citations to all years*	*Citations to 2006 and 2007*
Nr of cited references	24,865,358	3,898,851
Nr of abbreviated journal titles	14,367,745	2,936,157
Nr of abbreviated journal titles matching	9,702,753	2,422,430
Nr of cited references after matching	19,200,966	3,320,894
Nr of cited references fractionally counted	555,510.07	596,755.99 (103,828.70)
Average nr of references/paper	34.6	5.6

By setting a filter to the citations from 2006 and 2007 in the original download, the numerators of the weighted quasi-IFs can be calculated from the same 25M references; the same procedure was repeated for this subset. The third column of Table 1 shows the corresponding numbers.

For the denominator of our quasi-IFs, we used the sum of the numbers of citable issues in 2006 and 2007 as provided by the *JCR*s of these respective years. By setting a filter to the period 2003-2007, one could analogously generate a five-year IF, both weighted or without weighting. However, we limit the discussion here to the two-year IF and follow strictly the definitions of the ISI (Garfield, 1972). Of the 6,598 journals listed in the JCR-2008 only 5,794 could thus be provided with a value for the denominator of the IF in 2008 based on values for the number of citable items in the two preceding years larger than zero. In a next step, we use exclusively the references provided to the 2006 and 2007 volumes of the 5,742 journals which have both a non-zero value in the numerator (2008) and in both terms of the denominator (2006 and 2007, respectively). These 5,742 journals contain a subset of 3,255,133 (98.0% of 3,320,894)) references or fractionally counted 583,833.98 (97.8% of 596,755.99) references, to publications in 2006 and 2007.

Testing differences in citation behavior among journals for significance

The fractional counts of the citations provide us with distributions indicating citation behavior at the level of each journal. Which statistics could be useful to test these multiple citation distributions of different sizes for the significance of their homogeneity and/or differences? Post hoc pairwise comparisons can be performed after obtaining a significant omnibus F with ANOVA. Among the post hoc tests which are available in SPSS for multiple comparisons, one may prefer to choose one of the tests which do not *ex ante* assume equal variance (for example, Dunnett's C test). However, this assumption about homogeneity in the variance itself can first be tested using Levene's Test for Equality of Variances (available within ANOVA). If alternatively the assumption holds, one can use the Tukey test which—as implemented in SPSS—includes controls for testing the significance of the differences among *multiple* samples.

Testing for between-group variances among fields of science

We will test the extent to which the normalization implied by using fractional counting reduces the between-group variance in relation to the within-group variance for the case of the thirteen fields of science identified by ipIQ for the purpose of developing the *Science and Engineering Indicators 2010* (NSB, 2010, at p. 5-30 and Appendix Table 5-24). We chose this classification because it is reflexively shaped and regularly updated on a journal by journal basis without automatic processing. Furthermore, journals are uniquely attributed to a broad field. However, the attribution is made only for the 3,853 journals used as original source data in both this study and the *Science and Engineering Indicators* of the NSF.

A two-level regression model will be estimated in which the (quasi-)IFs of journals are level-1 units and the 13 fields are level-2 clusters. Various two-level regression models are possible—depending on the scale of the dependent variable (here: IFs). Since IFs for journals are based on citation counts for the papers published in these journals, citations can be considered as count data. In the case of count data, a Poisson distribution is the best assumption (Cameron & Trivedi, 1998). Thus, we shall calculate a two-level random-intercept Poisson model. In order to handle overdispersion at level 1 (measured as large differences between the mean and the variance of the IFs) in this model, we follow Rabe-Hesketh & Skrondal's (2008) recommendation to use the sandwich estimator for the standard errors.[33]

Results

Let us as first compare the ISI-IFs as provided by the JCR 2008 with the quasi-IFs retrieved from the CD-Rom version of the *SCI 2008*. Table 2 provides the Pearson and Spearman rank correlations between the ISI-IF, the quasi-IF derived from the download of 2008, and the corresponding quasi-IF based on fractional counting. Not surprisingly—because of the high value of *N*—all correlations are significant at the 0.01 level. In the rightmost column, we also added the fractionated citations/publications ratio for 2008, for reasons to be explained below.

