A two-dimensional bibliometric index reflecting both quality and quantity

Mark Levene, Martyn Harris and Trevor Fenner
Department of Computer Science and Information Systems
Birkbeck, University of London
London WC1E 7HX, U.K.
{mark,martyn,trevor}@dcs.bbk.ac.uk

This paper is dedicated to the memory of Judit Bar-Ilan

Abstract

We propose a two-dimensional bibliometric index that strikes a balance between quantity (as measured by the number of publications of a researcher) and quality (as measured by the number of citations to those publications). While the square of h-index is determined by the maximum area square that fits under the citation curve of an author when plotting the number of citations in decreasing order, the rec-index is determined by the maximum area rectangle that fits under the curve. In this context we may distinguish between authors with a few very highly-cited publications, who may have carried out some *influential* research, and *prolific* authors, who may have many publications but fewer citations per publication. The influence of a researcher may be measured via a restricted version of the rec-index, the rec_I-index, which is the maximum area vertical rectangle that fits under the citation curve. Similarly, the prolificity of a researcher may be measured via the rec_P-index, which is the maximum area horizontal rectangle that fits under the citation curve. This leads to the proposal of the two-dimensional bibliometric index (rec_I, rec_P), which captures both aspects of a researcher's output. We present a comprehensive empirical analysis of this two-dimensional index on two datasets: a large set of Google Scholar profiles (representing "typical" researchers) and a small set of Nobel prize winners. Our results demonstrate the potential of this two-dimensional index, since for both data sets there is a statistically significant number of researchers for whom rec_I is greater than rec_P. In particular, for nearly 25% of the Google Scholar researchers and for nearly 60% of the Nobel prize winners, rec_I is greater than rec_P.

Keywords: h-index, rec-index, bibliometric index, two-dimensional index, quantity versus quality, influential versus prolific

1 Introduction

The dilemma in bibliometrics [REG18] regarding the significance of quality versus quantity in evaluating academic research performance is still an ongoing debate [Sah11]. One common perspective is to view the number of publications of a researcher (P) as a measure of quantity and the total number of citations to these publications (C) as a measure of quality. Adopting this point of view, the output of a researcher is described in terms of P and C, although outputs other than publications should also be considered when assessing the worth of a researcher [Piw13]. Alternative metrics, known as altmetrics, include patents, software artifacts, and also online metrics such as download statistics and social media presence [BI18].

Several alternatives to P and C, such as the average number of citations per publication, the number of citations to the top, or the top 10, most cited publications, and the number of publications with at least 10 citations, have also been suggested [Hau16]. Although these simple metrics tend to emphasise only one facet of a researcher's impact, several other bibliometric indices, such as the h-index [Hir05], the g-index [Egg06] and generalisations of these [vW08], aim at a fairer balance between the numbers of citations and publications (see [TB16, REG18] for thorough reviews of these and other bibliometric indices).

We assume that a researcher publishes n cited publications, where $n \geq 0$. These are represented by a *citation vector* of *positive* integers, $\mathbf{x} = \langle x_1, x_2, \ldots, x_n \rangle$, where x_i is the number of citations to publication i, and these are sorted in descending order, i.e., $x_i \geq x_j$ for $1 \leq i < j \leq n$. The *citation curve* is the curve obtained by plotting the number of citations against the ranking of the publications as a histogram of the citation vector.

Whereas the square of the h-index, the h^2-index, is the area of the largest square that fits under the citation curve of a researcher, the *rec*-index (or *rectangle*-index) is the area of the largest rectangle that fits under the curve. (The standard formal definition of the h-index is given in Section 2.) Thus *rec* is a natural generalisation of h^2. They are both geometric indices that consider an area under the citation curve as encapsulating the essential citations for a set of *core* publications that in some sense represent the output of a researcher. Indeed, both these indices include the same number of citations for each core publication.

Our focus here is on an extension of joint research with the late Judit Bar-Ilan on the *rec*-index (which is the square of the χ-index) [FHLB18, LFB19] and, as mentioned above, can be viewed as a geometric generalisation of the h-index. In the context of joint work with the late Judit Bar-Ilan, we mention the *hw*-rank [BIL15], which is a variant of the h-index for ranking Web search engine results. Moreover, referring to the late Judit Bar-Ilan's research on the h-index, the highly-cited paper [BI08] highlights the important problem that computing the h-index from different bibliometric sources may lead to conflicting results because of subject-specific differences in the coverage of the sources.

To motivate the *rec*-index, consider the following three citation vectors, the citation curves of which are depicted in Figure 1: (i) $\langle 100 \rangle$, i.e. 1 publication with 100 citations, (ii) $\langle 10, 10, \cdots, 10 \rangle$, i.e. 10 publications with 10 citations each, and (iii) $\langle 1, 1, \cdots, 1 \rangle$, i.e. 100 publications with 1 citation each. (Note that the diagram in Figure 1 is not drawn to scale.)

