Scientometrics (2018) 116:147–160
https://doi.org/10.1007/s11192-018-2731-9

Rainbow ranking: an adaptable, multidimensional ranking method for publication sets

Georgios Stoupas[1] · Antonis Sidiropoulos[2] · Antonia Gogoglou[1] · Dimitrios Katsaros[3] · Yannis Manolopoulos[1]

Received: 5 October 2017 / Published online: 16 April 2018

Abstract Various scientometric indices have been proposed in an attempt to express the quantitative and qualitative characteristics of scientific output. However, fully capturing the performance and impact of a scientific entity (author, journal, institution, conference, etc.) still remains an open research issue, as each proposed index focuses only on particular aspects of scientific performance. Therefore, scientific evaluation can be viewed as a *multi-dimensional ranking problem*, where dimensions represent the assorted scientometric indices. To address this problem, the skyline operator has been proposed in Sidiropoulos et al. (J Informetr 10(3):789–813, 2016) with multiple combinations of dimensions. In the present work, we introduce a new index derived from the utilization of the skyline operator, called Rainbow Ranking or RR-index that assigns a category score to each scientific entity instead of producing a strict ordering of the ranked entities. Our RR-index allows the combination of any known indices depending on the purposes of the evaluation and outputs a single number metric expressing multi-criteria relative ranking and can be applied to any scientific entity such as authors and journals. The proposed methodology was

✉ Georgios Stoupas
grgstoupas@csd.auth.gr

Antonis Sidiropoulos
asidirop@it.teithe.gr

Antonia Gogoglou
agogoglou@csd.auth.gr

Dimitrios Katsaros
dkatsar@inf.uth.gr

Yannis Manolopoulos
manolopo@csd.auth.gr

1 Department of Informatics, Aristotle University, 54124 Thessaloniki, Greece

2 Department of Information Technology, Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, Greece

3 Department of Electrical and Computer Engineering, University of Thessaly, Volos, Greece

experimentally evaluated using a dataset of over 105,000 scientists from the Computer Science field.

Keywords Scientometrics · Ranking · *h*-Index · Skyline

Introduction

With the introduction of the well-known *h*-index by Hirsch (2005) a plethora of bibliometric indices have been proposed to better quantify the features of scientific output. Existing indices can be categorized according to the entity under evaluation, i.e. author performance, journal impact or publication level. In this direction, variations of the *h*-index have emerged to employ different ranges in citation accumulation, such as the *g*-index (Egghe 2006), *hg* (Alonso et al. 2010) and h_2 index (Kosmulski 2006). Other approaches perform normalizations of the calculated index over the number of publications, age of publications or citations and include but are not limited to the h_{norm} index (Sidiropoulos et al. 2007), the contemporary *h*-index (h^{cont}) and trend *h*-index (h^t) (Sidiropoulos et al. 2007), *m*-quotient (Hirsch 2005) and Bornmann's *m* (m_{Bor}) (Bornmann et al. 2008). Analogous efforts have been made to measure the distribution of citations that contribute to the calculation of the *h*-index, such as h_{rat} (Ruane and Tol 2008), tapered *h*-index (h_{tap}) (Anderson et al. 2008) and *w* of Wohlin (2009), while another set of works attempt to include different areas around the citation curve into a single number metric, such as the excess or *e*-index (Zhang 2009), the set of indices introduced by Jin (*A*, *R* and *AR* indices) (Jin et al. 2007) and the more recent Perfectionism Index or *PI*-index (Sidiropoulos et al. 2015), etc.

Given the wide range of bibliometric indices available in literature, their interrelations and correlations with one another have been extensively studied in Bornmann et al. (2011), Bollen et al. (2009), Schreiber et al. (2012) and (Bornmann et al. 2014). A taxonomy of 108 author-level indices has been performed in Wildgaard et al. (2014) investigating their usefulness, their complexity of calculation and the different publishing features they represent. All the aforementioned studies, report a high degree of correlation between the *h*-index and its variants, thus identifying overlapping information conveyed by the large variety of existing bibliometric indices. Consequently, to achieve a fair evaluation of scientific output one must consider multiple uncorrelated indices expressing different scientific qualities calculated from bibliometric data.

In an effort to achieve a fair and universal ranking of scientific entities. Wolcott et al. (2015) proposed the use of both time-dependent and -independent factors as part of a classification scheme to assign relative importance ranking to publications based on their probability of being highly cited. Relative performance at publication level has been also incorporated into a co-citation based indicator in Hutchins and Yuan (2016), whereas for journal ranking an attempt to broaden the evaluation of journals using altmetrics is presented in Tamar and Tim (2015). Tsai et al. (2014) identified domain dependent correlations between various journal rankings and produced a unified journal ranking as a combination of existing ones. In Franceschini (2014) the variation of different citers was used as a proxy for publication ranking, whereas Glanzel et al. (2014) proposes the application of performance classes to evaluate research at country and institutional level. Also in Alguliev et al. (2014) it is proposed a weighted index to combine the results from individual indices. The proposed index utilizes a policy of consensus that assigns the weights to indices and linearly combines them.

