COLLNET Journal of Scientometrics and Information Management

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsim20

When universities rise (Rank) high into the skyline

Georgios Stoupas, Antonis Sidiropoulos, Dimitrios Katsaros & Yannis Manolopoulos

To cite this article: Georgios Stoupas, Antonis Sidiropoulos, Dimitrios Katsaros & Yannis Manolopoulos (2021) When universities rise (Rank) high into the skyline, COLLNET Journal of Scientometrics and Information Management, 15:2, 241-258, DOI: 10.1080/09737766.2021.1955419

To link to this article: https://doi.org/10.1080/09737766.2021.1955419

Published online: 22 Nov 2021.

Submit your article to this journal

Article views: 48

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tsim20

COLLNET Journal of Scientometrics and Information Management
ISSN : 0973 – 7766 (Print) 2168 – 930X (Online)
Vol. 15(2) December 2021, pp. 241–258
DOI : 10.1080/09737766.2021.1955419

Taylor & Francis
Taylor & Francis Group

When universities rise (Rank) high into the skyline

Georgios Stoupas
Antonis Sidiropoulos
Dimitrios Katsaros
Yannis Manolopoulos

Georgios Stoupas
Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki
Greece

Antonis Sidiropoulos
Department of Information and Electronic Engineering
International Hellenic University
Thessaloniki
Greece

Dimitrios Katsaros *
Department of Electrical & Computer Engineering
University of Thessaly
Volos
Greece

* dkatsar@e-ce.uth.gr

Yannis Manolopoulos
Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki
Greece

and

Pure and Applied Sciences
Open University of Cyprus
Nicosia
Cyprus

* (Corresponding Author)

The quality of the education provided and the research impact produced by universities is continuously evaluated at national and international level. This phenomenon is not new. However, nowadays education is not only considered as a social value and right/privilege, but also as a big economic sector, which addresses to large portions of population worldwide. In this ecosystem, university rankings play a crucial role since they provide filtered information which is reproduced in surveys, newspapers, social media etc. All university rankings are based on a set of ad hoc evaluation criteria. Moreover, the final score is based on a set of arbitrary weights summing up to 1. Thus, at the end, these university rankings differ significantly producing ambiguities and doubts. In this paper, we propose a novel university ranking method based on the Skyline operator, which is used on multi-dimensional objects to extract the non-dominated (i.e., "prevailing") ones. Our method is characterized by several advantages, such as: it is transparent, reproducible, without any arbitrarily selected parameters, based on the research output of universities only and not on publicly not traceable or questionnaires. Our method does not provide absolute rankings, but rather it ranks universities categorized in equivalence classes. Thus, we develop a generic framework which can be used for ranking universities and departments, and even individual persons. For the proof of concept we apply the framework in our Greek academic space, providing a case study on ranking persons and departments on computer science and engineering using data extracted from Microsoft Academic.

Keywords: *University ranking, h-index, Rainbow ranking, Skyline, Greek computer science departments.*

© Taru Publications

1. Introduction

For modern societies, education is a value per se, since individuals can improve their social, cultural and economic status through education, at university level in particular. International and national organizations monitor the activities of universities, colleges and research institutions to create and publicize rankings of these academic institutions. Although evaluation organizations exist for decades, their number has increased dramatically nowadays[1]. Apart from such organizations, many academic efforts related to university ranking are flourishing. They draw their inspiration from traditional methodologies such as multicriteria sorting, rank fusion, Pareto frontier and the related Data Envelopment Analysis [10] (see Section 2.1 for definitions of Pareto frontier and DEA). For instance, recent academic efforts on the topic include a goal programming model [9], application of the Pareto frontier [16], application of the Data Envelopment Analysis [6, 25], clustering [14, 20], and merging (fusing) several ranked lists [12].

The ranking lists produced by evaluation organizations are of major importance for decision making by prospective students and their families, by academic staff in search of employment, and by funding agencies. As the general public is interested in university rankings, the same way universities are interested as well. The placement of universities in these lists is a crucial factor that can shape their future. The increasing competition among academic institutions has led many universities to adapt their strategy according to the particular criteria of each evaluation system. The reason for this inclination can be understood by considering the "Thomas theorem" from sociology, which states "if men define situations as real, they are real in their consequences" [22]. As mentioned in [3]: "if rank positions between two universities define performance differences as real, they are real in their consequences (although the university ranking shows only slight differences between the universities' scores)".

Some of the evaluation organizations are private enterprises, whereas others stem from university research centers or by even national research institutions. Probably, the three most prestigious global rankings are: ARWU, QS and THE. Another quite well-known ranking list is Webometrics, which is published by the Spanish National Research Council. In addition, we mention the so called NTU (National Taiwan University) ranking which was founded by the Higher Education Evaluation and Accreditation Council of Taiwan. Finally, there are very few ranking lists originated from universities. Notably, we mention the lists of: École Nationale Supérieure des Mines de Paris, Leiden University, Middle East Technical University, Wuhan University, and Shanghai Jiao Tong University, which founded the ARWU organization.

It is a fundamental practice of all the evaluation organizations to base their outcomes on a set of indicators, which differ from one university ranking to the other. From the long list of ranking organizations, here we focus on those which are mentioned more often in

[1] https://en.wikipedia.org/wiki/College_and_university_rankings

the literature and the newspapers. Thus, in the sequel we are going to introduce the indicators used by ARWU, QS, THE, and Webometrics.