It can be seen that all three researchers have a *rec*-index of 100, while researcher (ii) has an h^2-index of 100, but researchers (i) and (iii) both have an h^2-index of only 1. The h^2-index may be seen as balancing *quality*, on the one hand, by favouring publications with a higher number of citations and *quantity*, on the other hand, by taking into account all publications with a sufficient number of citations. However, such an approach disadvantages a researcher, such as (i), with a few very highly-cited publications, who may have carried out some *influential* seminal research, and it also disadvantages a *prolific* researcher, such as (iii), who may have many publications but with fewer citations per publication.

Now, as a further example, consider two citation vectors: (i) $\mathbf{x_1}$, with $n = 90$, having 1 publication with 100 citations and a further 89 publications with 1 citation each, and (ii) with $n = 100$, $\mathbf{x_2}$, having 1 publication with 90 citations and a further 99 publications with

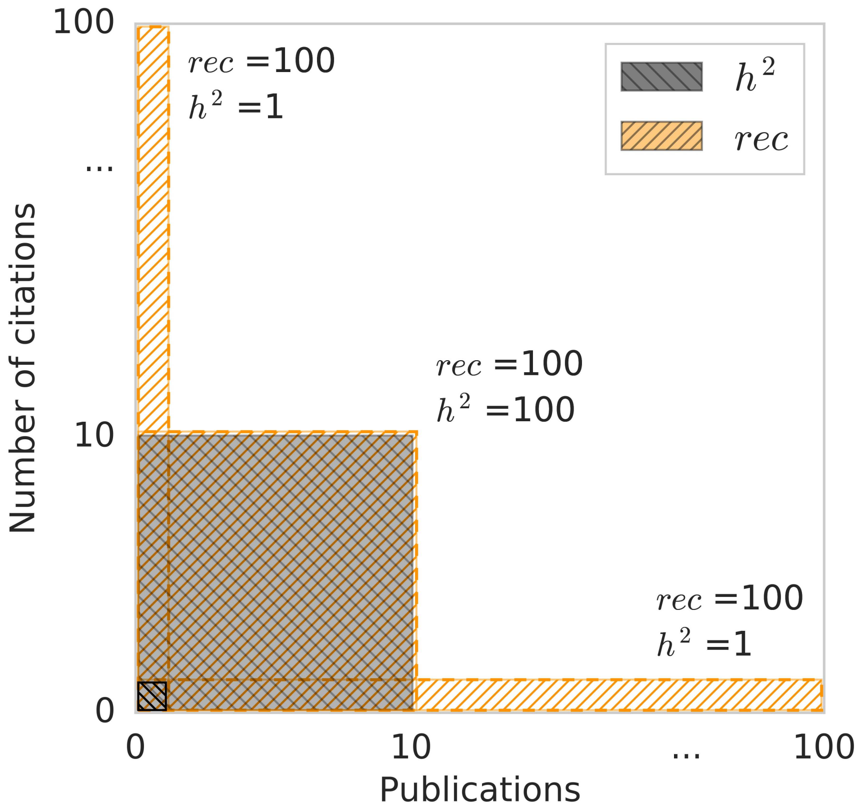


Figure 1: Example of the geometric interpretation of the h^2 and rec indices.

1 citation each. It is evident that $rec(\mathbf{x_1}) = 100$, which is represented by a vertical rectangle having width = 1 and height = 100, like that in Figure 1, and that $rec(\mathbf{x_2}) = 100$, which is represented by a horizonal rectangle having width = 100 and height = 1, like that in Figure 1.

A vertical rectangle (with height greater than or equal to width) represents the *influence* of a researcher, and a horizontal one (with width greater than or equal to height) the *prolificity* of the researcher. In fact, as suggested in [LFB19], we can represent a researcher by the pair of indices $(rec_I(\mathbf{x}), rec_P(\mathbf{x}))$, where, $rec_I(\mathbf{x})$ is the area of the largest vertical rectangle under the citation curve of the citation vector $\mathbf{x}$, and $rec_P(\mathbf{x})$ is the area of the largest horizonal rectangle under the citation curve of $\mathbf{x}$ (the formal definitions are given in Section 2). For simplicity of notation, we abbreviate these to rec_I and rec_P when $\mathbf{x}$ is apparent from context. We note that, when the width and height of either of the rectangles are equal, we have $rec_I = h^2$, $rec_P = h^2$ or both.

So, in our example, the two-dimensional index $(rec_I(\mathbf{x_1}), rec_P(\mathbf{x_1})) = (100, 90)$ and $(rec_I(\mathbf{x_2}), rec_P(\mathbf{x_2})) = (90, 100)$. This indicates that the first researcher is more influential than prolific and the second is more prolific than influential. We believe that the (rec_I, rec_P)-index is a sensible way of presenting the balance of quality and quantity of a researcher's publication output. Such multi-dimensional indices are in line with the conclusion in [WSL14] that several bibliometric indicators should be used to gauge the overall impact of a researcher.

