Journal of Informetrics 8 (2014) 310–317

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Counting and comparing publication output with and without equalizing and inflationary bias

Nils T. Hagen*

Faculty of Biosciences and Aquaculture, University of Nordland, N-8049 Bodø, Norway

ARTICLE INFO

Article history:
Received 20 November 2013
Received in revised form 3 January 2014
Accepted 6 January 2014
Available online 6 February 2014

Keywords:
Publication performance ranking
Coauthor credit
Harmonic formula
Bibliometric bias
Equalizing bias
Inflationary bias

ABSTRACT

This paper examines the effects of inflationary and equalizing bias on publication output rankings. Any identifiable amount of bias in authorship accreditation was detrimental to accuracy when ranking a select group of leading Canadian aquaculture researchers. Bias arose when publication scores were calculated without taking into account information about multiple authorship and differential coauthor contributions. The ensuing biased equal credit scores, whether fractional or inflated, produced rankings that were fundamentally different from the ranking of harmonic estimates of actual credit calculated by using all relevant byline information in the source data. In conclusion, the results indicate that both fractional and inflated rankings are misleading, and suggest that accurate accreditation of coauthors is the key to reliable publication performance rankings.

1. Introduction

To count publications is the most basic task in evaluative bibliometrics and scientometrics (De Bellis, 2009). The outcome of any such task, usually a ranking or other comparative assessment, is determined by how the countable units of publication are selected and accredited. Hence, when two rankings based on the same set of publications produce fundamentally different results, at least one must be misleading as a consequence of inaccurate accreditation. Systematic inaccuracies in accreditation arise when authorship credit is not divided among coauthors, i.e. inflationary bias, or when credit is divided equally among coauthors who have not contributed equally, i.e. equalizing bias (Hagen, 2008). Here, the question of interest is: How serious is the effect of these biases?

It is known from several studies comparing inflated and fractional credit that inflationary bias alters publication performance rankings (Aksnes, Schneider, & Gunnarson, 2012; Gauffriau & Larsen, 2005; Gauffriau et al., 2008; Huang & Lin, 2011; Huang, Lin, & Chen, 2011; Piro, Aksnes, & Rørstad, 2013; Pravdic & Oluic-Vukovic, 1991). However, the possibility that both inflated and fractional rankings are misleading cannot be dismissed as long as neither ranking is corrected for equalizing bias.

Less is known about the effect of equalizing bias. One study found that rankings of *h*-index scores were gravely distorted by both biases (Hagen, 2008), and another study estimated that equalizing bias accounted for approximately 60% of the

* Tel.: +47 755 17 398.
E-mail address: Nils.Hagen@uin.no

1751-1577
http://dx.doi.org/10.1016/j.joi.2014.01.003

variation in a composite empirical dataset (Hagen, 2013). These results suggest that the distortional effect of equalizing bias may be comparable to the effect of inflationary bias. Furthermore, it is important to account for equalizing bias as it may affect a majority of contemporary publications (Waltman, 2012), and because it controverts the purpose of performance ranking by diverting credit from primary authors to secondary authors (Hagen, 2008).

In this study I compare the separate and combined effects of equalizing bias and inflationary bias on the ranking of leading Canadian aquaculture researchers, and show that both biases have a detrimental effect on accuracy. I conclude by discussing the need to move from proxies of authorship credit to more accurate estimates based on all available relevant information in order to construct reliable publication performance rankings.

2. Materials and methods

2.1. Empirical data

The empirical data were independently derived from a ranking of leading Canadian aquaculture researchers (Picard-Aitken & Coté, 2010, table XIX, p. 34). The number of publications for each researcher was matched using the same time frame (1996-2008) and publication types (journal articles, conference papers, notes, and reviews from peer-reviewed journals) to extract data for each researcher from the same commercial database vendor (Scopus). This task was nontrivial as the number of obvious aquaculture publications for each researcher rarely corresponded to their number in table XIX (Picard-Aitken & Coté, 2010), and publications of possible relevance for aquaculture had to be omitted or added in order to obtain a matching number. This approach worked for 35 of the 36 researchers in the original table. But for one researcher who was listed with 19 "aquaculture" publications, I found it impossible to make a meaningful selection of papers among the approximately 70 publications provided by the Scopus database. This researcher had conducted basic research using cell cultures derived from the tissue of aquaculture species, and it is easy to see how a keyword search in a library database could deliver an imprecise result.