- ARWU (Academic Ranking of World Universities)[2] uses six indicators to evaluate universities, such as: (a) alumni who have won a Nobel or Field Medal, (b) staff who have won a Nobel or Field Medal, (c) highly cited researchers, (d) papers published in Nature or Science, (e) papers indexed in SCI and SSCI, and (f) per capita academic performance. For the above criteria, the weights are 10%, 20%, 20% 20%, 20% and 10%, respectively.
- **QS (Quacquarelli Symonds)[3]** uses six key pillars: (a) academic peer research based on an internal global academic survey, (b) faculty/student ratio to measure teaching commitment, (c) citations/faculty to measure research impact, (d) employment based on graduate employers, (e) international student ratio to measure the diversity of the student community, and (f) international staff ratio to measure the diversity of the academic staff. For the above criteria, the weights are 40%, 20%, 20%, 10%, 5% and 5%, respectively.
- **THE (Times Higher Education World University Rankings)[4]** uses 13 performance indicators, which are grouped into five categories: (a) industry income and innovation, (b) international diversity, (c) teaching and learning environment (d) volume, revenue and reputation, and (e) citations and research influence. For these five criteria, the weights are 2.5%, 5%, 30%, 30%, and 32,5%, respectively.
- **Webometrics Ranking of World Universities[5]** investigates the online presence of universities. Four markers are used: (a) presence measured by the size of the main web domain, (b) visibility measured by the external networks linking to the institution's webpages, (c) transparency measured by number of citations of the top researchers, and (d) excellence measured by the number of papers amongst the top 10% most each in each of 26 disciplines. For these five criteria, the weights are 5%, 50%, 10%, and 35%, respectively.
- **NTU ranking[6]** which is based on eight features categorized into three categories, namely: (a) research productivity, comprised by two features, i.e. number of articles in the last 11 years and number of articles in the current year, (b) research impact comprised by the citations in the last 11 years, the number of citations in the last 2 years, and the average number of citations in the last 11 years, and (c) research excellence comprised by the *h*-index of the last 2 years, the number of highly cited papers, and the number of articles in the current year in high-impact journals. For these eight features, the weights are 10%, 15%, 15%, 10%, 10%, 10%, 15% and 15%, respectively.

2 https://en.wikipedia.org/wiki/Academic_Ranking_of_World_Universities
3 https://en.wikipedia.org/wiki/QS_World_University_Rankings
4 https://en.wikipedia.org/wiki/Times_Higher_Education_World_University_Rankings
5 https://en.wikipedia.org/wiki/Webometrics_Ranking_of_World_Universities
6 http://nturanking.csti.tw

By looking closer to the set of indicators of each evaluating organization and the particular weights, one can draw several conclusions:

- The indicators of each ranking list are selected somehow arbitrarily. The intersection of indicators' sets between any two ranking systems is small. For example, when comparing ARWU and Webometrics, which have 6 and 4 indicators respectively, only the indicators about research production (papers) and research impact (citations) are common, although not defined exactly the same way.
- The weight of each is indicator is selected arbitrarily. For example, the teaching component is weighted by 30% at THE, 20% for QS, and 0% for ARWU and Webometrics.
- The results are not reproducible in all cases. For example, 50% of the final score of QS is based on questionnaires.
- The weight values are ranging from 2.5% up to 50%. Thus, it is questionable if there is balance between the assorted criteria. For example, according to THE, the minimum weight is 2.5 and the maximum is 30%.
- Intuitively it is understood that several indicators are correlated. For example, consider the cases e and f of QS (the international student ratio and the international staff ratio), and the cases c and d of Webometrics (the number of citations and the number of papers).
- Finally, it is remarked that among the indicators of the above rankings, it is only the production and impact of publications which are common, although defined differently.

Although university rankings have gained popularity, they have been heavily criticized for many reasons as summarized in [13]. To name a few [1]:

- they are not statistically robust;
- they are not stable but show inconsistent fluctuations from year to year;
- they favor universities of English-speaking countries focusing in hard science;
- they tweak their results to show movement and attract commercial interest;
- they compare apples-to-oranges (e.g., teaching versus research institutions).

Here, we propose a different approach to rank academic institutions. In general, evaluation criteria can be divided in two categories: academic (teaching and research performance) and non-academic ones (reputation, diversity, facilities etc). As academic issues, and research in particular, is a well understood field of evaluation, and not easily manipulated with questionnaires, we focus on this field along the lines of the CWTS ranking of Leiden University[7]. The technical problems which appear when universities are ranked by

7 https://en.wikipedia.org/wiki/CWTS_Leiden_Ranking

bibliometric methods have been analyzed in [24], a paper which has been characterized as disruptive [3]. In any case, the spirit of our method can be extended to any set of indicators.

It has been argued that university rankings can lead to erroneous conclusions. The top-15 or so universities are indeed a separate class of high-quality institutions. However, below this line the differences are not significant even for universities which are separated by many places [1, 8]. Similar critical remarks have been stated in [11, 15, 27]. Notably, it is mentioned that our method solves the particular problem of creating false impressions about the differences between the various universities [2].