The rest of the paper is organised as follows. In Section 2, we introduce and formally define the two-dimensional index (rec_I, rec_P). In Section 3, we discuss the two data sets that we analyse: a large Google Scholar data set, described in Subsection 3.1, and a small data set of Nobel prize winners, described in Subsection 3.2. In Section 4, we present the main analysis of the data sets and results obtained. In Subsection 4.1, we analyse the Google Scholar data set and, in Subsection 4.2, we consider the Nobel prize winners data set. Finally, in Section 5, we give our concluding remarks.

2 One- and two-dimensional bibliometric indices

To introduce the notion of a *bibliometric index*, we recall that for a citation vector $\mathbf{x}$ we only consider publications that have at least one citation; we assume that $\mathbf{x} = \langle\rangle$ in the absence of any citations. A *one-dimensional bibliometric index* is a function f that maps citation vectors to the set of non-negative real numbers; for convenience we refer to it as the f-index. As in [Woe08], we assume the *baseline condition* that, for the empty citation vector, $f(\langle\rangle) = 0$. (We will often write just f for $f(\mathbf{x})$.) A *two-dimensional bibliometric index* is a pair (f_1, f_2), where f_1 and f_2 are one-dimensional bibliometric indices.

The h-index [Hir05] is defined as the maximum number h of a researcher's publications such that each has at least h citations, i.e., for a citation vector $\mathbf{x}$, the h-index is the largest rank h for which $x_h \geq h$. In the absence of any citations $h = 0$, which is consistent with the baseline condition. The h-index is completely insensitive to the fact that a researcher's top few publications may have many more than h citations, and conversely that another researcher may have a large number of publications having slightly fewer than h citations [BD07]. For comparison with the rec-index, we will use the h^2-index, i.e., the square of the h-index.

The rec-index is defined as the maximum area rectangle that can fit under the citation curve, while the h^2-index is the maximum area square that can fit under the citation curve. Formally, the rec-index (or *rectangle*-index) of a researcher with citation vector $\mathbf{x}$ is defined by

$$rec(\mathbf{x}) = \max_i ix_i. \tag{1}$$

(For completeness, we mention that the χ-index, introduced in [FHLB18], satisfies $\chi(\mathbf{x}) = \sqrt{rec(\mathbf{x})}$.)

Now, if we let k denote a value of i that maximises ix_i in (1), the rec-index can distinguish between more *influential* researchers for whom $x_k > k$ (such as the researcher represented by the vertical rectangle in Figure 1) and more *prolific* researchers for whom $k > x_k$ (such as the researcher represented by the horizontal rectangle in Figure 1). In this sense, the rec-index avoids the debate of number of citations versus number of publications by awarding all three researchers depicted in Figure 1 the same rec-index of 100.

We now formally define the two-dimensional index (rec_I, rec_P) introduced in [LFB19]. We first define the *publication vector* $\mathbf{p} = \langle p_1, p_2, \ldots, p_m\rangle$ as the *conjugate partition* of the citation vector $\mathbf{x} = \langle x_1, x_2, \ldots, x_n\rangle$, where $m = x_1$ and p_i is the number of publications with at least i citations [Woe08]. Geometrically, the publication vector is obtained by reflecting the histogram of the citation vector along the main diagonal.

Some indices tend to emphasise *influence*, for example, the maximum citation index x_1, whereas others, such as the publication count n, emphasise *prolificity*. In order to emphasise influence rather than prolificity, we define a version of the rec-index, the rec_I-index, in which we restrict the maximum in (1) to be over those i for which $i \leq x_i$. Conversely, to emphasise prolificity, we define the conjugate index, the rec_P-index, in terms the corresponding publication vector $\mathbf{p}$, restricting i so that $i \leq p_i$.

It is evident that the following two inequalities hold:

$$rec \geq rec_I \geq h^2, \tag{2}$$

$$rec \geq rec_P \geq h^2, \tag{3}$$

and also that

$$rec = max(rec_I, rec_P).$$

From inequalities (2) and (3), it follows that if $rec = h^2$ then

$$rec = rec_I = rec_P = h^2.$$

The above inequalities will be useful in the analysis of the results in Section 4.

3 Data sets and preliminary analysis

We now introduce the two data sets, which were also used in [FHLB18]. We then give some basic statistics for these data sets and compute the Spearman rank-correlation coefficients [Ros11] (Spearman $\mathbf{r}$) between various indices for the researchers concerned. In Subsection 3.1, we consider the Google Scholar data set and then, in Subsection 3.2, we consider the data set of Nobel prize winners.