The end result was a replicate dataset consisting of 699 authorship contributions from 35 researchers, to 531 research papers from 120 different journals (The replicate dataset is available on request). Only 10 publications were single-authored (1.9% of 531 papers), and I was unable to detect any unequivocal indication of equality among the coauthored contributions (98.6% of 689 contributions from 521 papers). However, 125 contributions came from coauthored papers where senior authorship was indicated by the presence of a corresponding last author.

2.2. Authorship quantification

Rankings were constructed by tallying inflated, fractional and harmonic credit scores for each of the 35 researchers in the replicate data set.

Inflated credit was calculated by using contribution count as a proxy for authorship credit (cf. Picard-Aitken & Coté, 2010), i.e. by assigning one full unit of authorship credit repeatedly to every coauthor:

$$\text{Inflated } i\text{th author credit} = 1 \tag{1}$$

Fractional credit was obtained by dividing one full unit of credit equally among all N coauthors of a multi-authored publication as follows:

$$\text{Fractional } i\text{th author credit} = \frac{1}{N} \tag{2}$$

Harmonic authorship credit for the ith author of a publication with N coauthors was calculated according to the following formula (Hagen, 2008, 2013; Hodge & Greenberg, 1981):

$$\text{Harmonic } i\text{th author credit} = \frac{1/i}{1 + (1/2) + \cdots + (1/N)} \tag{3}$$

I assumed that the presence of a corresponding last author indicated a senior author whose contribution was equivalent to the contribution of the first author (cf. Buehring, Buehring, & Gerard, 2007; Mattsson et al., 2011). In such cases, the first and the senior author share the credit for the 1st and 2nd position, and this reduces the credit of intermediate coauthors by one position as follows (cf. Hagen, 2008, Fig. 5):

$$\text{1st and senior } (N\text{th}) \text{ author credit} = \frac{1 + (1/2)}{2(1 + (1/2) + \cdots + (1/N))} \tag{4}$$

$$\text{Intermediate } (i = 2, \ldots, N-1) \text{ author credit} = \frac{1/(i+1)}{1 + (1/2) + \cdots + (1/N)} \tag{5}$$

The accuracy of the harmonic formula was unrivalled by other formulations from the bibliometric literature, when assessed against an empirical baseline (Hagen, 2010, 2013).

2.3. Rank correlation

Kendall's rank correlation coefficient τ (*tau*) (Kendall, 1938) was determined for each pair of harmonic (H), fractional (F), and inflated (I) credit scores, τ_{HF}, τ_{HI}, and τ_{FI}, using the statistical software package R (http://www.r-project.org/). From these coefficient values I calculated Kendall's coefficient of partial rank correlation (Kendall & Gibbons, 1990; Kendall, 1942), as follows:

$$\tau_{HI.F} = \frac{\tau_{HI} - \tau_{HF}\tau_{IF}}{\sqrt{(1-\tau_{HF}^2)}\sqrt{(1-\tau_{IF}^2)}} \quad (6)$$

$\tau_{HI.F}$ is a measure of the correlation between the harmonic (H) and inflated (I) rankings after eliminating the effect of their correlation with the fractional ranking (F). The test of significance for $\tau_{HI.F}$ was carried out using the critical value 0.3041 for $N = 35$ and $(1-\alpha) = 0.995$ from Table VI in Maghsoodloo and Pallos (1981).

3. Results

3.1. Effect of equalizing and inflationary bias on credit scores

The inflated credit scores, i.e. the number of papers each researcher had contributed to during the time period 1996-2008, ranged from 13 to 35 (Fig. 1). According to the harmonic credit scores, each researcher's actual contribution ranged from 2.29 to 10.41 paper equivalents or between 14.9% and 52.5% of the authorship credit (Table 1). The difference between the harmonic and inflated credit scores represents the sum of contributions made by each researcher's collaborating coauthors. This difference is quantified by a combination of equalizing bias, ranging from −1.16 to 3.1, and inflationary bias ranging from 6.5 to 24.4.