The contribution of our method and its advantages over the popular university rankings are the following:

- It focuses on the research output of the universities, which is a mature field to take into consideration during an evaluation and does not rely on questionnaires.
- It uses a set of indicators well-known to the whole academic community from Google Scholar metrics, which are found at the personal accounts of individual researchers; so it supports reproducibility.
- It does not use arbitrary weights for each indicator to avoid tuning which can result in different outcomes; on the contrary, it treats all indicators equally in a symmetric manner.
- It avoids the absolute rankings, where there is no serious meaning in claiming that the i-th university is better than the $(i+1)$-th one.
- It provides a list with a single structure, contrary to the popular rankings, where paradoxically the first few hundreds of universities are ranked in absolute order, whereas the rest follow in groups. For example, QS ranks the first 500 universities in sequence, and then ranks the next 100 in groups of 10, the next 200 in groups of 50, and finally the next 200 in a single group.
- It is not prone to inconsistent fluctuations from year to year, which lead to inaccuracies and instability.

Thus, the focus of the present paper is the ranking of academic units, e.g., universities, departments, and towards such goal the article develops a generic framework which possesses the aforementioned advantages. To show the strengths and versatility of the framework we present a case study on the Greek academic space, and in particular on the Greek departments of computer science and engineering, and on the respective faculty members.

The structure of the remaining part of this paper is as follows. Section 2 explains the Skyline operator, which is used to extract the set of multidimensional objects that dominate all other objects of a dataset. Based on the Skyline operator is the technique of Rainbow Ranking which has been proposed in the past to cluster researchers in groups according to their production and impact. Section 3 presents the derivation of the dataset, and its cleaning along with the results of the application of the Rainbow Ranking method on the particular dataset of Greek computer science departments. In Section 4 we apply our methodology to a set of international universities. Finally, Section 5 concludes the paper.

Hotel	Price	Distance
h1	80	4
h2	140	4
h3	110	3
h4	300	2.7
h5	100	2.2
h6	140	2
h7	210	1.7
h8	200	1.1
h9	250	1.05
h10	340	1

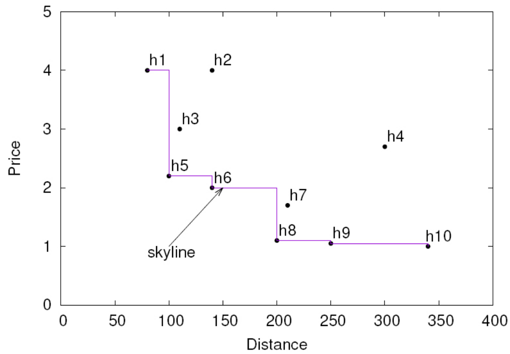

Figure 1

Skyline calculation for the hotels example.

2. Skyline and Rainbow Ranking

2.1 *The Skyline operator*

The Skyline operator is used to satisfy a database query by filtering results and keep only those objects that are not worse than any other, i.e., they are not dominated [5]. To understand the notion of Skyline, let's examine a particular example. Suppose that we want to spend our vacations at a beach hotel. The criteria to choose such a hotel is price and distance from the sea. Having this information for each available hotel we can produce two rank tables, one for each evaluation criterion, i.e. cost and distance. However, it is difficult to produce a global rank table by combining the two partial ones because we cannot define the relation between the two. For instance, is it worth to pay 10 euros for a hotel which is 100 meters closer to the sea? The Skyline operator detects the best hotels by combining these two (or more) criteria. The final Skyline set consists of the non-dominated hotels such that none of them is absolutely worst from any other.

Figure 1 illustrates a concrete example. The left part gives a dataset of 10 hotels with values for the two criteria in question. The right part depicts a 2-d plot, where axis x represents distance, axis y represents price, whereas every point represents a hotel. Apparently, the hotels $h1$, $h5$, $h6$, $h8$, $h9$ and $h10$ are members of the Skyline set, because any one of them is not worse than all other Skyline members in one only criterion. At the same time, these six hotels are no worse than the remaining four hotels ($h2$, $h3$, $h4$, $h7$) with respect to both criteria. In Skyline terminology, these four hotels are dominated.

The generic concept of 'skyline' set as a group of points not 'fully dominated' by any other point dates back several decades. It is similar to the term Pareto frontier in the economics studies, which is used to describe a set of options which are 'Pareto efficient'. Pareto efficiency refers to a set of resource allocations to a population where it is not possible to improve the allocation to one individual without hurting someone else. Therefore, the notions of Pareto frontier and skyline set are similar, but our application area is not related to any resource allocation problem. Nevertheless, the concept of (weak/strong) Pareto dominance [26] for options with multiple features as used in multicriteria optimization and decision making is alike to skyline dominance.

The technique of Data Envelopment Analysis [18] is also based on Pareto efficiency; it focuses on relative performance based on the equilibrium between resources (personnel, funding,etc.) and resulting productivity (number of published papers, citations acquired, etc.) and so on.

In the literature we can find several skyline computation algorithms [23], since this operator is used broadly in many applications. Elaborating further on algorithms and applications is out of the scope of this paper. However, it is vital to mention a useful application of Skylines in scientometrics. Sidiropoulos et al. [19] have harvested data from Microsoft Academic Search (MAS) and established a cleaned dataset with researchers' portfolios. Further, they tested several sets of three non-correlated bibliometric indices and produced a 3-d Skyline sets of 'dominating' researchers for each year of the period 1992-2013. The above study reveals that the Skyline operator may be used to assess scientific excellence, for grant allocation, for hiring and promoting academic staff.

