İKTİSADİ İDARİ VE SİYASAL ARAŞTIRMALAR DERGİSİ
JOURNAL OF ECONOMICS BUSINESS AND POLITICAL RESEARCHES

e-ISSN: 2564-7466	Cilt:10, Sayı:26, Şubat 2025
https://dergipark.org.tr/tr/pub/iktisad	Vol:10, No:26, February 2025

Towards Global Academic Performance Rankings: A Dynamic and Integrated Decision Support System Based on Scientometric Indicators in Different Databases

♦♦♦

Küresel Akademik Performans Sıralamalarına Doğru: Farklı Veri Tabanlarındaki Bilimsel Göstergelere Dayalı Dinamik ve Entegre Bir Karar Destek Sistemi

https://doi.org/10.25204/iktisad.1582267

Mahmut BAYDAŞ*

Article Info

Paper Type:
Research Paper

Received:
09.11.2024

Accepted:
27.12.2024

Abstract

This study aims to propose a decision support system based on multi-criteria decision-making (MCDM) methodologies in order to reach individual and global-scale academic performance, which is a neglected subject. Unlike previous classical applications and past studies, in this study, different science indicators (citation counts, article counts, and field-based impact) taken from different databases (Scopus, Web of Science, InCites, Google Scholar) were combined, and objective weights were assigned to each criterion. These indicators, weighted with the entropy method, were analyzed with CRADIS and other alternative methods. The analysis results showed that the Q1 article count and field-based impact scores were of high importance, whereas Google Scholar citations had lower weight. In accordance with the recommendation of the Leiden manifesto, which had a great impact on the academic community, to take into account multi-indicator and being field-based, the system proposed in this study also allows for the dynamic (updatable) and comprehensive evaluation of individual researcher performance. Compared to one-sided and limited performance measurements in literature or applications, this study fills a serious gap. Moreover, this system will help the parties to make accurate and updatable strategic decisions.

Keywords: *Individual and global academic performance, multi-criteria decision making, cradis method, entropy method.*

Makale Bilgileri

Makale Türü:
Araştırma Makalesi

Geliş Tarihi:
09.11.2024

Kabul Tarihi:
27.12.2024

Öz

Bu çalışma, ihmal edilen bir konu olan küresel ölçekte bireysel akademik performansa ulaşma yolunda çok kriterli karar verme (ÇKKV) metodolojilerine dayalı bir karar destek sistemi önermeyi amaçlamaktadır. Daha önceki klasik uygulama ve geçmiş çalışmaların aksine bu çalışmada farklı veri tabanlarından (Scopus, Web of Science, InCites, Google Akademik) alınan farklı bilim göstergeleri (atıf sayıları, makale sayısı ve alan bazlı etki) birleştirilmiş ve her bir kritere nesnel ağırlıklar atanmıştır. Entropi yöntemiyle ağırlıklandırılan bu göstergeler, CRADIS ve diğer alternatif yöntemlerle analiz edilmiştir. Analiz sonuçları, Q1 makale sayısı ve alan bazlı etki puanlarının yüksek öneme sahip olduğunu, buna karşın Google Akademik atıflarının daha düşük ağırlık taşıdığını göstermiştir. Akademik camiada oldukça etki bırakan Leiden manifestosunun çok göstergeli ve alan bazlılığın dikkate alınması önerisine uygun olarak bu çalışmada önerilen sistem, bireysel araştırmacı performansının dinamik (güncellenebilir) ve kapsamlı bir şekilde değerlendirilmesine de olanak tanımaktadır. Literatür veya uygulamalardaki tek yönlü ve kısıtlı performans ölçümlerle kıyaslandığında bu çalışma ciddi bir boşluğu doldurmaktadır. Dahası bu sistem taraflara isabetli ve güncellenebilir stratejik karar vermeye yardımcı olacaktır.