Inflationary bias accounts for the large and obvious difference between inflated and fractional credit scores, whereas equalizing bias accounts for the numerically smaller but seemingly equally disruptive difference between harmonic and

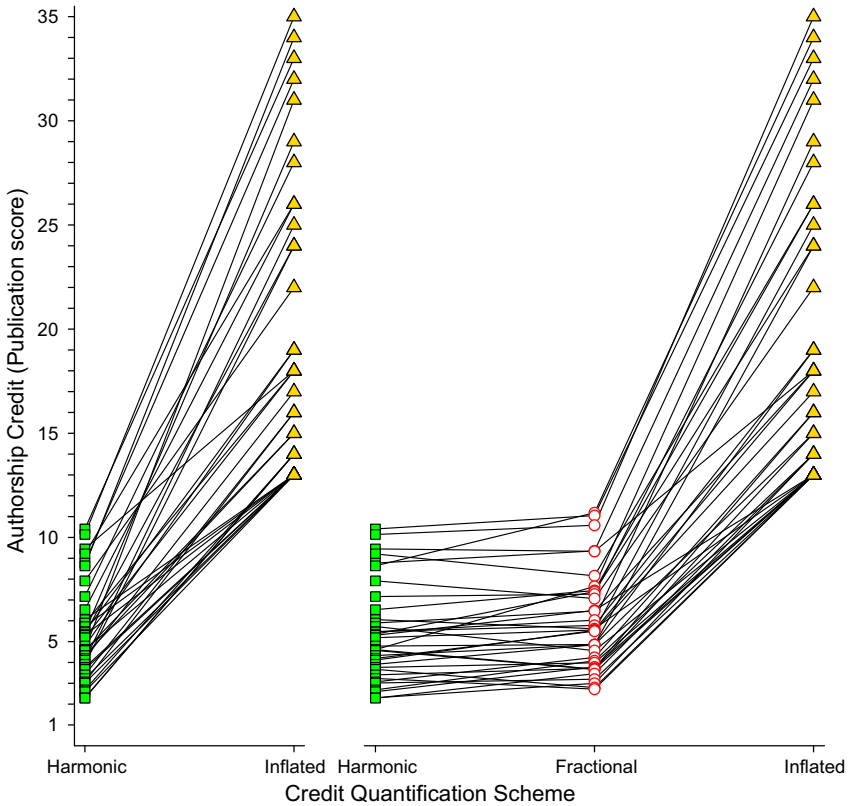

Fig. 1. Comparison of harmonic, inflated and fractional credit scores for a select group of 35 leading Canadian aquaculture researchers. Harmonic credit scores are estimates of actual authorship credit based on each coauthors position in the byline hierarchy, as well as additional information indicating the presence of a senior author. Inflated credit scores illustrate the effect of using contribution count as a proxy for authorship credit. Fractional credit scores show the effect of assuming that all coauthors have contributed equally. Slanted lines indicate bias-induced alterations of credit scores, and crossing lines indicate bias-induced rank displacement.

fractional credit scores (Fig. 1). Note that the range of fractional credit scores, 2.72–11.18, is comparable to the range of the harmonic credit scores, which indicates that the effect of equalizing bias is greatest for intermediate credit scores. Note also that crossing lines in Fig. 1 indicate credit scores which as a direct result of added bias have changed sufficiently to induce rank displacement. These changes are examined in the next section.

3.2. Bias-induced rank displacement

There was no coincidence between the harmonic and inflated rankings. The ranks of all researchers were displaced by the combination of equalizing bias and inflationary bias that was generated by using contribution count as a proxy for coauthor credit. Notably, the inflated ranking had a large number of tied ranks, all of which were resolved in the harmonic ranking (Fig. 2). The tied ranks represent researchers who had contributed differentially to an identical number of papers (Table 1), and the corresponding range of harmonic ranks reflects differences in the magnitude of their estimated contribution to those papers. In all, 72.2% of the inflated ranks were tied. The largest tied group consisted of 9 researchers who shared the bottom position of the inflated ranking. In the harmonic ranking they were redistributed over a range of 23 steps, from rank 10 to rank 33, on a scale with only 35 ranks (Fig. 2).