3.1 Google scholar data set

For our main analysis, we made use of a large data set of Google Scholar profiles compiled and made available by Radicchi and Castellano [RC13]. This data set was used to investigate the rec-index in [FHLB18]; however, for completeness we repeat the relevant details. The full data set contains approximately 90,000 citation vectors of authors across all disciplines, collected between June 29 and July 4, 2012. As in [RC13], we only included authors who had validated their Google Scholar accounts and we removed authors with fewer than twenty publications. We also removed all uncited publications and publications dated before 1945. We then filtered the data set by only including authors having a career of at least five years, where the career is deemed to have began from the year of the first paper published after 1944. After this preprocessing step, the reduced data set contained 34,393 citation profiles. For 71.70% of these profiles $rec = rec_I > rec_P$, for 24.39% $rec = rec_P > rec_I$, and $rec = rec_I = rec_P$ for the remaining 3.91%.

In Table 1, we present the basic statistics for h^2, rec, rec_I and rec_P. It can be seen that the statistics of rec_I and rec are similar, as are those of rec_P and h^2. This observation is supported by Table 2, which gives the Spearman rank-correlation coefficients between the four indices. The correlations are all high, and are exceptionally high between rec_I and rec, and between rec_P and h^2.

Statistics	h^2	rec	rec_I	rec_P
mean	510.20	814.47	800.63	544.75
median	225.00	363.00	357.00	255.00
min	4.00	21.00	9.00	20.00
max	45369.00	48702.00	48702.00	45780.00
std	900.79	1587.92	1576.65	943.02

Table 1: Basic statistics for the Google Scholar data set.

Spearman r	h^2	rec	rec_I	rec_P
h^2	1.000	0.932	0.928	0.994
rec		1.000	0.998	0.926
rec_I			1.000	0.916
rec_P				1.000

Table 2: Spearman rank-correlation between the indices for the Google Scholar data set.

3.2 Nobel prize winners data set

For our second data set, we collected the citation vectors of 99 Nobel prize winners across a variety of disciplines from the Web of Science platform [Cla18]. Like the Google Scholar data set, this data set was also used to investigate the rec-index in [FHLB18], but again we repeat the relevant details. We only included authors having twenty or more publications, and only those publications with citations. However, for this data set, we considered their full careers without a cutoff date. For 84.85% of these profiles $rec = rec_I > rec_P$, for 15.15% $rec = rec_P > rec_I$, and there were no profiles for which $rec_I = rec_P$.

In Table 3, we present the basic statistics for the Nobel laureates, while in Table 4 we present the Spearman rank-correlation coefficients. As one would expect, the statistics are much higher than for the Google Scholar data set. Again, the correlations between rec_I and rec, and between rec_P and h^2 are exceptionally high. However, the other four correlations are significantly lower than for the Google Scholar data set, but are all quite similar. This is not surprising since Nobel prize winners tend to be far more influential than prolific.

Statistics	h^2	rec	rec_I	rec_P
mean	5680.73	8931.26	8726.54	5985.53
median	4225.00	7105.00	6342.00	4225.00
min	144.00	308.00	308.00	144.00
max	38025.00	45736.00	45736.00	41664.00
std	6078.78	8560.01	8372.19	6511.60

Table 3: Basic statistics for the Nobel prize winners data set.

Spearman r	h^2	rec	rec_I	rec_P
h^2	1.000	0.879	0.858	0.996
rec		1.000	0.995	0.865
rec_I			1.000	0.848
rec_P				1.000

Table 4: Spearman rank-correlation between the indices for the Nobel prize winners data set.

4 Analysis and results

We now analyse the data sets introduced in Section 3 with the aim of exploring how authors are separated into three different classes. More specifically, we consider three classes according to whether (*C1*) $rec_I > rec_P$, (*C2*) $rec_I < rec_P$ or (*C3*) $rec_I = rec_P$, or rather which of these holds at a significance level of 99%.

The corresponding confidence intervals were constructed using the *bootstrap method* [DH97], which is a technique for computing a statistic that relies on random resampling with replacement from a given data set. The bootstrap method is usually nonparametric, making no distributional assumptions about the data set. The specific version that we use to classify the authors is the *basic bootstrap percentile method* [DH97, Section 5.3.1]; see also [AH15], which used the bootstrap method in the context of bibliometrics. More specifically, we resample with replacement from an author citation vector 1000 times, compute the four indices for each resample, and then compute a one-sided 99% confidence interval for each index. The bootstrap confidence intervals allow us to determine, for a given author, whether one index is *approximately equal to* ($\approx$) or *approximately greater than* ($\succ$) another index at a significance level of 99%.

4.1 Results for Google scholar data set

In Table 5 we show the breakdown into the three classes: we see that more than 75% of the authors fall into *C3* (for which $rec_I \approx rec_P$), while nearly 25% fall into class *C1* (for which $rec_I \succ rec_P$), and only 0.5% fall into class *C2* (for which $rec_I \prec rec_P$). We see that there is a significant number of more influential researchers who are clearly separated from the others, and who are fairly balanced in terms of being influential or prolific.

Class		# authors	% authors
C1:	$rec_I \succ rec_P$	8204	23.85
C2:	$rec_I \prec rec_P$	181	0.53
C3:	$rec_I \approx rec_P$	26008	75.62

Table 5: Breakdown of the three classes for the Google Scholar data set, where the total number of authors is 34393.