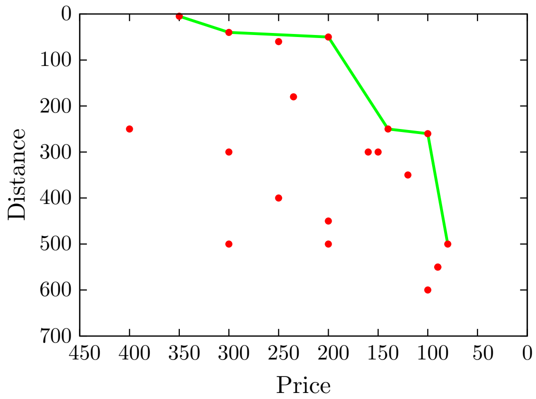


Fig. 1 Skyline example

Each bibliometric index produces a different rank list with strict scores for the same group of scientists. Having to deal with this storm of valuable indicators, the need arises for a general classification scheme of scientific entities according to multiple evaluation metrics. The quantification of abstract concepts such as “scientific impact” allow for tolerance limits to be fair. Therefore, it would be more appropriate to create ranking levels instead of strict scoring, where multiple researchers can be ranked at the same level. In this direction, the use of the *skyline operator* was proposed in a previous work (Sidiropoulos et al. 2016) to select from a set of researchers those that cannot be outperformed by any other from a pool of scientists. In other words, outstanding scientists are identified in one or more dimensions (i.e., features) but not necessarily in all of the dimensions. This way credit can be attributed to various publishing patterns and, thus, scientists that outperform others in certain attributes can also be distinguished. Of crucial importance to the skyline operator’s computation is the selection of dimensions, i.e., the evaluation indicators. Dependent on the perspective of the evaluation to be performed various appropriate metrics may be selected and the resulting skyline set will adjust accordingly.

In the present work we expand upon the concept of the skyline operator by incorporating the dominating groups of scientists into universal ranking levels, and introducing a new indicator based on the relative ranking a scientist has achieved, namely Rainbow Ranking scheme; Fig. 1 justifies the name.

The rest of the paper is organized as follows: The next section describes the dataset used in our analysis. The applied methodologies are presented in the third section, in which we introduce the new indicator and the process of its calculation. In the fourth section, the results of the study are given, whereas conclusions appear in the final section.

Dataset description

For the purposes of this study, we collected data from Microsoft Academic Search (MAS)[1]; in particular, we extracted full citation data starting from year 1950 up to 2015 for scientists publishing in the Computer Science area, as identified by the domain categorization of MAS. All related meta-data about their publication, citation and collaboration network

[1] We appreciate the offer of Microsoft to gratis provide their database API. The API that was used in this work has been discontinued by Microsoft in the summer of 2016.

were also retrieved. The initial query to MAS was intended to select all authors in Computer Science with more than 5 publications. Unfortunately, the publication year is missing for 12% of the publications in our datasets. To complete this missing information, we gathered data from DBLP[2] by using the XML search DBLP API. This way, we recovered about 6% of the missing information. After that step, the publications with missing year were ignored as well as the citations from publications with missing year. In the end, we collected 104,190 authors with a complete network of 4,805,131 publications and 20,877,029 citations. For illustrative purposes we have also utilized a small subset of our dataset that contains 700 Computer Science researchers from Greek Universities.

Methods

Skyline operator

As discussed in the introductory section, given a set of attributes that characterize scientific performance, the *skyline operator* outputs the ones that cannot be surpassed by any other scientist in the dataset. We will explain how this works by presenting an example from Börzsönyi et al. (2001). Assuming that we want to book a relatively cheap hotel nearby the sea. Having the information of the cost and the distance from sea for each hotel we can produce two rank tables, one for each evaluation metric (cost, distance). It is difficult to produce a global rank table by combining the existing two. That is because we cannot define the relation between cost and distance. We cannot define for example that for 10 meters less distance we have to pay 10 euros more. Any try for defining such relation will be arbitrary. The skyline notion enables us to detect the best hotels (given the requirements) by combining the two metrics (or more metrics). The skyline consists of the set of hotels (generally the set of objects) that none of them is absolutely worst from any other one. A geometrical view is shown in Fig. 1. In this plot every point represents a hotel (an object). The coordinates of each object are defined by the score of the object for each metric. Each metric corresponds to one dimension. A two metrics rank can be presented with a 2D plot.

The definition of skyline set and a basic, efficient algorithm for its computation, as presented by Borzsonyi adding the mathematical notation.