The fractional ranking resolved the ties of the inflated ranking, and 4 of its ranks (11.4%) coincide with the harmonic ranking. Nevertheless, 31 fractional ranks, representing 88.6% of the evaluated researchers, were displaced as a direct result of equalizing bias generated by dividing credit equally among coauthors who had not contributed equally. The displacement ranged from an advantage of 13 steps to a disadvantage of 21 steps, on a scale with only 35 ranks (Fig. 3).

Although the fractional ranking exhibited a significant rank correlation with both the harmonic and the inflated rankings (Table 2), the rank correlation between the harmonic and inflated rankings was weaker, and dwindled to almost zero (Kendall's partial rank correlation coefficient $\tau_{Harmonic\ Inflated.Fractional} = 0.0225$), when the effect of these rankings mutual correlation with the fractional ranking was eliminated. This result suggests that the added effect of inflationary bias is largely independent of the effect of equalizing bias.

Table 1
Summary data for 35 Canadian aquaculture researchers. Harmonic (H), fractional (F) and inflated (I) credit scores and ranks with associated estimates of bias and actual contribution percentage.

Credit score			Rank			Bias		Contribution (%)
H	*F*	*I*	*H*	*F*	*I*	Equalizing	Inflationary	
10.41	11.05	33	1	2	3	0.64	21.95	31.54
10.14	10.59	35	2	3	1	0.44	24.41	28.98
9.45	9.33	18	3	5	18	−0.12	8.67	52.50
9.21	8.16	26	4	6	8.5	−1.05	17.84	35.42
8.77	9.35	32	5	4	4	0.58	22.65	27.41
8.64	11.18	34	6	1	2	2.55	22.82	25.40
7.92	7.07	22	7	11	13	−0.85	14.93	35.98
7.16	7.28	28	8	10	7	0.12	20.72	25.58
6.53	7.45	26	9	8	8.5	0.92	18.55	25.13
6.06	5.62	13	10	16	32	−0.44	7.38	46.59
5.89	6.47	18	11	13	18	0.57	11.53	32.74
5.72	4.57	13	12	23	32	−1.16	8.43	44.04
5.48	5.78	19	13	15	15	0.30	13.22	28.83
5.47	6.03	29	14	14	6	0.56	22.97	18.85
5.36	7.40	24	15	9	11.5	2.04	16.60	22.32
5.30	6.50	13	16	12	32	1.20	6.50	40.79
5.19	5.56	18	17	17	18	0.37	12.44	28.82
4.79	4.86	19	18	20	15	0.08	14.14	25.19
4.62	3.68	13	19	29	32	−0.93	9.32	35.50
4.61	7.65	31	20	7	5	3.04	23.35	14.88
4.58	3.67	15	21	30	23.5	−0.91	11.33	30.50
4.33	4.85	15	22	21	23.5	0.52	10.15	28.87
4.19	5.50	17	23	19	20	1.31	11.50	24.65
4.11	5.54	25	24	18	10	1.43	19.46	16.44
3.91	4.85	24	25	22	11.5	0.94	19.15	16.29
3.76	3.98	13	26	26	32	0.22	9.02	28.93
3.67	2.79	13	27	34	32	−0.88	10.21	28.23
3.40	3.72	16	28	28	21.5	0.31	12.28	21.27
3.23	2.72	13	29	35	32	−0.51	10.28	24.87
3.06	4.24	16	30	24	21.5	1.18	11.76	19.15
3.01	3.19	13	31	32	32	0.18	9.81	23.18
2.68	4.08	14	32	25	26	1.39	9.92	19.16
2.60	3.78	13	33	27	32	1.19	9.22	19.97
2.30	3.00	14	34	33	26	0.69	11.00	16.44
2.29	3.46	14	35	31	26	1.17	10.54	16.34