A more detailed analysis of *C1* in Table 6 shows that in the vast majority of cases rec_P is approximately equal to h^2 (*C12*) rather than being approximately greater than h^2 (*C11*); see two examples of researchers in *C1* in Figure 2. For those few researchers in *C2* who tend to be prolific, we see from Table 7 that in all cases rec_I is approximately equal to h^2 (*C22*) rather than being approximately greater than h^2 (*C21*); see an example of a researcher in *C2* in Figure 3. In Table 8 we see that for *C3* there is an overwhelming number of researchers for whom $rec_I \approx rec_P \approx h^2$ (*C32*). On the other hand, there is a small minority for whom $rec_I \approx rec_P \succ h^2$ (*C31*); see two examples of researchers in *C3* in Figure 4.

C1:	$rec_I \succ rec_P$	# authors	% authors
C11:	$rec_P \succ h^2$	19	0.23
C12:	$rec_P \approx h^2$	8185	99.77

Table 6: Breakdown of the class *C1* for the Google Scholar data set, where the total number of authors is 8204.

C2:	$rec_P \succ rec_I$	# authors	% authors
C21:	$rec_I \succ h^2$	0	0.00
C22:	$rec_I \approx h^2$	181	100.00

Table 7: Breakdown of the class *C2* for the Google Scholar data set, where the total number of authors is 181.

C3:	$rec_I \approx rec_P$	# authors	% authors
C31:	$rec_I \approx rec_P \succ h^2$	26	0.10
C32:	$rec_I \approx rec_P \approx h^2$	25982	99.90

Table 8: Breakdown of the class *C3* for the Google Scholar data set, where the total number of authors is 26008.

4.2 Results for Nobel prize winners data set

In Table 9 we show the breakdown into the three classes, analogous to Table 5. We see that nearly 60% fall into *C1* (for which $rec_I \succ rec_P$), while nearly 40% fall into class *C3* (for which $rec_I \approx rec_P$), and only 2% fall into class *C2* (for which $rec_I \prec rec_P$). Thus for Nobel prize winners, as opposed to "typical" Google Scholar authors, there is a clear majority of influential researchers.

A more detailed analysis of *C1* in Table 10 shows that rec_P is always approximately equal to h^2 (*C12*); see an example of a researcher in *C1* in Figure 5. For the two researchers in *C2* who tend to be prolific, we see from Table 11 that in this case rec_I is always approximately equal to h^2 (*C22*); see an example of a researcher in *C2* in Figure 6. In Table 12 we see that for *C3* it is always the case that $rec_I \approx rec_P \approx h^2$ (*C32*); see an example of a researcher in *C3* in Figure 7.

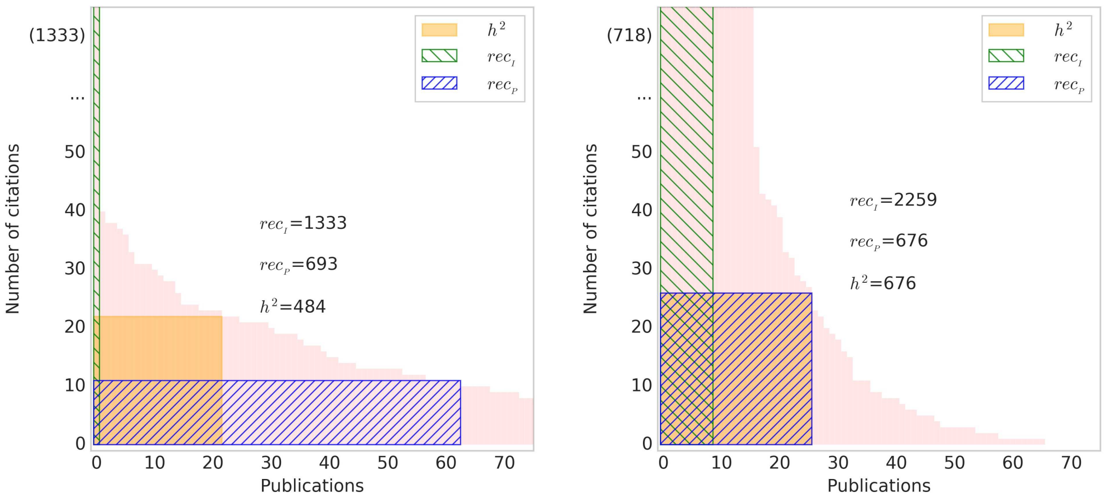


Figure 2: Examples for class *C1* of Google Scholar authors for which: $rec_I \succ rec_P \succ h^2$ (left), $rec_I \succ rec_P \approx h^2$ (right).