Table 2: Correlations between the ISI-IF, quasi-IFs based on integer and fractional counting, and fractionally counted citations divided by publications in 2008.[34] The lower triangle provides the Pearson correlations (*r*) and the upper triangle the Spearman rank-order correlations (*ρ*).[35]

	ISI-IF	*Quasi-IF (integer)*	*Quasi-IF (fractional)*	*Fractional c/p 2008*
ISI-IF		.898(**)	.835(**)	.669(**)
		5742	5742	5687
Quasi-IF (integer)	.971(**)		.937(**)	.770(**)
	5742		5742	5687
Quasi-IF (fractional)	.926(**)	.937(**)		.813(**)
	5742	5742		5687
Fractional c/p 2008	.746(**)	.771(**)	.818(**)	
	5687	5687	5687	

** Correlation is significant at the 0.01 level (2-tailed).

As could be expected, the quasi-IF based on integer counting correlates higher with the ISI-IF than the one based on fractional counting. These correlations confirm that our quasi-IFs can be considered similar to the ISI-IF in nature, although there may be important differences at lower levels of aggregation. The two IFs (based on integer and fractional counting, respectively) are very different in terms of the numerators of the IFs. Yet, the quasi-IF based

[33] We calculated also a normal regression analysis after lognormalizing the dependent variable. This procedure provides results that have the same tendency.

[34] Of these 5,742 journals, 55 journals did not contain a number of issues in the *JCR* 2008. (Of the 6,598 journals contained in the *JCR* 2008, 133 were not attributed a number of issues.)

[35] Using the Kolmogorov-Smirnov test, it could be inferred that the distributions for all four variables cannot be assumed to follow a normal distribution.

on fractional counting can explain more than 85% of the variance in the ISI-IF ($r^2 = (.926)^2 = .857$).

Is field normalization accomplished by fractional counting?

Ex-post estimation of pair-wise comparisons (ANOVA in SPSS, v. 15) allows for testing 50 cases at a time. How to select 50 from among the 5,742 journals in our domain? Most ISI Subject Categories contain more than 50 journals, but fortunately, the most problematic one of "multidisciplinary" journals contains only 42 journals. Preliminary testing of the fractional citation distributions of this set provided us with both counter-intuitive and intuitively expectable results. However, we saw no obvious way of validating the quality of the distinctions suggested by using Dunnett's C-test within this set.

Thus, we devised another test extending and generalizing from the differences studied in Leydesdorff & Opthof (2010) between three mathematics journals and two other journals. Can journals in mathematics and cellular biology (including *Molecular Cell*) be sorted separately using this method? For this purpose we used the 20 journals with highest ISI-IFs in the ISI Category *Mathematics*[36] and the 20 journals with highest ISI-IFs in the category of *Cell Biology*.[37]

In 2008, the top-20 mathematics journals range in terms of their ISI-IFs from 1.242 for *Communications in Partial Differential Equations* to 3.806 for *Communications on Pure and Applied Mathematics. Annals of Mathematics* and *Inventiones Mathematicae* are part of this set, but *Mathematical Research Letters* (with an ISI-IF of 0.524) is not.

[36] The ISI Category *Mathematics* contains 214 journal names with ISI-IFs ranging from zero to 3.806 for *Communications on Pure and Applied Mathematics.*

[37] The ISI Category *Cell Biology* contains 157 journal names with ISI-IFs ranging 0.262 for *Biologischeskie Membrany* to 35.423 for *Nature Reviews of Molecular Cell Biology.* (No ISI-IF 2008 is provided for *Animal Cells and Systems.*)

Table 3: IFs and quasi-IFs for the twenty journals with highest ISI-IFs in the ISI Subject Categories of *Mathematics* and *Cell Biology*.