Definition 1 (Dominance relationship) *Given two multidimensional points s1 and s2 with attributes (dimensions) from a space D, if s1 is equal to or better than s2 in all dimensions, and s1 is better than s2 in at least one attribute, we would say that s1 dominates s2 and write* $s1 \succ s2$. *That is:*

$$s1 \succ s2 : (\forall \alpha \in D, s_1.a \geq s_2.a) \wedge (\exists a \in D, s_1.a > s_2.a).$$

Definition 2 (Skyline set) *The skyline set comprises the set of points not dominated by any other point.*

The concept of skyline, calculated by the respective operator, has been utilized in the field of Computer Science for decades and dates back to the definition of the *Pareto frontier* in economics (Voorneveld 2003). However, the skyline set does not refer to

[2] http://dblp.uni-trier.de/.

efficient resource allocation; rather it provides a multi-criteria selection of distinguished scientists. The algorithm by Chomicki et al. (2003) known as Sort-First-Skyline (SFS) was employed in Sidiropoulos et al. (2016) to experimentally verify the identified elite groups of scientists and it will be utilized in the present analysis as well, due to its minimal computational cost and efficiency.

To appropriately identify the dimensions that will serve as attributes to the skyline calculation various experiments were performed in Sidiropoulos et al. (2016) and Stoupas (2017). The results of this analysis comply with literature in the sense that there exist groups (clusters) of highly correlated scientometric indicators and therefore the skyline does not vary significantly with different combinations of dimensions, as long as they are derived from the same group of indices. In Table 1 we present the Spearman's correlation coefficients between selected rank methods, based on various scientometric indicators. For the experiments conducted in the present work, we selected the *h*-index as a ranking method because it is the most commonly used performance indicator. We also selected the Perfectionism Index (Sidiropoulos et al. 2015) because it is the most dissimilar index with *h*-index based on Table 1. Finally, the *A*-index was included, since it is also dissimilar with *h*-index and offers a different counting of citations in the *h*-core.

Rainbow-ranking (RR-index)

The skyline operator selects the best performing scientific entities based on multiple criteria, but does not assign a meaningful and comparable ranking score to every scientist. The skyline operator, given a set of scientists, just extracts an elite set. Therefore *Rainbow-Ranking* is introduced to apply the skyline operator iteratively until all scientists of a dataset are classified into a skyline level. More specifically, given a set of scientists $A = X_1$, the first call of skyline produces the first skyline level. We denote this first set of scientists as set S_1. In the next step, we compute set $X_2 = X_1 - S_1$, which contains the scientists in the dataset that were not classified in the first skyline set S_1. For the set X_2 the skyline operator is applied once more and the result is the second skyline level (S_2). The process continues until all the scientists of the dataset are assigned a value that corresponds to the skyline level they have been ranked in. It is obvious that the set S_i is dominating over S_j ($S_i \prec S_j$) if $i < j$. Also, for researchers α and β it holds that $\alpha \prec \beta$ if $\alpha \in S_i$ and $\beta \in S_j$ and $S_i \prec S_j$.

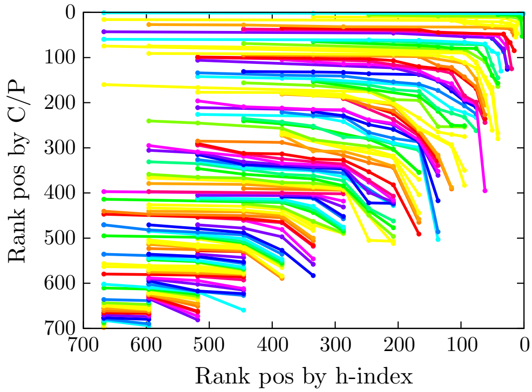
Figure 2 shows a graphical representation of the skyline levels with two dimensions (features): citations per publication and the *h*-index. We have selected the aforementioned subset of our original dataset containing 700 researchers from Greek Universities and have ranked them according to these two dimensions. Every point in Fig. 2 corresponds to a scientist. Each line connecting the points corresponds to a different skyline level. The *x*-axis represents ranking positions of each scientist according to their *h*-index, whereas on the *y*-axis the respective ranking positions according to citations per publication. Since this iterative procedure results into a plot with grouped curves as shown in Fig. 1, and the procedure is built over the notion of skyline, we have selected the name *Rainbow Ranking*. Also, in the above plot (and in all our experiments), in case of a tie, the average position of all tie members is used to express the rank position. This is the fairest method, as well as the method that is used in Spearman's coefficient computation. Note that *h*-index produces a lot of ties.

Inevitably, a score value should be produced for each rank level. If this score simply represent the level number, it would provide limited interpretability for the relative ranking