Class		# authors	% authors
C1:	$rec_I \succ rec_P$	59	59.60
C2:	$rec_I \prec rec_P$	2	2.02
C3:	$rec_I \approx rec_P$	38	38.38

Table 9: Breakdown of the three classes for the Nobel prize winners data set, where the total number of authors is 99.

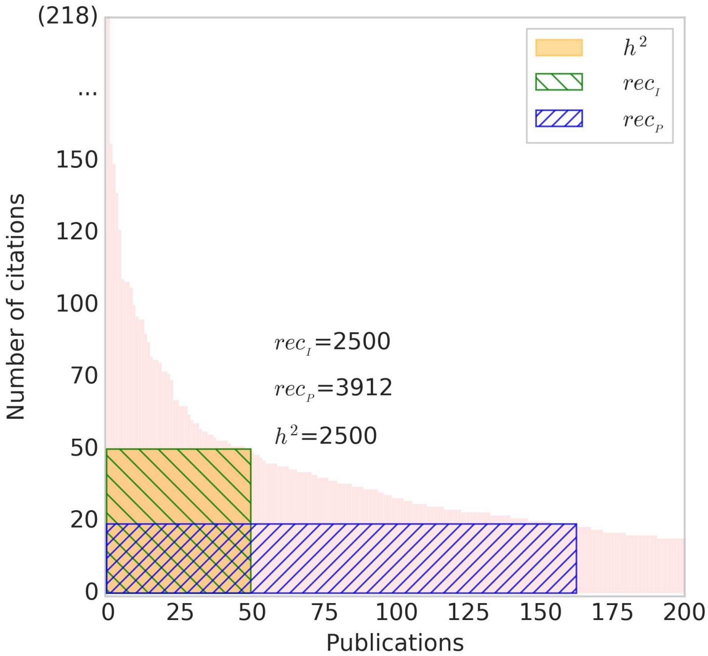
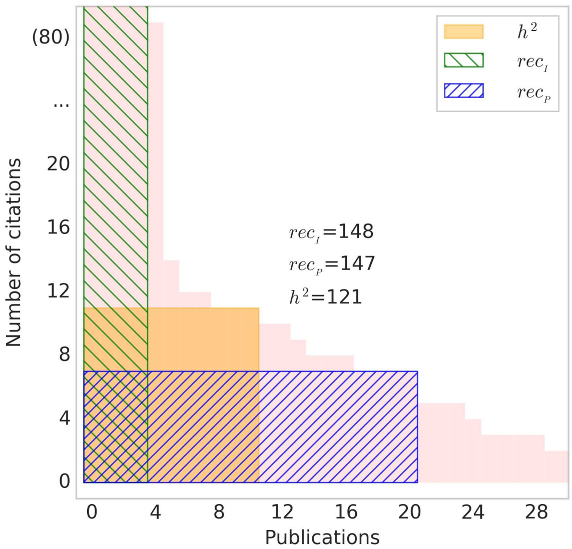


Figure 3: Example for class *C2* of a Google Scholar author for which: $rec_P \succ rec_I \approx h^2$.



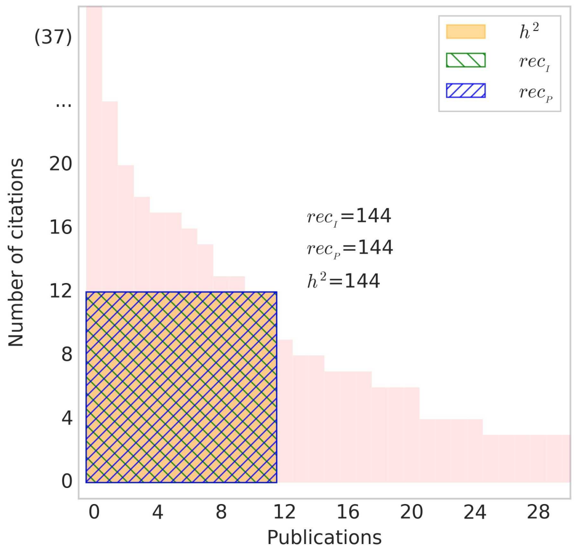


Figure 4: Examples for class *C3* of Google Scholar authors for which: $rec_I \approx rec_P \succ h^2$ (left), $rec_I \approx rec_P \approx h^2$ (right).

C1: $rec_I \succ rec_P$	# authors	% authors
C11: $rec_P \succ h^2$	0	0.00
C12 $rec_P \approx h^2$	59	100.00

Table 10: Breakdown of the class *C1* for the Nobel prize winners data set, where the total number of authors is 59.

C2: $rec_P \succ rec_I$	# authors	% authors
C21: $rec_I \succ h^2$	0	0.00
C22: $rec_I \approx h^2$	2	100.00

Table 11: Breakdown of the class *C2* for the Nobel prize winners data set, where the total number of authors is 2.

C3: $rec_I \approx rec_P$	# authors	% authors
C31: $rec_I \approx rec_P \succ h^2$	0	0.00
C32: $rec_I \approx rec_P \approx h^2$	38	100.00

Table 12: Breakdown of the class *C3* for the Nobel prize winners data set, where the total number of authors is equal to 38.