Journal	*ISI-IF 2008*	*Quasi-IF (integer counting)*	*Quasi-IF (fractional counting)*	*fractionated c/p ratio 2008*	*Journal*	*ISI-IF 2008*	*Quasi-IF (integer counting)*	*Quasi-IF (fractional counting)*	*fractionated c/p ratio 2008*
Commun Pur Appl Math	3.806	2.151	0.750	2.390	*Nat Rev Mol Cell Bio*	35.423	28.339	3.129	4.416
B Am Math Soc	3.500	1.667	0.575	3.909	*Cell*	31.253	25.226	2.499	7.354
Ann Math	3.447	2.688	1.416	4.794	*Nat Med*	27.553	20.669	2.284	6.156
J Am Math Soc	2.476	1.667	0.803	1.429	*Annu Rev Cell Dev Bi*	22.731	18.385	1.967	5.168
Mem Am Math Soc	2.367	1.469	0.729	1.313	*Nat Cell Biol*	17.774	14.392	1.408	2.829
Invent Math	2.287	1.294	0.595	2.543	*Cell Stem Cell*	16.826	*n.a.*	*n.a.* [38]	0.447
Acta Math Djursholm	2.143	1.526	0.748	3.201	*Cell Metab*	16.107	12.994	1.347	0.890
Found Comput Math	2.061	1.121	0.422	0.207	*Gene Dev*	13.623	10.684	1.015	2.759
Comput Complex	1.562	0.357	0.175	0.144	*Trends Cell Biol*	13.385	11.212	1.186	2.088
Duke Math J	1.494	0.924	0.465	1.412	*Mol Cell*	12.903	11.011	1.143	2.151
Publ Math Paris	1.462	0.273	0.098	0.103	*Dev Cell*	12.882	10.566	1.095	1.516
J Differ Equations	1.349	0.992	0.382	0.659	*Curr Opin Cell Biol*	12.543	10.266	1.018	2.259
Am J Math	1.316	0.789	0.406	1.881	*Nat Struct Mol Biol*	10.987	9.695	1.000	0.887
Constr Approx	1.308	0.723	0.281	0.439	*Curr Opin Genet Dev*	9.677	7.156	0.727	1.368
Nonlinear Anal Theor	1.295	0.540	0.217	0.179	*Trends Mol Med*	9.621	6.961	0.742	1.215
B Symb Log	1.294	0.618	0.422	0.263	*Plant Cell*	9.296	8.213	0.890	2.030
Adv Math	1.280	0.797	0.409	0.487	*J Cell Biol*	9.12	7.743	0.827	2.977
Random Struct Algor	1.253	0.663	0.310	0.444	*Curr Opin Struc Biol*	9.06	7.337	0.883	1.674
J Differ Geom	1.244	0.791	0.369	1.684	*Embo J*	8.295	7.055	0.769	4.390
Commun Part Diff Eq	1.242	0.856	0.307	0.648	*Aging Cell*	7.791	5.345	0.501	0.316
Mean	1.909	1.095	0.494	1.407	Mean	15.343	12.276	1.286	2.645
Standard deviation	0.835	0.612	0.295	1.356	Standard deviation	7.949	6.449	0.694	1.928

[38] There is no number of issues listed for *Cell Stem Cell* in the *JCR* in 2006. This number is part of the denominator of an IF. However, the journal can be compared in terms of the citations provided in 2008 (that is, the numerator of the IF).

The top-20 journals in the ISI Subtject Category *Cell Biology* range in terms of their ISI-IF 2008 from 7.791 for the journal *Aging Cell* to 35.423 for *Nature Reviews of Molecular Cell Biology*. Thus, one can expect the two groups (*Mathematics* and *Cell Biology*) to be very different in terms of both their ISI-IFs—there is no overlap in the two ranges—and their citation practices. Table 3 provides the values for the ISI-IFs and our quasi-IFs—based on integer and fractional counting, respectively—for the two groups.

Table 3 shows that the mean of the quasi-IFs based on fractional counting remains more than twice as high for the 20 journals in molecular biology (1.286) than for the 20 journals in mathematics (0.494). Thus, the correction for the field level seems not complete. In a private email communication (23 June 2010), Ludo Waltman suggested that the remaining difference might be caused by the different rates at which papers in the last two years are cited in these two fields. In the journals classified as *Cell Biology* almost all papers contain references to recent (that is in this context, the last two years) publications, while this is less than half of the papers in journals classified as *Mathematics*.[39]

On the basis of this reasoning, a citation window longer than two years would attenuate this remaining difference. For example, the IF-5 can be expected to do better for this correction than the IF-2. More radically, the accumulation of all citations—that is, "total cites"—divided by the number of publications (the c/p ratio) for all years would correct for the differences among journals in terms of their cited half-lives.[40] The right-most columns in each category of Table 3, however, show that a difference between the mathematics set and the cell-biology set remains even when fractionated c/p ratios—which include citations from all years—are used. Thus, the field-specific effects are further mitigated, but do not disappear. In other words, these differences cannot be fully explained by the citation potentials of the two different fields; the fields remain different.