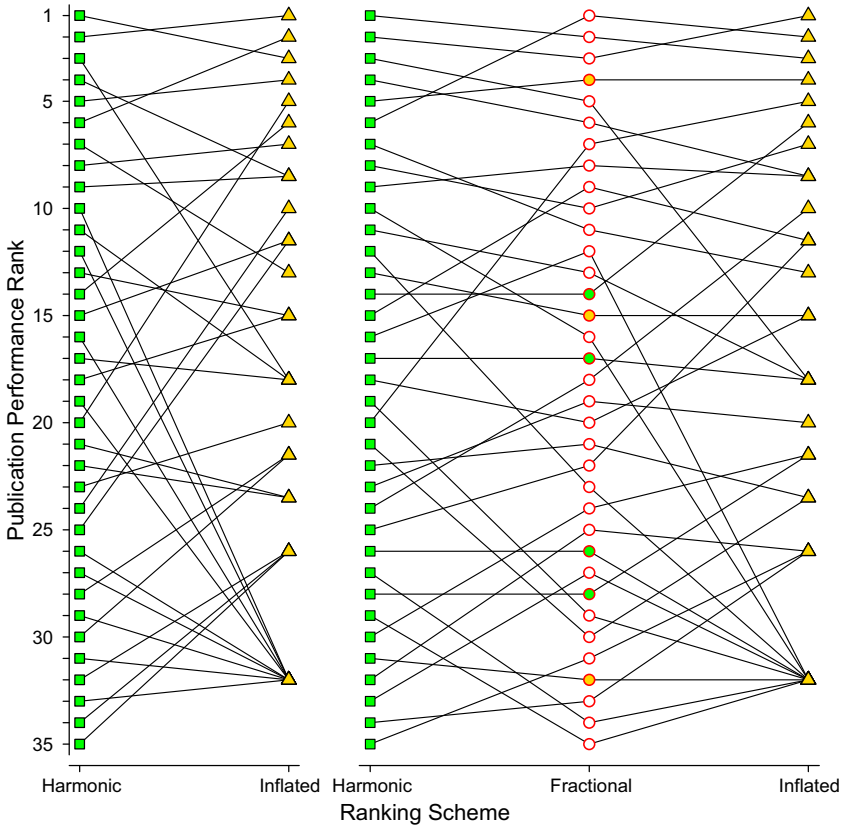

Fig. 2. Comparison of harmonic, fractional and inflated rankings of a select group of 35 leading Canadian aquaculture researchers. The harmonic ranking is based on estimates of actual authorship credit. Fractional ranking shows the effect of equalizing bias generated by assuming that all coauthors have contributed equally. The inflated ranking illustrates the effect of inflationary bias generated by using contribution count as a proxy for authorship credit. Lines from harmonic to inflated show the combined effect of equalizing and inflationary bias, lines from harmonic to fractional show the effect of equalizing bias, and lines from fractional to inflated show the added effect of inflationary bias relative to the fractional ranking. Horizontal lines indicate coinciding ranks, and crossing lines indicate bias-induced rank displacement.

Table 2
Rank correlation analysis of Canadian aquaculture researchers. The relationship between rankings of harmonic (H), fractional (F), and inflated (I) estimates of coauthor credit is quantified using Kendall's rank correlation coefficient τ. Kendall's coefficient of partial rank correlation $\tau_{HI.F}$ is a measure of the correlation between the harmonic and inflated rankings after eliminating the effect of their correlation with the fractional ranking.

Comparison	Coefficient	Rank correlation	P-value
Harmonic vs. Inflated	τ_{HI}	0.4623	<0.001
Harmonic vs. Fractional	τ_{HF}	0.6908	<0.001
Inflated vs. Fractional	τ_{IF}	0.6515	<0.001
Partial rank correlation			
Harmonic vs. Inflated	$\tau_{HI.F}$	0.0225	>0.05[a]

[a] Maghsoodloo and Pallos (1981), Table VI therein.

4. Discussion

Biased equal credit scores, whether fractional or inflated, produced rankings that were fundamentally different from the ranking based on harmonic estimates of actual coauthor contribution. The harmonic, fractional and inflated rankings would have coincided if all of the Canadian aquaculture researchers had contributed equally and had the same number of coauthors. But there was no unequivocal indication of equal contribution and the number of coauthors varied, and so the lack of agreement between the rankings is explained entirely by the effects of equalizing bias and inflationary bias.

Equalizing bias skews credit scores by systematically favouring secondary authors at the expense of primary authors, and inflationary bias favours researchers who have many coauthors and make small contributions, over researchers who have few coauthors and make larger contributions (Hagen, 2008). Both biases had detrimental effects on the accuracy of the rankings.