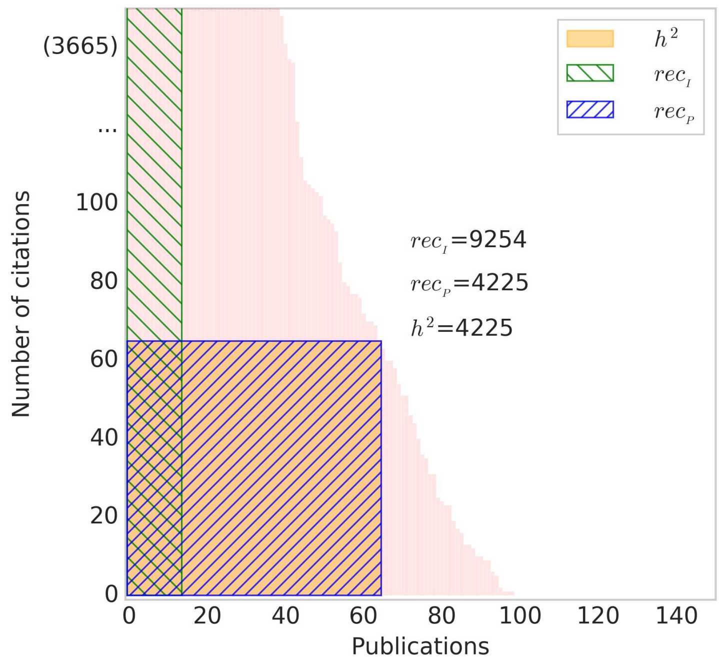


Figure 5: Example for class *C1* of one of the two Nobel prize winners for which: $rec_I \succ rec_P \approx h^2$, Greider, Physiology/Medicine, 2009.

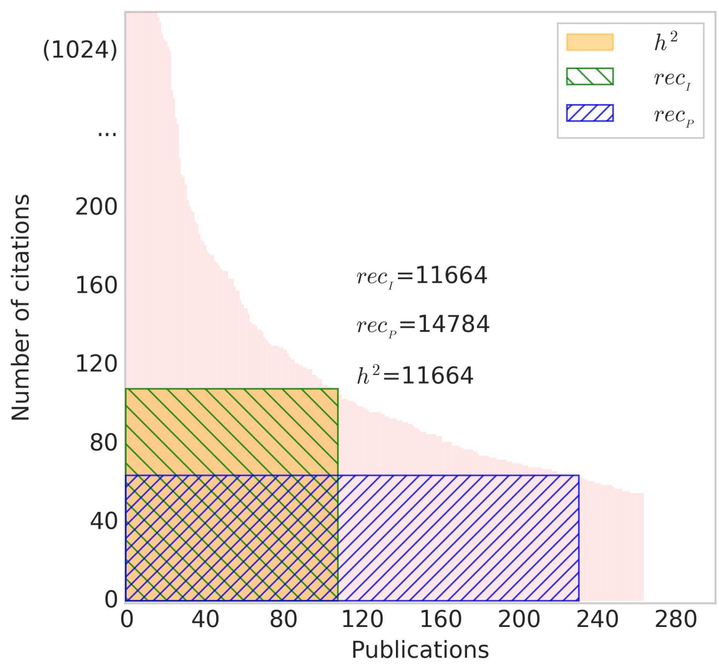


Figure 6: Example for class *C2* of one of the two Nobel prize winners for which: $rec_P \succ rec_I \approx h^2$, Schrock, Chemistry, 2005.

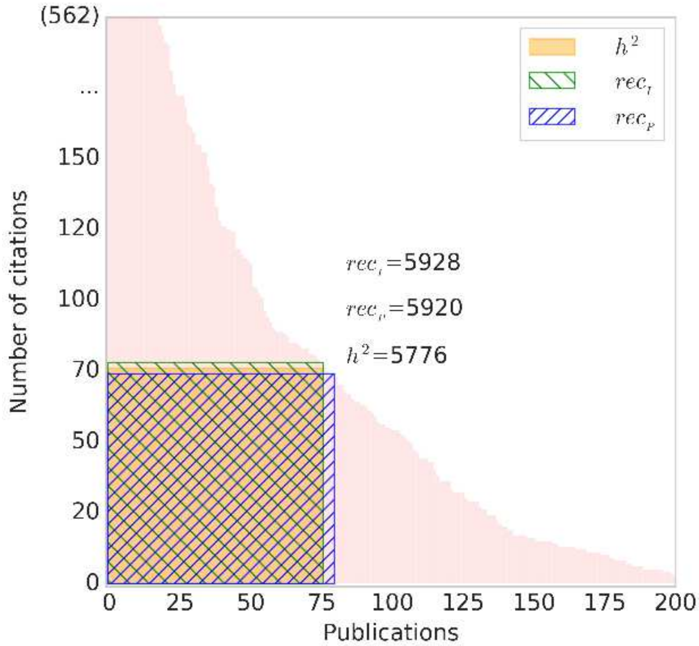


Figure 7: Example for class *C3* of a Nobel prize winner for which: $rec_I \approx rec_P \approx h^2$, Gurdon, Physiology/Medicine, 2012.