Table 1 Rank methods correlation based on Spearman's coefficient

	C	P	C/P	h	g	A	R	AR	h_2	hg	$m_{quotient}$	h_{cont}	h^t	h_{norm}	h_{tap}	h_{rat}	e	PI	mBornmann
P	0.398																		
C/P	0.776	− 0.216																	
h	0.904	0.451	0.649																
g	0.977	0.446	0.726	0.906															
A	0.897	0.233	0.790	**0.651**	0.869														
R	0.996	0.361	0.798	0.885	0.975	0.920													
AR	0.973	0.380	0.763	0.892	0.957	0.872	0.973												
h_2	0.923	0.379	0.719	0.927	0.916	0.762	0.920	0.912											
hg	0.963	0.457	0.703	0.976	0.975	0.774	0.951	0.946	0.944										
$m_{quotient}$	0.759	0.349	0.575	0.884	0.764	0.510	0.740	0.829	0.805	0.844									
h_{cont}	0.905	0.405	0.681	0.924	0.902	0.730	0.896	0.933	0.937	0.935	0.884								
h^t	0.911	0.422	0.675	0.950	0.909	0.714	0.899	0.927	0.947	0.953	0.884	0.970							
h_{norm}	0.323	0.652	0.791	0.338	0.274	0.300	0.348	0.332	0.362	0.315	0.358	0.328	0.330						
h_{tap}	0.899	0.546	0.575	0.963	0.900	0.647	0.865	0.875	0.892	0.954	0.840	0.898	0.922	0.212					
h_{rat}	0.909	0.445	0.657	*0.998*	0.910	0.662	0.891	0.897	0.934	0.977	0.882	0.930	0.956	0.342	0.962				
e	0.975	0.319	0.808	0.817	0.953	*0.959*	0.988	0.953	0.886	0.905	0.671	0.855	0.851	0.341	0.800	0.825			
PI	0.174	0.654	0.657	**0.031**	0.136	**0.324**	0.214	0.181	0.135	0.086	0.047	0.094	0.074	*0.753*	− 0.056	0.039	0.259		
mBornmann	0.888	0.269	0.758	0.775	0.864	0.839	0.899	0.872	0.887	0.841	0.653	0.846	0.834	0.377	0.753	0.785	0.901	0.240	
w	0.928	0.340	0.751	0.840	0.914	0.865	0.935	0.912	0.899	0.897	0.707	0.878	0.871	0.335	0.816	0.847	0.931	0.200	0.906

Fig. 2 Rainbow ranking graph



of each scientist compared to his peers; therefore, a normalization of this value is required. To summarize the ranking levels into a single number metric, given a set of scientists A and a set of dimensions *dims*, we define the RR-index of a scientist a based on *dims* as follows:

$$RR(a, dims) = 100 - 100 * \left(\frac{|A_{\text{above}}(a, dims)|}{|A|} + \frac{|A_{\text{tie}}(a, dims)|}{2 * |A|} \right) \tag{1}$$

In Eq. (1) $|A|$ is the total number of scientists in our dataset, $|A_{\text{above}}(a, dims)|$ is the number of scientists ranked at higher skyline levels than scientist a based on dimensions *dims*. Note that level 1 is considered higher than level 2 in a rank table. Additionally, $|A_{\text{tie}}(a, dims)|$ is the number of scientists who are ranked at the same level with scientist a, excluding scientist a. Consequently, the following holds for the *RR*-index:

$$0 < RR(a, dims) \leq 100 \tag{2}$$

The case when $RR(a, dims) = 100$ means that scientist α is ranked in the first skyline level alone. Since all the members of a skyline level should be assigned with the same score, we have chosen to assign a score analogous to the average rank position of all tie members normalized to the range 0–100.

Also, it is obvious that the following condition holds:

$$\sum_{\forall a \in A} RR(a, dims) = |A| * (|A| - 1)/2 \tag{3}$$

The key components for the calculation of the *RR*-index are the skyline dimensions and the ranking positions assigned according to each dimension. By selecting different bibliometric indices as skyline dimensions, the calculated *RR*-index can be fully customizable. However, since bibliometric indices are highly correlated with each other, as depicted in Table 1, selecting highly correlated indices would yield analogous results in the final skyline ranking. As the number of dimensions (criteria) increases the skyline's sizes increases as well because the number of individuals that are not clearly bypassed by others increases.

In the next section we present our experiments with multiple dimension combinations and investigate the distinguishing power of the RR-index.

Results

The computation of RR-index in our dataset with the dimensions *h*-index, *A* and *PI*, i.e., *RR(h, A, PI),* produces a classification into $\sim$ 700 levels of skylines. In Fig. 3 the size of each skyline level is illustrated. The plot follows a relatively stable distribution with about 200 authors per skyline level and a peak at level 500. The first 20 skyline levels are significantly smaller in size (less than 50 entities per skyline level); the same applies for the last sky-line levels (higher than 600). This observation indicates that when a unified relative ranking is produced though comparison with one's peers the "Matthew effect" in scientific output is mitigated in the sense that few scientists are ranked at the top, but also few of them are assigned at the lowest ranks. The majority of them are dispersed across the moderate skyline levels, which better explains the actual distribution of performance (Brzezinski 2015). Please, note that Table 2 gives abbreviations for the tested RR-indices.

In a similar manner, Fig. 4 shows the distribution of skyline sizes for *RR(h, A, PI,* h^t*).* In this computation we added the dimension of trend *h*-index (h^t). The trend *h*-index is clustered very close to the original *h*-index. We argue that this happened due to the fact that Computer Science, the domain selected, is a relatively new and fast paced field, where not enough "sleeping beauties" (Van Raan 2004). Sidiropoulos et al. (2007) are present or even when they do exist, they do not "sleep" for long periods of time. Therefore, the addition of the trend index will help assign merit to scientists publishing in up and coming fields of research. With one added dimension, we see that the levels of skylines decreased and sky-line sizes increased. However, in the total ranking, the similarity of *RR(d1)* and *RR(d3)* is still very high (0.997 Spearman, 0.957 Kendall tau), as shown later in Tables 3 and 4. This means that, although we added a new dimension, the number of levels and the ties changed but the ranking remains almost the same.