2.2 *Rainbow Ranking*

According to the previous, the Skyline operator can extract the 'top' (dominating) researchers based on multiple bibliometric criteria. However, it can not assign a meaningful and comparable ranking to all researchers. To this end, the method of Rainbow Ranking was introduced in [21], where the Skyline operator is applied iteratively until all scientists of a dataset have been classified into a Skyline level. More specifically, given a set of scientists $X1$, the first call of the Skyline operator produces the first Skyline level, which we denote as $S1$. Next, we apply the Skyline operator on the dataset $X1 - SI$, to derive the second Skyline layer, denoted as $S2$. This process continues until all the scientists of the dataset have been assigned to a particular Skyline level Si.

The notion of Ranking will be better explained with the following example. The authors in [21] have selected the bibliographic portfolios of 539 academic staff from 19 CS faculties from 17 Greek universities[8]. The Rainbow Ranking has been applied on this dataset, that is an iterative Skyline operator on diminishing datasets by adopting two bibliographic indices: citations/publication and h-index; thus, Figure 2 has been produced. Axis x depicts the position of a scientist in the range [1..539] according to her/his h-index value; axis y depicts their position in the same range according to the ratio c/p, where c is the number of citations and p is the number of publications. Every point corresponds to a scientist and each line connects the points of the same Skyline level. Since this iterative procedure results into a plot with grouped curves, the method has been called 'Rainbow Ranking'.

8 By 'CS faculties' we denote all CS, CEng, ECE 'Departments' and 'Schools' according to the terminology used in Greek universities, offering four and five years studies.

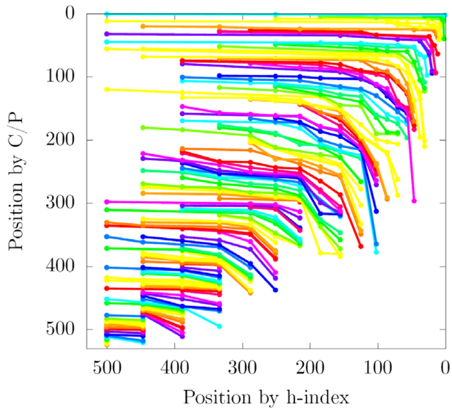

Figure 2

Rainbow Ranking plot for scientists.

To give more semantics to the method, a particular value should characterize the Skyline levels. Should this value be the iteration number, then this would convey limited interpretability since the relativeness would be lost. It is crucial to designate the position of scientist among their peers. Therefore, a normalization of this value is necessary. Thus, for a research a, her/his RR-index is defined as:

$$RR(a) = 100 - 100 \times \frac{|A_{above}(a)| + |A_{tie}(a)|/2}{|A|} \tag{1}$$

where A is the set of scientists, $A_{above}(a)$ is the number of scientists at higher Skyline levels than scientist a, and $A_{tie}(a)$ is the number of scientists at the same Skyline level with scientist a, excluding scientist a. Apparently, it holds that: $0 < RR(a) \leq 100$.

A key component for the RR-index concept is the number and identity of the Skyline dimensions. By selecting different bibliometric indices as Skyline dimensions, RR-index can be fully customizable. However, since bibliometric indices are correlated [4], selecting such indices would yield analogous results in the final Skyline ranking. Also, as the number of dimensions increases, the Skyline size increases as well, which decreases the discrimination power of the RR-index.

3. The Case Study of Greek Computer Science Departments and Faculties

In [21] the effectiveness and reliability the RR-index has been tested experimentally on bibliographic data for authors. In the present work, the RR-index is generalized to higher conceptual levels. In this chapter, first we present the dataset used for our experiments with RR-index, then we introduce the Skyline features used by the RR-index, and finally we present the experimental results at three levels: at author, faculty and institutional level.

3.1 *Datasets*

For our experiments we have used the Microsoft Academic Search (MAS[9]) database. We have obtained the Microsoft Academic Graph from the Open Academic Graph workgroup[10]. The initial dataset consisted of 253,144,301 authors with 208,915,369 publications. Out of this initial dataset we kept only the publications having a Document Object Identifier (DOI[11]) as well as the publication year. This cleaning led to selecting 77,080,039 publications authored by 84,818,728 distinct researchers.

For the needs of the experiments, the authors of the Greek Universities were identified in the Microsoft Academic Graph database and two data sets were created:

- The first dataset consists of the academic staff of 19 Computer Science faculties of 17 major Greek universities. More specifically, in total, 539 authors were selected and verified according to the official website of each department.
- The second dataset consists of all authors/researchers with affiliation in the aforementioned 17 Greek universities. It should be noted that this dataset consists not only of the academic staff of the universities but also of all authors of the database, affiliated to any of these universities according to the Microsoft Academic Graph.

Moreover, we created a dataset consisting of international universities, namely the top-500 universities of the (NTU) ranking. This ranking is based on eight features categorized into three categories, namely research productivity, research impact, and research excellence. We collected all relevant data for replicating this ranking and feeding them into our method.