5 Concluding remarks

We have proposed the two-dimensional index (rec_I, rec_P) and shown that it is able to capture two aspects of a reseacher's profile. On the one hand, rec_I measures the influence of a researcher and, on the other hand, rec_P measures the prolificity of a researcher. The results of analysing the two data sets demonstrate the potential of the (rec_I, rec_P)-index, since for both data sets there is a significant number of researchers for whom $rec_I \succ rec_P$: nearly 25% of Google Scholar researchers and nearly 60% of Nobel prize winners. Moreover, the very high correlations between rec_I and rec, and between rec_P and h^2 (see Tables 2 and 4), suggest that we could consider employing (rec, h^2) as a two-dimensional index. However, this index may undervalue the small number of researchers who are prolific. We note that $rec_I \prec rec_P$ only infrequently, as can be seen from Tables 5 and 9. We also see from the tables in Section 4 that, for all three classes *C1*, *C2* and *C3*, the smaller of rec_I and rec_P is approximately equal to h^2 for over 99% of the authors in both data sets.

We believe our investigation of the two-dimensional index (rec_I, rec_P) contributes usefully to the quantity-versus-quality debate. However, more research needs to be done on multi-dimensional indices that are able to capture multiple aspects of a researcher's citation profile.

References

[AH15] J.P. Andersen and S. Haustein. Bootstrapping to evaluate accuracy of citation-based journal indicators. In *Proceedings of the 15th International Society of Scientometrics and Informetrics Conference*, pages 413–414, Istanbul, Turkey, 2015.

[BD07] L. Bornmann and H.-D. Daniel. What do we know about the h index? *Journal of the American Society for Information Science and Technology*, 58:1381–1385, 2007.

[BI08] J. Bar-Ilan. Which h-index? – A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74:257–271, 2008.

[BI18] J. Bar-Ilan. The Journal of Altmetrics is launched – Editorial. *Journal of Altmetrics*, 1:Article 1, 5 pages, 2018.

[BIL15] J. Bar-Ilan and M. Levene. The hw-rank: an h-index variant for ranking web pages. *Scientometrics*, 102:2247–2253, 2015.

[Cla18] Clarivate Analytics. Web of science. See `www.webofknowledge.com`, 2018.

[DH97] A.C. Davison and D.V. Hinkley. *Bootstrap Methods and their Applications.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, UK, 1997.

[Egg06] L. Egghe. Theory and practise of the g-index. *Scientometrics*, 69:131–152, 2006.

[FHLB18] T. Fenner, M. Harris, M. Levene, and J. Bar-Ilan. A novel bibliometric index with a simple geometric interpretation. *PLoS ONE*, 13(7):e0200098, 2018.

[Hau16] K. Hausken. The ranking of researchers by publications and citations: Using RePEc data. *Journal of Economics Bibliography*, 3:530–558, 2016.

[Hir05] J.E. Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 98:16569–16572, November 2005.

[LFB19] M. Levene, T. Fenner, and J. Bar-Ilan. Characterisation of the χ-index and the *rec*-index. *Scientometrics*, 120:885–896, 2019.

[Piw13] H. Piwowar. Altmetrics: Value all research products. *Nature*, 493:159, 2013.

[RC13] F. Radicchi and C. Castellano. Analysis of bibliometric indicators for individual scholars in a large data set. *Scientometrics*, 97:627–637, 2013.

[REG18] R. Rousseau, L. Egghe, and R. Guns. *Becoming Metric-Wise: A Bibliometric Guide for Researchers.* Chandos Information Professional Series. Chandos Publishing, Cambridge MA, 2018.

[Ros11] B. Rosner. *Fundamentals of Biostatistics.* Brooks/Cole, Cengage Learning, Boston, MA, 7th edition, 2011.

[Sah11] J.-A. Sahel. Quality versus quantity: Assessing individual research performance. *Science Translational Medicine*, 3:84cm13, 2011.

[TB16] R. Todeschini and A. Baccini. *Handbook of Bibliometric Indicators: Quantitative Tools for Studying and Evaluating Research.* Wiley-VCH, Weinheim, Germany, 2016.

[vW08] N.J. van Eck and L. Waltman. Generalizing the h- and g-indices. *Journal of Informetrics*, 2:263–271, 2008.

[Woe08] G.J. Woeginger. A symmetry axiom for scientific impact indices. *Journal of Informetrics*, 2:298–303, 2008.

[WSL14] L. Wildgaard, J.W. Schneider, and B. Larsen. A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics*, 101:125–158, 2014.