In Fig. 5 we doubled the dimensions used in *RR(d1).* For each one of the three dimensions of *RR(h, A, PI)* an additional one was introduced that displays a high correlation with one of the original dimensions based on Table 1. The rationalized h_{rat}-index was selected due to its high similarity to *h*-index, *e*-index as it is similar to *A*, whereas h_{norm} was included because it is correlated with *PI.* In this point, it must be noticed that h_{norm} is not strongly correlated with *PI* as the correlation value is less than 0.9 but h_{norm} is the closer index to *PI*. As a result, the skyline levels decreased and the skyline sizes increased. We see exactly the same behavior in Fig. 6. As expected, the more dimensions we include the more members are placed in each skyline level. We observe a denser ranking in

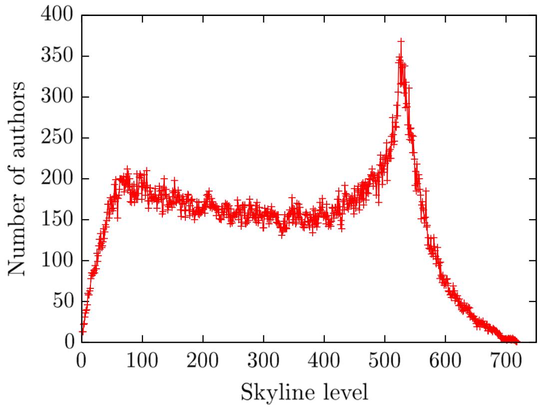


Fig. 3 Skylines' sizes for *RR(h, A, PI)* or *RR(d1)*

Table 2 Abbreviations for Rainbow Rankings used in this publication

Abbreviation	Abbreviation for
RR(d1)	*RR(h-index, A-index, PI)*
RR(d2)	*RR(h-index,* h_{rat}*, A-index, e-index, PI,* h_{norm}*)*
RR(d3)	*RR(h-index, A-index, PI,* h^t*)*
RR(d4)	*RR(h-index,* h_{rat}*, A-index, e-index, PI,* h_{norm}*,* h^t*)*

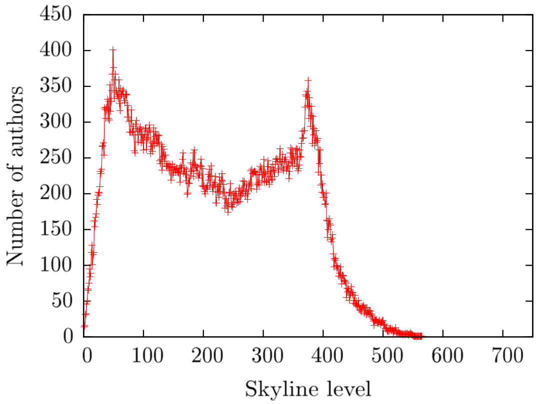


Fig. 4 Skylines' sizes for *RR(h, A, PI,* h^t*)* or *RR(d3)*

Table 3 Spearman coefficient of RRs with their generators

	h	A	h^t	h_{norm}	h_{rat}	e	PI	*RR(d1)*	*RR(d2)*	*RR(d3)*
h										
A	0.651									
h^t	0.950	0.714								
h_{norm}	0.338	0.300	0.330							
h_{rat}	0.998	0.662	0.956	0.342						
e	0.817	0.959	0.851	0.341	0.825					
PI	0.031	0.324	0.074	0.753	0.039	0.259				
RR(d1)	**0.806**	**0.935**	0.831	0.400	0.813	0.963	**0.307**			
RR(d2)	**0.818**	**0.918**	0.835	**0.476**	**0.825**	**0.954**	**0.362**	0.991		
RR(d3)	**0.836**	**0.925**	**0.862**	0.399	0.843	0.966	**0.287**	0.997	0.990	
RR(d4)	**0.830**	**0.913**	**0.851**	**0.491**	**0.837**	**0.955**	**0.367**	0.984	0.997	0.988

Bold values indicate the corresponding generators

Table 4 Kendall Tau coefficient

	RR(d1)	*RR(d2)*	*RR(d3)*
RR(d2)	0.920		
RR(d3)	0.957	0.920	
RR(d4)	0.894	0.961	0.912

Fig. 5 Skylines' sizes for *RR(h,* h_{rat}, *A*, *e*, *PI*, h_{norm}) or *RR(d2)*