3.2 *Skyline Dimensions*

For the needs of the implementation of the Skyline operator and the creation of the rankings produced by the *RR*-index, we selected the scientific and metric indicators used by Google Scholar:

- *Cit* : the number of citations to all publications.
- *Cit*-5: the number of citations during the last 5 years to all publications.
- *h*-index: the largest number *h* such that *h* publications have at least *h* citations.
- *h* -index-5: the largest number *h* such that *h* publications have at least *h* new citations during the last 5 years.
- *i*10 : the number of publications with at least 10 citations.
- *i*10-5: the number of publications that have received at least 10 new citations during the last 5 years.

9 https://academic.microsoft.com
10 https://www.aminer.cn/oag2019
11 https://doi.org

Table 1

Cardinality of distinct values for each feature.

RR	*Cit*	*h*	*i*10	*Cit*-5	*h*-5	*i*10-5
539	440	38	74	334	24	41

The specific indicators were used for the following reasons. Firstly, because their combination gives a quite global and comprehensive view of the scientific impact of the author, the institution and in general any scientific entity, which is the focus of scientific evaluation. This is because with these indicators we take into account both the dimension of quantity and the dimension of the quality of the scientific work, as well as the dimension of time in which these achievements have taken place. Also, the fact that these indicators are used by Google Scholar means that they are generally accepted by the scientific community and they consist a daily measure of comparison and evaluation by a large number of scientists. Finally, their selection supports the reproduction of our results.

3.3 *RR-index for Faculty Members*

Initially the *RR*-index was calculated at the level of individuals, i.e., the 539 members of the CS faculties of the Greek universities. These 539 individuals were grouped in 539 ranking levels. In particular, Table 1 presents the number of distinct values of the *RR*-index (i.e. levels) along with the number of distinct values of the six features under consideration. One could argue that probably there is a meaning in choosing one (any) of the six features and try to rank these individuals. However, the first comment is that such a selection

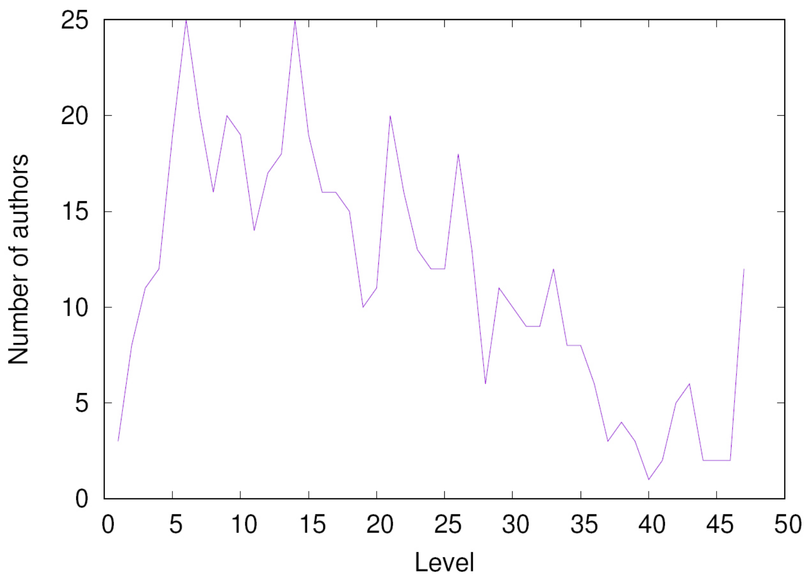

Figure 3

***RR*-level distribution of authors.**

Table 2

Rainbow Ranking for authors, the 4 top *RR*-levels.

Author	*RR*-level	*RR*-index	*Cit*	*h*	*i*10	*Cit*-5	*h*-5	*i*10-5
Nikos Hatziargyriou	1	99.81	11009	42	115	5501	26	73
George Karagiannidis	1	99.81	8079	49	155	3856	34	98
Ioannis Pitas	1	99.81	12568	55	226	2774	25	77
K.A. Antonopoulos	2	98.77	1744	23	38	948	20	26
Minos Garofalakis	2	98.77	4824	40	80	871	18	29
Yannis Manolopoulos	2	98.77	5651	33	104	1569	17	40
Petros Maragos	2	98.77	7499	42	122	1500	17	47
Konstantina Nikita	2	98.77	3405	29	98	1246	18	34
John Psarras	2	98.77	3212	30	89	1342	19	39
Grigorios Tsoumakas	2	98.77	3383	24	38	1785	18	28
Ioannis Vlahavas	2	98.77	3468	28	64	1560	18	32
Aggelos Bletsas	3	96.98	4285	19	34	1205	13	21
Pavlos Georgilakis	3	96.98	2367	25	57	1439	14	29
Aggelos Kiayias	3	96.98	5862	25	33	554	14	16
Stefanos Kollias	3	96.98	4783	31	95	1007	14	19
Aristidis Likas	3	96.98	4068	34	64	1350	17	32
Sotiris Nikoletseas	3	96.98	2331	26	75	679	13	21
Stavros Papathanassiou	3	96.98	2620	25	41	1286	19	27
Ioannis Pratikakis	3	96.98	2729	28	55	1153	19	38
Anastasios Tefas	3	96.98	2675	26	64	1178	17	37
Sergios Theodoridis	3	96.98	3535	27	72	1128	16	30
Yannis Theodoridis	3	96.98	3920	31	65	1095	17	32
Dimitris Achlioptas	4	94.80	2935	25	40	930	15	21
Lefteris Angelis	4	94.80	2093	25	54	834	17	28
Antonis Argyros	4	94.80	2794	24	49	1223	16	24
Hercules Avramopoulos	4	94.80	2615	26	70	532	10	11
Elias Glytsis	4	94.80	2624	30	68	291	9	6
Yannis Ioannidis	4	94.80	3959	33	58	727	15	25
Kostas Kalaitzakis	4	94.80	3242	20	26	1204	17	21
Elias Kosmatopoulos	4	94.80	2693	23	47	1130	18	27
Constantine Kotropoulos	4	94.80	2791	28	61	777	13	20
Konstantinos Parsopoulos	4	94.80	3787	24	41	1003	15	23
George Polyzos	4	94.80	3971	29	56	1103	14	19
Dimitris Syvridis	4	94.80	3129	24	68	1080	13	23