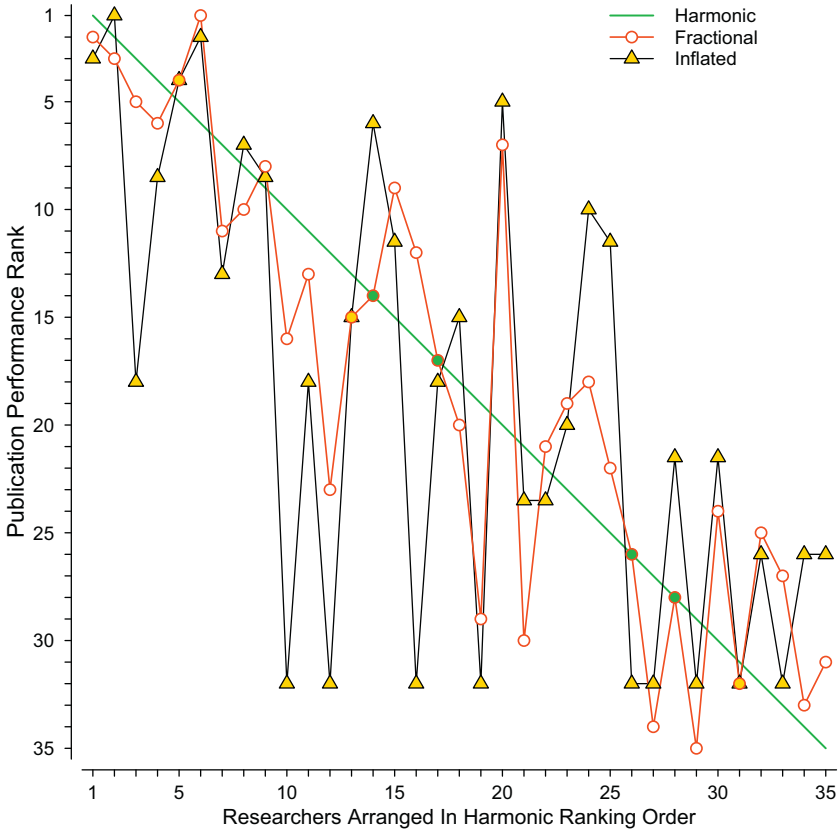

Fig. 3. Comparison of bias-induced rank displacement. The green diagonal reference line represents the harmonic ranking of the Canadian aquaculture researchers. Rank displacement is indicated by distance from the diagonal line. Inflated ranks show the combined effect of equalizing and inflationary bias. Fractional ranks show the residual effect of equalizing bias after correcting for inflationary bias.

Canadian aquaculture researchers' publication performance has been ranked with inflated publication scores on two previous occasions (Picard-Aitken & Coté, 2010; Sylvain, 1993). The present analysis is based on a replicate data set matching the number of papers for researchers included in the most recent ranking. It shows that both inflated and fractional rankings were seriously affected by bias-induced rank displacement (Figs. 2 and 3). In other words, fractional credit scores, although corrected for inflationary bias, are still confounded by equalizing bias.

Interestingly, when Pravdic and Oluic-Vukovic (1991) removed inflationary bias from the "normal" inflated publication score of 27 Croatian chemistry researchers, then 6 of the researchers were no longer included among the top 27 (Fig. 4). Pravdic and Oluic-Vukovic's (1991) result shows that inflationary bias not only has the potential to rearrange rankings, but that nearly 25% of the prospective competitors had been prematurely excluded from the ranking because they were missed by the inflated ranking procedure. By analogy, their result implies that some leading Canadian Aquaculture researchers may have been prematurely excluded from the ranking as a consequence of a biased selection procedure relying on inflated credit scores (Picard-Aitken & Coté, 2010).

Previous studies extolling the advantages of fractional credit did not examine the effect of equalizing bias (e.g. Aksnes et al., 2012; Gauffriau et al., 2008; Lindsey, 1980; Piro et al., 2013; Põder, 2010; Price, 1981). For example, after fractionalizing Piro et al. (2013) observed large rank displacement with a trend towards rank reversal when Norwegian researchers in 37 subfields of science were ranked according to the average number of publications per person in each subfield (Fig. 4). But neither Piro et al. (2013) nor Pravdic and Oluic-Vukovic (1991) provided a comparative reference ranking based on estimates of actual coauthor contribution. Their inflated and fractional rankings are both influenced by equalizing bias, and objective assessment of accuracy relative to a biased reference ranking is not possible (Fig. 4).