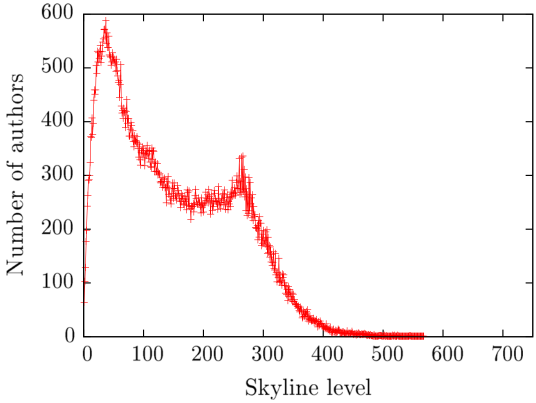
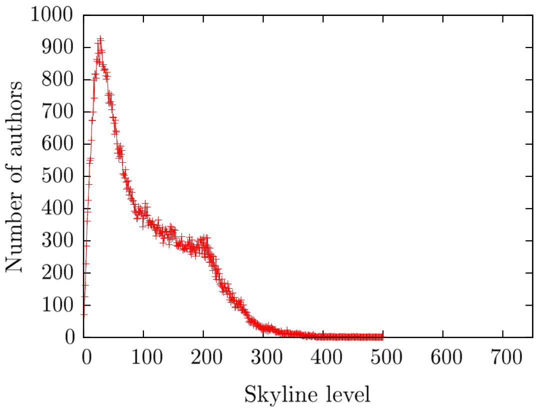


Fig. 6 Skylines' sizes for *RR(h,* h_{rat}, *A, e, PI,* h_{norm}, h^t*)* or *RR(d4)*



moderate skyline levels, meaning that with the addition of correlated dimensions the segmentation in performance levels is less detailed due to the similarity in rankings produced by correlated indices. As depicted at level 40, there is a peak at 900 skyline members. In Tables 3 and 4 the correlation values indicate that all four variations of RR-index produce similar overall ranking, even though the respective sizes of individual skyline levels may differ. This means that if representative rank methods are selected, then there is no need for a large number of them to be used in producing a unified representative ranking. Additionally, when less correlated dimensions are selected the resulting segmentation becomes more detailed and precise.

One can argue that ranking into levels produces a lot of ties. This claim is not true. As it is shown in Fig. 7, *h*-index, which is the most commonly used method, produces much fewer levels and much greater number of ties. The peak of *h*-index cardinality is more than 12,000 while in RR the worst case (Fig. 6) is less than 1000.

Finally, Table 3 illustrates that all variations of *RR* are not necessarily similar with their generators. For example, although *PI* is one of the generating dimensions of *RR(d1),* the coefficient between them is only 0.307. This result can be attributed to the interconnected nature of bibliometric indicators, meaning that the resulting relative ranking of a scientist may significantly differ from her/his individual ranking on each one of the generator dimensions. However, the information conveyed by the RR-index provides the relative ranking as a result of overall performance, including different aspects of scientific output as expressed by the generator dimensions. Consequently, the RR ranking can be considered a more unified and representative evaluation metric, compared to its individual generator dimensions.

Table 5 illustrates the Rank Table according to *RR(d1)*. We have included columns *h*, *A* and *PI* which are the generators of *RR(d1)* as well as the values for *C* (number of citations), *P* (number of publications) and *C/P* (citations per publication). The last column shows the skyline level to which each respective scientist has been assigned.

The first 13 scientists were ranked in the first skyline level and they are assigned the same value for *RR(d1)*. In this list of top ranked researchers, we encounter scientists who can be grouped into two subsets based on their work; one group is comprised of those who have worked in core computer science (e.g., networking, compilers, databases), and the second group of those who have contributed to the field of computational methods for biology, with the latter group being the largest of the two. More specifically, in the former group, we see Scott Shenker well known for his contributions in networking theory and practice, Ian Foster of the community of high performance computing (grid and cloud computing), and Jeffrey Ullman whose work spans across several research areas (compilers, programming languages, databases). In the latter group, David Haussler was a member of the team who sequenced the human genome, Robert Tibshirani made solid contributions to statistical learning theory and its application to biological problems, Altschul Stephen, David Lipman and Web Miller, co-developers of the well-known BLAST family of tools for sequence comparison. Additionally, there is Higgins Desmond, developer of the Crystal-W and Crystal-X tools for multiple sequence alignment, Gish Warren co-worker of David Haussler to human genome sequencing, and of Stephen Altschul to the development of BLAST.

In the second skyline level, we mainly encounter core computer scientists, namely Hector Garcia-Molina of databases, Deborah Estrin of embedded systems (sensors), David Culler of networking, Simon Herbert of political science and economics, Ronald Rivest of crypto-graphy and co-developer of the RSA cryptosystem, Vladimir Vapnik the father of statistical learning theory, Thomas Cormen well-known for his work on distributed algorithms, Claude Shannon the father of information theory, and so on. Finally, Eugene Myers a computational biologist who co-authored the famous Science-Nature paper on the sequence of the human genome is also ranked at the same skyline level. The reason why there is a large number of computational biologists/bioinformatics researchers in the top