would be arbitrary and probably biased. On the other hand, ranking these individuals by using *h*-5 or *i*10-5 could lead to many ties. It is here that the *RR*-index comes: it *encompasses all* these six features and clusters the individuals in 47 groups.

As mentioned before, Rainbow Ranking creates 47 levels for the 539 authors, which means 11 authors per level on the average, approximately. Figure 3 illustrates the distribution of the number of authors per level. The 1st level includes only three authors, whereas the most populous levels are the 6th and 14th, with 25 authors. Table 2 presents the top-4 ranking levels of the members of the CS faculties as derived by their *RR*-index.

We note that at every level there are authors who, in at least one Skyline feature, have a better score than other authors of the same level. However, no author is dominated by any other author at the same level, as conceptually it is expected from the Skyline notion. In addition, at each level there is at least one author with better scores with respect to all features, in comparison to at least one author of the next/lower level, and therefore dominates him. In other words, any author at level *i* is dominated by at least one author at level ($i - 1$). With this method, a grouping of authors is performed at subsequent levels, so that the members of one group are collectively "better" than those of the next level. In other words, instead of having an absolute ranking of individuals, which most probably have similar research output profiles, we introduce a group ranking, with clear-cut and tunable rules, which take into account many dimensions without using arbitrary weights.

For example, according to the *Cit* metric, which expresses the total number of citations received by an author, Ioannis Pitas should be ranked first. On the other hand, if the *Cit*5 metric, which expresses the citations received by an author during the last five years, is adopted, then it is Nikos Hatziargyriou who should be ranked first. Finally, according to the *i*105 dimension, which expresses the number of publications with at least 10 citations during the last five years, then George Karagiannidis should take the lead.

3.4 *RR-index for CS Faculties*

By stepping to a higher conceptual level and generalizing the previous approach, we compute the *RR*-index of the 19 largest CS faculties, where the previous 539 individuals belong. This generalization is achieved by accumulating all the values of the adopted 6 features of all the faculty members belonging to each faculty. Thus, for example, the *Cit* value expresses the total number of citations received by the publications of all faculty members of each department. Table 3 shows the *RR*-index of these 19 departments, which are grouped in nine Skyline levels.

Table 3

Rainbow Ranking for CS faculties

Fac-Univ	**#Staff**	***RR*-level**	***RR*-index**	***Cit***	***h***	***i*10**	***Cit*-5**	***h*-5**	***i*10-5**
ece-ntua	71	1	100	91077	105	1822	30979	60	421
di-uoa	39	2	92.11	46897	88	899	12066	40	184

Contd...

inf-auth	29	2	92.11	53740	98	877	17588	50	197
csd-uoc	24	3	78.95	29925	77	552	8188	36	120
ece-tuc	24	3	78.95	28858	75	432	8902	38	115
ee-auth	28	3	78.95	28842	72	612	10542	44	191
ee-duth	38	4	65.79	22106	63	513	8360	34	111
inf-aueb	25	4	65.79	25623	72	497	8128	35	108
ceid-upatras	23	5	52.63	17740	52	464	4654	23	36
cs-uoi	25	5	52.63	23768	70	426	6934	35	105
ece-upatras	38	5	52.63	20764	57	462	6068	26	60
dit-uop	24	6	36.84	13736	53	315	4214	27	55
icsd-aegean	17	6	36.84	10243	41	249	4848	32	60
inf-unipi	21	6	36.84	14072	56	305	5370	31	84
ainf-uom	31	7	26.32	9040	40	206	4204	26	66
ece-uth	18	8	21.05	7687	38	161	2629	23	44
ice-uniwa	23	9	10.53	4673	31	105	1952	18	28
iee-ihu	28	9	10.53	5287	31	115	1705	18	28
inf-ionio	13	9	10.53	4460	34	99	2031	20	22

The first ranking level consists of one faculty member only: the School of Electrical and Computer Engineering of National Technical University of Athens. This faculty is the largest faculty in Greece and apparently is favored by its size. Apparently, normalization is necessary to debias the results with respect to size. However, debiasing is not the issue of this current study.

On the other hand, we also notice that there are 4 ranking levels consisting three CS faculties. This fact is a proof of concept, i.e., these faculties have the same *RR*-index and belong to the same group, without any of them dominating the others. Still one of the members of a level dominates one of the members of the lower level, thus forming a distinct hierarchy of groups.