Aggregate rankings from the literature also document large effects of fractionalizing (e.g. Aksnes et al., 2012; Gauffriau & Larsen, 2005; Gauffriau et al., 2008; Huang & Lin, 2011; Huang et al., 2011). Aggregate rankings may include additional inaccuracies because they rely on proxy indicators of publication output which accredit nations or institutions according to a variety of inexact "counting methods" (Gauffriau & Larsen, 2005; Larsen, 2008). Such "methods" rely on imprecise heuristic approximations of authorship credit, i.e. proxy indicators based on the kind of information that can be conveniently extracted from a database. For example, by accrediting every nation or institution mentioned in the authors' address list with one full

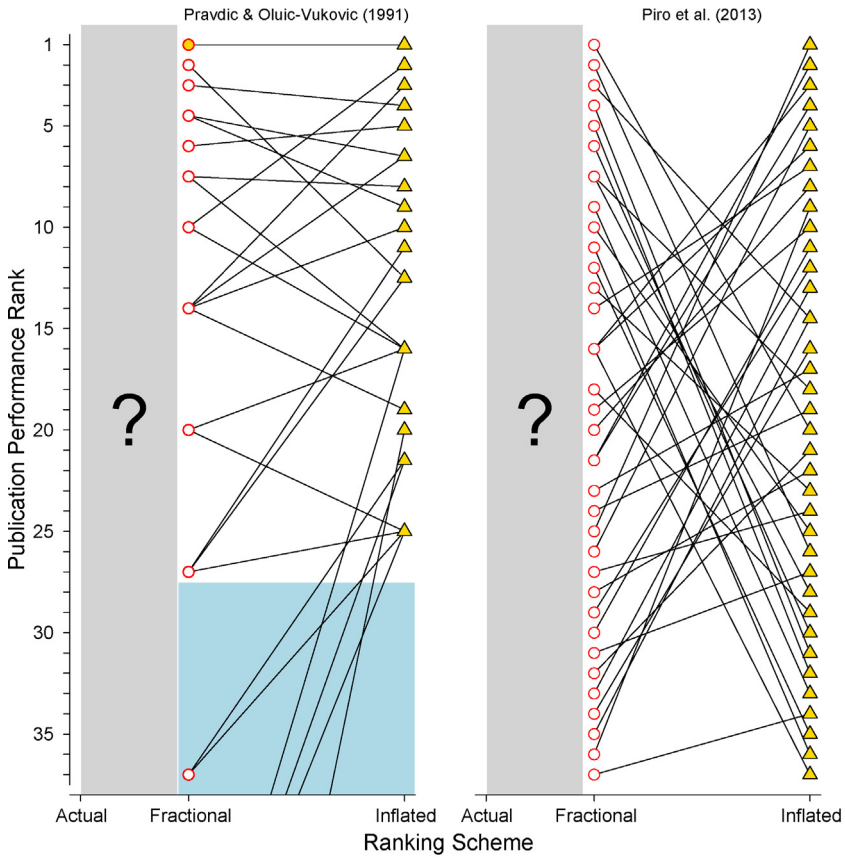

Fig. 4. Comparison of fractional and inflated rankings from the literature. Pravdic and Oluic-Vukovic (1991) ranked a subset of 27 Croatian chemistry researchers. The light blue rectangle identifies 6 researchers who were no longer among the top 27 when publication scores were fractionalized. Piro et al. (2013) ranked Norwegian researchers in 37 subfields of science according to the average number of publications per person in each subfield. Inflated ranks are affected by a combination of equalizing and inflationary bias. Fractional ranks are corrected for inflationary bias. Question marks indicate that the effect of equalizing bias is unaccounted for in the absence of a ranking based on estimates of actual coauthor contribution. Lines show the added effect of inflationary bias relative to the fractional ranking. Horizontal lines indicate coinciding ranks, and crossing lines indicate displacement of fractional ranks caused by inflationary bias.

credit, or one equal fraction of credit, irrespective of the number of coauthors from the respective nation or institution, and irrespective of the coauthors' actual contribution.