Fig. 7 *h*-index levels' sizes

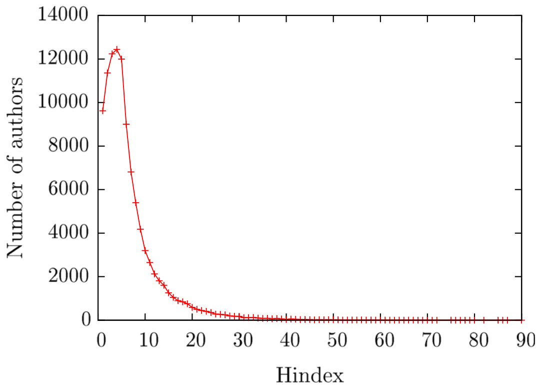


Table 5 First two skyline levels' members

Name	*C*	*P*	*C*/*P*	*h*-index	*A-index*	*PI*	*RR*(*d1*)	#Skyline
Shenker Scott	38,557	473	81.52	90	361.02	12,187	99.99	1
Foster Ian	39,052	730	53.50	87	365.48	− 9320	99.99	1
Ullman Jeffrey	38,019	445	85.44	82	394.98	14,977	99.99	1
Haussler David	27,799	320	86.87	78	305.29	15,007	99.99	1
Tibshirani Robert	47,661	344	138.55	69	636.06	33,447	99.99	1
Miller Webb	54,262	532	102.00	42	1272.76	35,446	99.99	1
Higgins Desmond	41,527	190	218.56	21	1974.43	38,419	99.99	1
Lipman David	48,638	97	501.42	20	2425.05	47,498	99.99	1
Altschul Stephen	46,730	78	599.10	19	2453.42	45,970	99.99	1
Gish Warren	26,065	33	789.85	9	2894.11	25,930	99.99	1
Thompson Julie	36,441	450	80.98	8	4552.50	32,969	99.99	1
Gibson Toby	36,329	427	85.08	8	4538.63	33,041	99.99	1
Zhang Jinghui	28,638	94	304.66	5	5727.20	28,218	99.99	1
Garcia-molina Hector	25,743	578	44.54	86	205.88	− 9173	99.98	2
Estrin Deborah	34,706	446	77.82	85	344.86	11,246	99.98	2
Culler David	27,360	363	75.37	77	296.17	11,267	99.98	2
Simon Herbert	31,620	1194	26.48	75	389.40	− 46,680	99.98	2
Lander Eric	42,201	430	98.14	67	612.10	22,369	99.98	2
Rivest Ronald	38,336	294	130.39	58	615.40	28,012	99.98	2
Vapnik Vladimir	31,324	123	254.67	49	618.14	30,099	99.98	2
Leiserson Charles	23,147	155	149.34	36	627.36	20,159	99.98	2
Myers Eugene	32,210	286	112.62	33	954.42	24,950	99.98	2
Cormen Thomas	16,707	50	334.14	14	1189.57	16,399	99.98	2
Shannon Claude	13,554	32	423.56	7	1935.57	13,428	99.98	2
Woods Richard	11,642	41	283.95	6	1940.33	11,468	99.98	2
Schaffer Alejandro	24,096	42	573.71	5	4818.20	23,936	99.98	2

ranked group could be the recent popularity and fast growth of the field, as compared to the more "old-fashioned" domains of core computer science. It can be observed that even though this group may have scored relatively lower values in *h*-index, the values of *A* and *C*/*P* are significantly higher, meaning that these bioinformatics researchers accumulate citations at a faster pace compared to the more mature researchers of the core computer science group. As a result, different publishing patterns can be identified and rewarded using the RR-index.

Discussion and conclusions

During the last decade, mainly due to the development of open access online databases, which maintain large scale publication records of individual scientists and their citations, the introduction of new scientometric indicators comes as a storm; a number of them are highly correlated with each other, but there exist several indices that capture independent aspects of the publishing behaviour and/or impact of a scientist's output. Despite the fact that we would wish to have a single numeric metric to tell us everything about a scientist's publishing patterns and their impact, this can be a real challenge that remains to be addressed.

In this article we follow a different path compared to earlier relevant works, and investigate the following scenario: Given a set of indicators, selected in any algorithmic way (e.g., by clustering, by administrative decisions, etc.), can we successively rank scientists into layers based on the given indicators, such that the scientists in each layer outperform those of the lower layers according to (at least one) indicator?

To address this problem, we employed the notion of skyline introduced in Sidiropoulos et al. (2016). We iteratively apply this method, finding successive skylines, which produces a grouping of scientists into layers. We call this ranking scheme the *Rainbow Ranking*. The merits of Rainbow Ranking are the following:

(a) it works for any set of dimensions (i.e. scientometric indicators), and thus it relieves the evaluator from the burden of selecting just one indicator to perform the ranking; however, at the extreme case of a singleton set of dimensions, then the Rainbow Ranking reduces to the ranking imposed by that indicator,
(b) it is not correlated to existing schemes,
(c) it allows for multiple ties (i.e., when scientists outperform each other in different aspects of their evaluated work),
(d) it proves to be a practical ranking scheme that can be scaled up to thousands of entities by inheriting the decreased computational complexity of skyline calculation, and finally
(e) It is a general index that can be used in the classification of scientific journals, conferences, researchers, universities and of any other scientific entity.