3.5 *RR-index for 17 Greek Universities*

Finally, the *RR*-index values for the above 17 Greek universities were calculated. We use the second dataset which comprises of all authors affiliated with these universities. Again, note that each feature value was accumulated over the total number of the academic staff in each university. Table 4 shows the accumulated results for the six Skyline features as well as the values of the *RR*-index and corresponding level for the Greek universities. Table A1 shows the full names of the universities.

We observe that these 17 universities are grouped in twelve ranking levels. The grouping created by applying our Rainbow Ranking method was relatively limited. This is due to the fact that the number of universities is small and the feature values vary widely. In

Table 4

Rainbow Ranking for Greek Universities

University	*RR*-level	*RR*-index	*Cit*	*h*	*i*10	*Cit*-5	*h*-5	*i*10-5
uoa	1	100	7078897	841	61172	3016544	552	24711
auth	2	91.18	3356467	548	32761	1578225	368	11249
ntua	2	91.18	2663388	498	19926	1416386	378	6886
uoi	3	79.41	2156665	513	20445	952141	344	8387
uoc	3	79.41	2125246	466	24320	805642	264	8769
upatras	4	70.59	1773229	376	23733	679132	216	7062
duth	5	61.76	633733	270	7395	283990	164	2901
uth	5	61.76	659359	275	8749	293147	165	2898
aegean	6	50.00	342804	205	5118	173323	151	1907
tuc	6	50.00	372983	226	5074	149161	136	1858
aueb	7	41.18	246569	183	3701	95860	107	1218
unipi	8	35.29	128987	138	2121	55475	84	660
uom	9	26.47	83994	111	1627	38222	70	456
uop	9	26.47	60460	101	1119	27508	72	434
uniwa	10	17.65	67113	100	1238	31924	62	301
ihu	11	11.76	18160	67	340	9666	47	150
ionio	12	5.88	13776	48	252	6859	33	42

turn, the latter fact is due to the different sizes of the universities both in terms of the number of faculties as well as the number of academic staff (besides the inherent differences of the qualities/strengths of the personnel).

We emphasize that although in the ranking of 539 authors there are other indicators that can create groupings (due to ties) similar to *RR*-index, in the ranking of CS faculties and universities, no indicator can create groups, whereas *RR*-index is the only method that produces levels by creating groups.

4. The Case Study of the Top-500 International Universities of the NTU Ranking

In this section we apply our methodology to the top 500 universities of the National Taiwan University Ranking (NTU) list. We choose to work with this ranking list, because it is based exclusively on verifiable research performance indicators. From the NTU ranking we focused on the top-500 universities, and extracted their involved data. There were no missing values. We applied our skyline methodology to these 500 universities and compared our ranking with the NTU ranking. The Spearman correlation coefficient [7] of the two ranked lists was 0.91 and the Rand correlation index [17] was 0.93, which are quite high, but not really close to a perfect positive correlation. There were some significant

Table 5

Comparison of the first two skyline levels with the respective NTU ranking.

University	skyline level	NTU rank (value)
Harvard University	1	1 (98.6)
Massachusetts Institute of Technology	1	7 (59.9)
University of California, San Francisco	1	15 (54.7)
University of California, Berkeley	1	16 (54.5)
Stanford University	2	2 (65.7)
University of Toronto	2	3 (62.2)
Johns Hopkins University	2	4 (61.3)
University of London, University College London	2	5 (61.2)
University of Oxford	2	6 (60.9)
Tsinghua University	2	17 (54.5)
Shanghai Jiao Tong University	2	25 (50.6)
Zhejiang University	2	30 (49.8)
Washington University in St. Louis	2	32 (49.2)
Mayo Medical School	2	40 (47.9)
University of Texas, M.D. Anderson Cancer Center	2	58 (45.5)
California Institute of Technology	2	65 (44.9)
University of California, Santa Cruz	2	153 (39.1)

discrepancies among the two rankings, especially in the top positions. In Table 5 we show the first two skyline "levels" – which correspond to RR-index (Equation 1) values to 99.6 and 97.9, respectively – and contrast these universities' position to their respective position in NTU ranking.

5. Conclusions

There are several university rankings that regularly publicize their annual results, which are then reproduced in newspapers and social media. Although these university rankings have been intensively criticized in the literature related to research evaluation and scientometrics, they are very popular and they seriously affect the ecosystem of higher education.

Here, we propose an alternative approach to rank universities by elaborating on the multidimensional Skyline operation, and the Rainbow Ranking methodology. In particular, our method provides ranked sets, like equivalent classes, instead of ranked lists as provided by the traditional scientometrics indicators. Thus, it will be safe to claim that a university is better than another, if it is found in a higher equivalence class according to the set of selected dimensions. It should be noted that our method classifies with the same respect in all institutions from the first to the last, without creating arbitrary sets of 100 or 500 institutions such as other ranking lists. At the same time, the score of the *RR*-index gives us a direct perception of the institution's position in the overall ranking since it is in the range 0 – 100.

We have tested our method with data extracted from Microsoft Academic Graph concerning Greek computer science departments and faculty members. We have also enriched the dataset by visiting university websites and extracting real and fresh data related to the author identification and affiliation, where it was possible. We have cleaned our data by removing unreliable publications and citations such as publications in unverified authorities or with incomplete and erroneous metadata. We have applied our method on the cleaned data and we have come up with a fair ranking of the universities, which is well received if we look as professional peers.