In conclusion, any identifiable amount of bias in authorship accreditation was detrimental to the accuracy of the ranking of the Canadian aquaculture researchers' publication output. Bias arose when publication scores were calculated without taking into account information about multiple authorship and differential coauthor contributions. The ensuing biased equal credit scores, whether fractional or inflated, produced rankings that were fundamentally different from the ranking of harmonic estimates of actual credit based on all relevant byline information in the source data. The results demonstrate that accurate accreditation of coauthors is the key to reliable publication performance rankings.

Acknowledgements

Thanks to H. K. Marshall for improving the style and flow of the manuscript, and thanks to the University of Nordland for facilitating data analysis and manuscript preparation.

References

Aksnes, D. W., Schneider, J. W., & Gunnarson, M. (2012). Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, *6*, 36–43.
Buehring, G. C., Buehring, J. E., & Gerard, P. D. (2007). Lost in citation: Vanishing visibility of senior authors. *Scientometrics*, *72*(3), 459–468.
De Bellis, N. (2009). *Bibliometrics and citation analysis. From the science citation index to cybermetrics*. Plymouth, UK: Scarecrow Press.
Gauffriau, M., & Larsen, P. O. (2005). Counting methods are decisive for rankings based on publication and citation studies. *Scientometrics*, *64*(1), 85–93.
Gauffriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., & von Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, *77*(1), 147–176.

Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLoS ONE*, *3*(12), e4021: 4021–4027
Hagen, N. T. (2010). Harmonic publication and citation counting: Sharing authorship credit equitably – Not equally, geometrically or arithmetically. *Scientometrics*, *84*(3), 785–793.
Hagen, N. T. (2013). Harmonic coauthor credit: A parsimonious quantification of the byline hierarchy. *Journal of Informetrics*, *7*(4), 784–791.
Hodge, S. E., & Greenberg, D. A. (1981). Publication credit. *Science*, *213*(4511), 950.
Huang, M. H., & Lin, C. S. (2011). Counting methods & university ranking by H-Index. *Proceedings of the American Society for Information Science and Technology*, *48*(1), 1–6.
Huang, M. H., Lin, C. S., & Chen, D. Z. (2011). Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact. *Journal of The American Society for Information Science and Technology*, *62*(12), 2427–2436.
Kendall, M., & Gibbons, J. D. (1990). *Rank correlation methods* (5th ed.). London: Charles Griffin.
Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), 81–93.
Kendall, M. G. (1942). Partial rank correlation. *Biometrika*, *32*(3/4), 277–283.
Larsen, P. O. (2008). The state of the art in publication counting. *Scientometrics*, *77*(2), 235–251.
Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. *Social Studies of Science*, *10*(2), 145–162.
Maghsoodloo, S., & Pallos, L. L. (1981). Asymptotic behaviour of Kendall's partial rank correlation coefficient and additional quantile estimates. *Journal of Statistical Computation and Simulation*, *13*, 41–48.
Mattsson, P., Sundberg, C. J., & Laget, P. (2011). Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position. *Scientometrics*, *87*, 99–105.
Picard-Aitken, M., & Coté, G. (2010). *Bibliometric analysis of aquaculture research at DFO and in Canada. Final report*. Montréal, Canada: Science-Metrix., http://www.science-metrix.com/pdf/SM_DFO_Aquaculture_Research.pdf.
Piro, F. N., Aksnes, D. W., & Rørstad, K. (2013). A macro analysis of productivity differences across fields: Challenges in the measurement of scientific publishing. *Journal of The American Society for Information Science and Technology*, *64*(2), 307–320.
Põder, E. (2010). Let's correct that small mistake. *Journal of The American Society for Information Science and Technology*, *61*(12), 2593–2594.
Pravdic, N., & Oluic-Vukovic, V. (1991). Distribution of scientific productivity – Ambiguities in the assignment of author rank. *Scientometrics*, *20*(1), 131–144.
Price, D. D. S. (1981). Multiple authorship. *Science*, *212*(4498), 986.
Sylvain, C. (1993). Canadian research activity in aquaculture: A bibliometric analysis. *Scientometrics*, *27*(3), 295–316.
Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, *6*(4), 700–711.