Let us take a closer look into these differences and to the issue of whether we should include all or more previous years or only the last two years? Can this distinction be retrieved by testing the fractionally counted numerators of the quasi-IFs of the 40 journals using a relevant *post-hoc* test? Among the (2 x 20 =) 40 journals 65,223 references were exchanged in 2008 to the volumes of 2006 and 2007. Between each two citation patterns of these 40 journals, one can test the differences for their statistical significance with ANOVA. Since the variances are again not homogeneous (Levene's test), we use the same Dunnett's C as the *post-hoc* test on the (40 * 39)/ 2 = 780 possible pairwise comparisons.

If two journals are not significantly different in terms of their fractionated citation patterns, they will be considered as belonging to the same group. Figure 1 shows the results for using these two groups of journals—with the black and white colors of the nodes indicating the *a priori* group assignment to mathematics or cellular biology—using Pajek and a spring embedded algorithm (Kamada & Kawai, 1989) for the visualization.

[39] Waltman & Van Eck (2010b) therefore suggests an additional normalization based on the *average* number of references in the citing journal rather than straightforwardly using the citing publications as the reference standard.

[40] The assumption implied is that the fields grow proportionally in terms of the database. Since this is not likely, a shorter citation window may also have advantages.

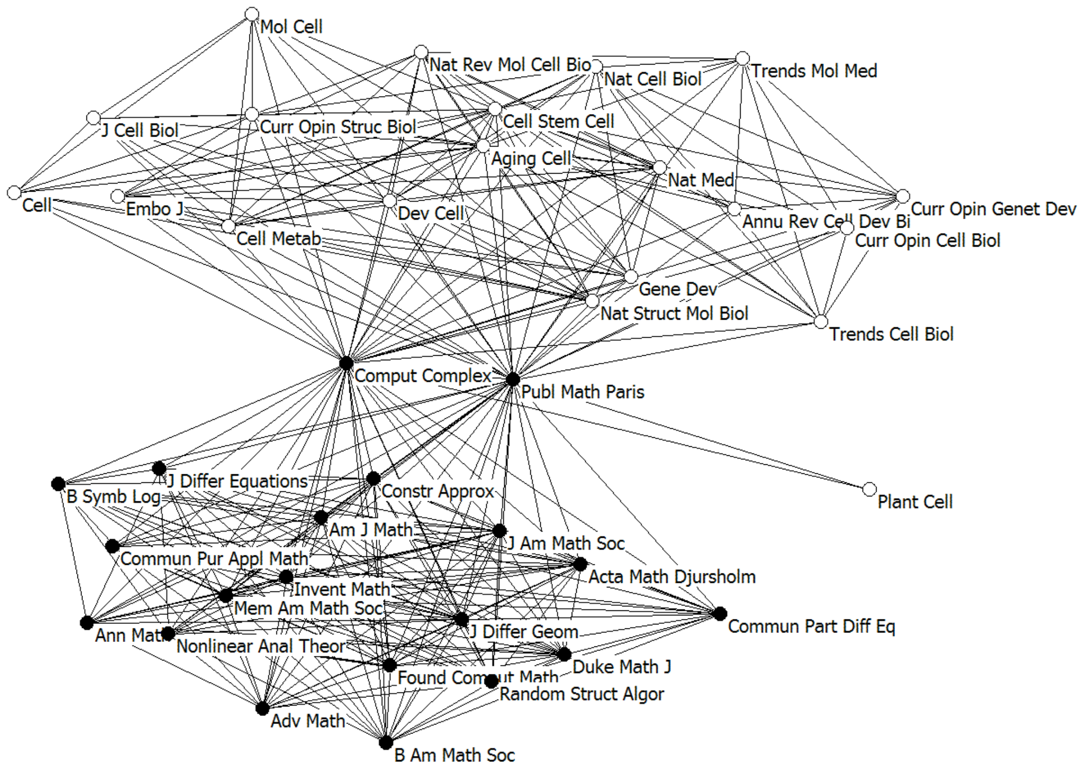

Figure 1. Dunnett's C test on fractionally counted citation impacts (2006 and 2007) for two groups of journals.