Note

A preliminary version of this manuscript [20] appeared in the proceedings of the 1st International Workshop on Assessing Impact and Merit in Science (AIMinScience), 2020.

Appendix A. Greek universities full names

Table A1

Greek Universities full names

Acronym	University Name
aegean	University of Aegean
aueb	Athens University of Economics & Business
auth	Aristotle University of Thessaloniki
duth	Democritus University of Thrace
ihu	International Hellenic University
ionio	Ionian University
ntua	National Technical University of Athens
tuc	University of Crete
unipi	Universtity of Piraeus
uniwa	University of West Attica
uoa	National & Kapodistrian University of Athens
uoc	University of Crete
uoi	University of Ioannina
uom	University of Macedonia
uop	University of the Peloponnese
upatras	University of Patras
uth	University of Thessaly

References

[1] Lefteris Angelis, Nick Bassiliades, and Yannis Manolopoulos. On the necessity of multiple university rankings. *COLLNET Journal of Scientometrics & Information Management*, 13(1):11–36, 2019.

[2] Lutz Bornmann and Wolfgang Glänzel. Which differences can be expected when two universities in the Leiden ranking are compared? Some benchmarks for institutional research evaluations. Scientometrics, 115:1101–1105, 2018.

[3] Lutz Bornmann and Werner Marx. Thomas theorem in research evaluation. Scientometrics, 123(1):553–555, 2020.

[4] Lutz Bornmann, Rüdiger Mutz, Sven E. Hug, and Hans-Dieter Dieter Daniel. 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, 2011.

[5] Stephan Börzsönyi, Donald Kossmann, and Konrad Stocker. The skyline operator. In Proceedings 17th IEEE International Conference on Data Engineering (ICDE), pages 1–20, Heidelberg, Germany, 2001.

[6] M.-L. Bougnol and J. H. Dula. Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. Annals of Operations Research, 145:339–365, 2006.

[7] Spearman C. The proof and measurement of association between two things. *American Journal of Psychology*, 15(1):72–101, 1904.

[8] Christopher Claassen. Measuring university quality. Scientometrics, 104(3):793–807, 2015.

[9] F. García, F. Guijarro, and J. Oliver. A multicriteria goal programming model for ranking universities. *Mathematics*, 9(5), 2021.

[10] F. Guijarro, M. Martínez-Gomez, and D. Visbal-Cadavi. A model for sector restructuring through genetic algorithm and inverse DEA. Expert Systems with Applications, 154, 2020.

[11] Jill Johnes. University rankings: What do they really show? Scientometrics, 115(1):585–606, 2018.

[12] T. Luque-Martínez and N. Faraoni. Meta-ranking to position world universities. Studies in Higher Education, 45(4):819–833, 2020.

[13] Yannis Manolopoulos and Dimitrios Katsaros. Metrics and rankings: Myths and fallacies. In Revised Selected Papers, 18th International Conference on Data Analytics & Management in Data Intensive Domains (DAMDID/RCDL), pages 265–280, Moscow, Russia, 2017.

[14] P. Perchinunno and M. Cazzolle. A clustering approach for classifying universities in a world sustainability ranking. Environmental Impact Assessment Review, 85, 2020.

[15] Fredrik Niclas Piro and Gunnar Sivertsen. How can differences in international university rankings be explained? Scientometrics, 109(3):2263–2278, 2016.

[16] S. Rahnamayan, S. Mahdavi, K. Deb, and A. Asilian-Bidgoli. Ranking multimetric scientific achievements using a concept of Pareto optimality. Mathematics, 8, 2020.

[17] W. M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336) : 846–850, 1971.

[18] S. Rousseau and R. Rousseau. Data envelopment analysis as a tool for constructing scientometric indicators. Scientometrics, 40(1):45–56, 1997.

[19] Antonios Sidiropoulos, Antonia Gogoglou, Dinitrios Katsaros, and Yannis Manolopoulos. Gazing at the skyline for star scientists. *Journal of Informetrics*, 10(3):789–813, 2016.

[20] G. Stoupas, A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos. Skyline-based university rankings. In Proceedings of the ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium, pages 347–352, Lyon, France, 2020.

[21] Georgios Stoupas, Antonis Sidiropoulos, Antonia Gogoglou, Dimitrios Katsaros, and Yannis Manolopoulos. Rainbow ranking: an adaptable, multidimensional ranking method for publication sets. Scientometrics, 116(1):147–160, 2018.

[22] William Isaac Thomas and Dorothy S. Thomas. The Child in America. Oxford:Knopf, 1928.

[23] Eleftherios Tiakas, Apostolos N. Papadopoulos, and Yannis Manolopoulos. Skyline queries: An introduction. In Proceedings 6th International Conference on Information, Intelligence, Systems & Applications (IISA), pages 1–6, Corfu, Greece, 2016.

[24] Anthony F.J. Van Raan. Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. Scientometrics, 62(1):133–143, 2005.

[25] D. Visbal-Cadavid, M. Martínez-Gomez, and F. Guijarro. Assessing the efficiency of public universities through DEA: A case study. Sustainability, 9(8), 2017.

[26] M. Voorneveld. Characterization of Pareto dominance. Operations Research Letters, 31(1):7–11, 2003.

[27] Ludo Waltman, Paul Wouters, and Nees Jan van Eck. Ten principles for the responsible use of university rankings, 2017.