In the future, we plan to investigate the ranking stability of Rainbow Ranking, i.e., how fast and to what extent the contents of the layers change over time. We further plan to incorporate the element of fuzziness to Rainbow Ranking thus allowing for more flexible rankings or even overlapping rankings to occur.

References

Alguliev, R. M., Aliguliyev, R. M., Fataliyev, T. K., & Hasanova, R. S. (2014). Weighted consensus index for assessment of the scientific performance of researchers. *Collnet Journal of Scientometrics and Information Management, 8*(2), 371–400.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2010). *hg*-index: A new index to characterize the scientific output of researchers based on *h*- and *g*-index. *Scientometrics, 82,* 391–400.

Anderson, T. R., Hankin, R. K., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the *h*-index to score total publication output. *Scientometrics, 76*(3), 577–588.

Bollen, J., van de Sompel, H., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE, 4*(6), e6022.

Bornmann, L., Mutz, R., & Daniel, H. D. (2008). Are there better indices for evaluation purposes than the *h* index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology, 59*(5), 830–837.

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. P. (2014). *h*-index research in scientometrics: A summary. *Journal of Informetrics, 8*(3), 749–750.

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. (2011). A multilevel meta-analysis of studies reporting correlations between the *h*-index and 37 different *h*-index variants. *Journal of Informetrics, 5*(3), 346–359.

Börzsönyi, S., Kossmann, D., & Stocker, K. (2001). The skyline operator. In *Proceedings of the IEEE international conference on data engineering* (*ICDE*) (pp. 421–430).
Brzezinski, M. (2015). Power laws in citation distributions: Evidence from scopus. *Scientometrics*, *103*(1), 213–228.
Chomicki, J., Godfrey, P., Gryz, J., & Liang, D. (2003). Skyline with presorting. In *Proceedings of the IEEE international conference on data engineering* (*ICDE*) (pp. 717–719).
Egghe, L. (2006). Theory and practice of *g*-index. *Scientometrics, 69*(1), 131–152.
Franceschini, F. (2014). The citer-success-index: A citer-based indicator to select a subset of elite papers. *Scientometrics, 101*(2), 963–983.
Glanzel, W., Thijs, B., & Debackere, K. (2014). The application of citation-based performance classes to the disciplinary and multidisciplinary assessment in national comparison and institutional research assessment. *Scientometrics, 101*(2), 939–952.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569–16572.
Hutchins, B., & Yuan, X. A. (2016). Relative citation ratio (RCR): A new metric that uses citation rates to measure influence at the article level. *PLoS Biology, 14*(9), e1002541.
Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *r*- and *ar*-indices: Complementing the *h*-index. *Chinese Science Bulletin, 52*(6), 855–863.
Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newsletter, 2*(3), 4–6.
Ruane, F., & Tol, R. (2008). Rational (successive) *h*-indices: An application to economics in the Republic of Ireland. *Scientometrics, 75*(2), 395–405.
Schreiber, M., Malesios, C., & Psarakis, S. (2012). Exploratory factor analysis for the hirsch index, 17 *h*-typevariants, and some traditional bibliometric indicators. *Journal of Informetrics, 6*(3), 347–358.
Sidiropoulos, A., Gogoglou, A., Katsaros, D., & Manolopoulos, Y. (2016). Gazing at the skyline for star scientists. *Journal of Informetrics, 10*(3), 789–813.
Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch *h*-index for disclosing latent facts in citation networks. *Scientometrics, 72*(2), 253–280.
Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2015). Ranking and identifying influential scientists versus mass producers by the Perfectionism Index. *Scientometrics, 103*(1), 1–31.
Stoupas, G. (2017). Citation graph analysis and application of prototype evaluation and ranking methods. Master Thesis. Alexander Technological Institute of Thessaloniki.
Tamar, V. L., & Tim, S. E. (2015). Ranking journals using altmetrics. *CoRR*. abs/1507.00451.
Tsai, C. F. (2014). Citation impact analysis of top ranked computer science journals and their rankings. *Journal of Informetrics, 8*(2), 318–328.
Van Raan, A. F. J. (2004). Sleeping beauties in science. *Scientometrics, 59*(3), 467–472.
Voorneveld, M. (2003). Characterization of Pareto dominance. *Operations Research Letters, 31*(1), 7–11.
Wildgaard, L., Schneider, J. W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics, 101*(1), 125–158.
Wohlin, C. (2009). A new index for the citation curve of researchers. *Scientometrics, 81*(2), 521–533.
Wolcott, H. N., Fouch, M. J., Hsu, E., Bernaciak, C., Corrigan, J., & Williams, D. (2015). Modeling time-dependent and -indpendent indicators to facilitate identification of breakthrough research papers. In *Proceedings of the international society of scientometrics & informetrics conference* (*ISSI*) (pp. 403–408).
Zhang, C. (2009). The *e*-index, complementingthe *h*-index for excess citations. *PLoS ONE, 4*(5), e5429.