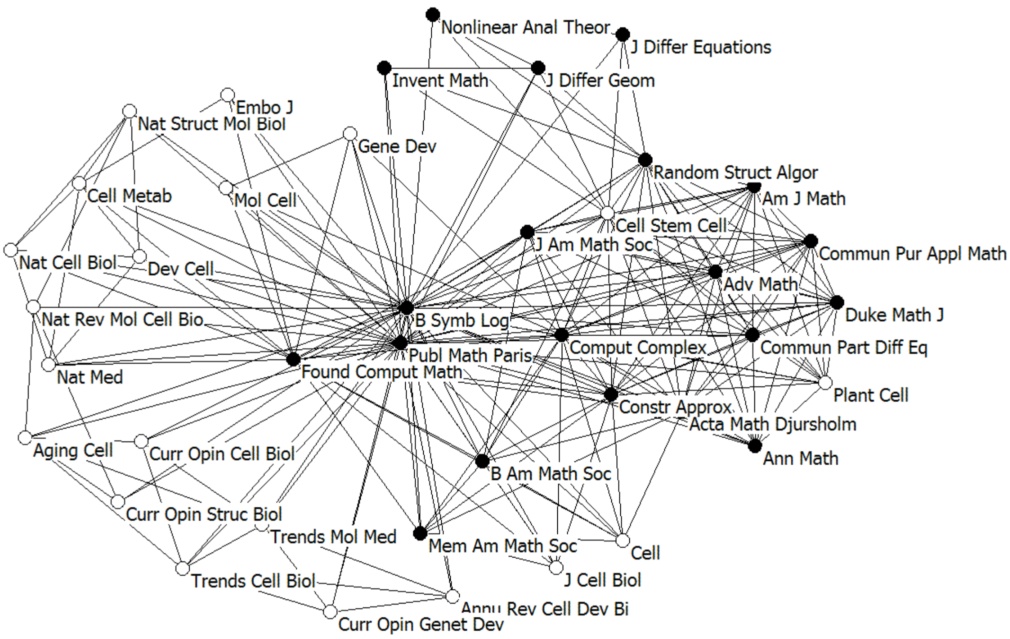

Figure 2: Mapping based on fractional counting of total cites in 2008; *N* = 40; Dunnett's C test; visualization in Pajek using Kamada & Kawai (1989).

Journals are linked in Figure 1 when these statistics are *not* significantly different—in other words, the journals can statistically be considered as a group—in terms of their fractional citation patterns (being cited in 2008). Although these results are motivating on visual inspection, they are not completely convincing. The journal *Plant Cell* is set apart—as it perhaps should be—but its relationships to the mathematics journals *Computational Complexity* and *Publications Mathématiques de l'IHÉS (Paris)* are unexpected. The patterns in these latter two journals deviate from their group (of mathematics journals) and accord also with other groupings.

We repeated the same exercise using not only the citations to the two previous years—that is, the numerators of the IFs—but the total cites to these 40 journals: 270,595 references are provided in 2008 to papers in these 40 journals. The larger size of this sample (415%) and the inclusion of citations to all previous years might make it easier to distinguish the two sets, but it did not! In Figure 2, some journals (e.g., *Cell Stem Cell*—a relatively new journal—but also *Cell*) are misplaced within the mathematics set.

In summary, the relations at the research front as indicated by the fractionated IF—that is, using only the last two years—are more distinctive than the total cites (that is, taking a longer time span into account). Similarly, a representation based on integer counting in the numerator of the IF (not shown) confirmed that this methodology can only be used for this purpose on the fractionally counted numerator of the quasi-IF.

Variance-component model

The research question is whether the differences among fields of sciences (that is, the between-field variance) can be reduced significantly by the normalization of the numerators of the IFs in terms of fractional citation counts. In order to answer this question, a variance-component model was calculated. The thirteen fields (NSB, 2010, at p. 5-30 and Appendix Table 5-24) provide the level-2 clusters, and the (quasi-) IFs of the journals are the level-1 units for this test. For reasons specified above, we defined additionally a model using the fractional c/p ratios as the dependent variable.

The results of the model estimations are presented in Table 4. We calculated four models (M1 to M4)—each using a different method of measuring journal impact: ISI-IFs 2008, quasi-IFs based on integer counting, quasi-IFs based on fractional counting, and fractionated c/p ratios for 2008. The models assume the intercept as a fixed effect and the variance of the intercepts across fields as a random effect. There are 3,923 (M1 to M3) or 3,869 (M4) IFs of journals, respectively, that are clustered within the 13 fields.[41]

Our assumption is that the level-2 (between-field) variance is reduced (or near zero) by using the IF based on fractional counting (M3) or the fractionated c/p ratio (M4), respectively, compared to the IF based on integer counting (M2). A reduction of this variance coefficient to close to zero would indicate that systematic field differences no longer play a role. The model for the ISI-IF (M1) is additionally included in Table 4; however, only the models M2 to M4 can be compared directly, because for these models the values for each journal are calculated on the basis of the same citation impact data.

[41] 54 journals contained in the CD-Rom version of the *SCI* are not provided with a number of issues in the *JCR* 2008.

Table 4. Results of four two-level random-intercept Poisson models

	M1: ISI-IF 2008	M2: IF (integer counting)	M3: IF (fractional counting)	M4: Fractionated c/p ratio 2008
Term	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Fixed effect				
Intercept	.67 (.11)*	.02 (.20)	-1.28 (.10)*	-.75 (.19)*
Random effect				
Level 2	.15 (.06)*	.48 (.21)*	**.09 (.05)**	.28 (.15)
$N_{journal}$	3923	3923	3923	3869
$N_{fields\ (clusters)}$	13	13	13	13

* $p < .05$

The results in Table 4 show that the variance component in the models M1 and M2 are statistically significant. In other words, both sets of data contain statistically significant differences between the fields. However, the variance component is not statistically significant in the models M3 and M4: field differences are no longer significant when the comparison is made in terms of fractionally counted citations. In the comparison of models M3 and M4 with model M2, the level 2-variance component is reduced by ((.48 – .09)/.48)*100) = 81% in model M3 and by ((.48 – .28)/.48)*100) = 42% in model M4.

In summary, the largest reduction of the in-between group variance is associated with model M3; in this case, the in-between group variance component is close to zero. This result provides a validation of our assumption: field differences in IFs are significantly reduced—to near zero—when the IFs are based on fractional counting. Using the longer time window as in the case of the c/p ratios does not improve on this result. In other words, these results point out that the quasi-IF based on fractional counting of the citations provides a solution for the construction of an IF where journals can be compared across broadly defined fields of science.

Conclusions and discussion

Fractionally counted impact factors can be used for comparison across field borders such as the thirteen broadly defined fields used by the NSF (NSB, 2010). Integer-counted IFs cannot be compared across fields of science because of differences in citation potentials among fields (Garfield, 1979b). However, the reasoning cannot be reversed: differences in citation potentials cannot be used for classifying fields or subfields.

First, other factors play a role such as the differences among document types (e.g., reviews versus research articles and conference proceedings) which are also unevenly distributed among fields of science. Relevant citation windows can also be expected to vary both among fields and over time. In addition to citation behavior, publication behavior varies among fields of science. In other words, the intellectual organization can be expected to affect the textual organization in ways that are different from the statistical expectations based on regularities in the observable distributions (Leydesdorff & Bensman, 2006; Milojević, 2010).

The remaining source of variance perhaps could be found in different portfolios among disciplines in terms of document types (reviews, proceedings papers, articles, and letters).[42] Moed (2010) proposed omitting letters when developing the SNIP indicator arguing that letters and brief communications inflate the representation of the research front by using more references to the last few years. Similarly, one could argue against using reviews because they

[42] The ISI (Thomson Reuters) decided to divide the category of "articles" into "articles" and "proceedings papers" as of October 2008.

may deflate the citation potential based on the most recent years (Leydesdorff, 2008, at p. 280, Figure 3). Review articles, however, are currently defined by Thomson Reuters among others as articles that contain 100 or more references.[43] One could focus exclusively on articles and proceedings papers, but in this study we wished to compare the effects of fractionation directly with the ISI-IF which is based on integer counting of the citations of all "citable items".

References

Althouse, B. M., West, J. D., Bergstrom, C. T., & Bergstrom, T. (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and Technology, 60*(1), 27-34.

Bensman, S. J. (2007). Garfield and the impact factor. *Annual Review of Information Science and Technology, 41*(1), 93-155.

Bornmann, L. (2010). Towards an ideal method of measuring research performance: Some comments to the Opthof and Leydesdorff (2010) paper. *Journal of Informetrics, 4*(3), 441-443.

Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge, UK: Cambridge Univ Press.

Garfield, E. (1972). Citation Analysis as a Tool in Journal Evaluation. *Science 178*(Number 4060), 471-479.

Garfield, E. (1979a). *Citation Indexing: Its Theory and Application in Science, Technology, and Humanities*. New York: John Wiley.

Garfield, E. (1979b). Is citation analysis a legitimate evaluation tool? *Scientometrics, 1*(4), 359-375.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters, 31*(1), 7-15.

Leydesdorff, L. (2008). *Caveats* for the Use of Citation Indicators in Research and Journal Evaluation. *Journal of the American Society for Information Science and Technology, 59*(2), 278-287.

Leydesdorff, L., & Bensman, S. J. (2006). Classification and Powerlaws: The logarithmic transformation. *Journal of the American Society for Information Science and Technology, 57*(11), 1470-1486.

Leydesdorff, L., & Bornmann, L. (2011). How fractional counting affects the Impact Factor: Normalization in terms of differences in citation potentials among fields of science. *Journal of the American Society for Information Science and Technology, 62*(2), 217-229.

Leydesdorff, L., & Opthof, T. (2010). *Scopus*' Source Normalized Impact per Paper (SNIP) *versus* the Journal Impact Factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology, 61*(11), 2365-2396.

Milojević, S. (2010). Modes of Collaboration in Modern Science: Beyond Power Laws and Preferential Attachment. *Journal of the American Society for Information Science and Technology, 67*(7), 1410-1423.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*(3), 265-277.

National Science Board. (2010). *Science and Engineering Indicators*. Washington DC: National Science Foundation; available at http://www.nsf.gov/statistics/seind10/ (retrieved on July 19, 2010).

Opthof, T., & Leydesdorff, L. (2010). *Caveats* for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance. *Journal of Informetrics, 4*(3), 423-430.

Plomp, R. (1990). The significance of the number of highly cited papers as an indicator of scientific prolificacy. *Scientometrics, 19*(3), 185-197.

[43] "In the *JCR* system any article containing more than 100 references is coded as a review. Articles in 'review' sections of research or clinical journals are also coded as reviews, as are articles whose titles contain the word 'review' or 'overview.'" At http://thomsonreuters.com/products_services/science/free/essays/impact_factor/ (retrieved June 18, 2010). See for problems with these delineations among document types in the Scopus database Leydesdorff & Opthof (2010c).

Pudovkin, A. I., & Garfield, E. (in print). Percentile Rank and Author Superiority Indexes for Evaluating Individual Journal Articles and the Author's Overall Citation Performance. *CollNet Journal*; available at http://garfield.library.upenn.edu/papers/aipegdalianchina2009.pdf (retrieved on June 18, 2010).
Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. College Station, TX: Stata Press.
Seglen, P. O. (1992). The Skewness of Science. *Journal of the American Society for Information Science, 43*(9), 628-638.
Sher, I. H., & Garfield, E. (1965). New tools for improving and evaluating the effectiveness of research. *Second conference on Research Program Effectiveness,* Washinton, DC, July, 27-29.
Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. N., (2010) Statistical validation of a global model for the distribution of the ultimate number of citations accrued by papers published in a scientific journal. *Journal of the American Society for Information Science and Technology, 61*(7), 1377-1385.
Testa, J. (2010). Web of Science Coverage Expansion. Available at http://community.thomsonreuters.com/t5/Citation-Impact-Center/Web-of-Science-Coverage-Expansion/ba-p/10663 (retrieved on July 19, 2010).
Van Raan, A. F. J., van Leeuwen, T. N., Visser, M. S., van Eck, N. J., & Waltman, L. (2010). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics, 4*(3), 431-435.
Waltman, L., & Van Eck, N. J. (2010). *A general source normalized approach to bibliometric research performance assessment*. Paper presented at the 11th Conference on Science & Technology Indicators (STI), Leiden, September 8-11.
Zitt, M. (2010). Citing-side normalization of journal impact: A robust variant of the Audience Factor. *Journal of Informetrics, 4*(3), 392-406.
Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and Technology, 59*(11), 1856-1860.