Anahtar Kelimeler: *Bireysel ve küresel akademik performans, çok kriterli karar verme, cradis yöntemi, entropi yöntemi.*

Atıf/ to Cite (APA): Baydaş, M. (2025). Towards global academic performance rankings: A dynamic and integrated decision support system based on scientometric indicators in different databases. *Journal of Economics Business and Political Researches, 10*(26), 154-174. https://doi.org/10.25204/iktisad.1582267

* ORCID Lecturer Dr., Necmettin Erbakan University, Faculty of Applied Sciences, Department of Accounting and Financial Management, mbaydas@erbakan.edu.tr

1. Introduction

'Performance measurement' (whether with a systematic scientific background or superficially) has been done from the past to the present. If it is not well-defined what is intended to be measured, and where and when it will be used, there will be a lot of doubts about the reliability of the correct measurement. Incorrect performance measurements that are not made in a comprehensive, systematic, and dynamic manner will have many harms. The main purposes of performance measurement can be listed as determining the level of success and achievability of targets, understanding processes, scanning for problems and bottlenecks, and ensuring that improvement decisions are based on objective data. Thus, measurement is used to improve performance. Moreover, effective performance measurement should reflect real results, should not make false comparisons, should include normalized metrics that can be used in comparison, should be practical and easy to understand, and finally should provide a benefit that exceeds the cost (Parker, 2000). As is known, the 'scientific performance' of an institution or an individual plays a locomotive and vital role in the development of countries' industry, technology, politics, economy, etc., and especially in the production and dissemination of knowledge. In this respect, the global scientific performance of universities and, in part, individuals are updated and published every year, and success rankings are made. Thanks to this performance and the metrics that determine performance, prestigious universities strive to attract high-quality students and academics from all over the world. However, it is interesting that even when common criteria are considered, there may be differences between the ranking index results. This shows that measuring the knowledge production and research performance of universities is not as easy as it seems (Olcay and Bulu, 2017).

The purpose of rankings should of course be to compare institutions or individuals according to their merit and performance. Various methodologies are still used for these rankings. Performance indicators are generally used from data obtained from some databases and other primary sources. On the other hand, the accuracy or adequacy of the data and criteria used for performance in reflecting real performance can be questioned. There are many performance criteria for university rankings such as teaching, research, citations, knowledge transfer, international faculty member ratio, international student ratio, per capita performance, regional research reputation, faculty member/student ratio, and academic reputation, and these are evaluated in an integrated manner. The weight of importance of these criteria and how much they really reflect being a performance criterion are important. For example, the data used regarding 'research performance', which is a very important criterion, should adequately reflect the adequacy, scope, and purpose. University administrations and academics undoubtedly want to see information about their own research performance comparatively based on metrics, understand their current situation, and use it in decision-making beyond this. As is known, the general and integrated performance of universities with multiple criteria is published annually by some global-ranking organizations. In these rankings, numerical and objective data (such as the number of publications, and citation numbers) such as scientometric or bibliometric indicators, as well as perception-based (such as reputation surveys) or structural measures (such as the faculty/student ratio) can be used (Szluka et al., 2023). A very important component of these rankings is the parameters related to scientific research productivity, which are scientometric or bibliometric indicators. In other words, the most important parameters affecting the ranking positions include citations and prestigious scientific publications.

An important point here is that the meticulousness and careful focus in 'university performance measurement' is withheld from 'individual researcher performance measurement'. In other words, some partial and incomplete dimensional rankings based on ready-made metrics for 'individual global research performance measurement' may not be as healthy and well-thought-out as institutional performance measurement. In addition, various databases and the scientific indicators they contain are often scattered, waiting to be transformed into an aggregate performance metric. Therefore, the performance of individuals should be examined sufficiently and even deeply. While the public knows the performance of universities well thanks to many ranking organizations, the same public does not have comprehensive, versatile, and detailed information about the special performance of individuals. This situation prevents universities from distinctively discovering, recognizing, recognizing, and identifying scientists with the high-impact value, effectiveness, capacity, and productivity they need. As is known, scientific impact and rankings are important in determining funding as well as providing global recognition and prestige. The factor that restricts or prevents this to some extent is the lack of a correct, dynamic, multi-criteria, integrated global-based individual performance measurement ranking metric.

Especially large academic institutions move like slow elevators in similar ranking ranges every year. It is not an easy task for the performance of such institutions to change even slightly. On the contrary, the ranking performance of individuals can show extraordinary changes in the same time intervals. Individual academic performance, which constitutes one of the most important performance dimensions of universities today, is related to research productivity, effectiveness, and efficiency of individuals. Today, the academic performance of individuals is generally published by updating them around indicators based on a different database, with each ranking being within the framework of the 'concepts of scientific publication and citation'. For example, the list known as 'World's Best Scientists' can be given as an example of this, which uses the Scopus database when creating this list. (Ioannidis, 2024). In fact, it is an interesting deficiency that a multidimensional, comprehensive, and integrated success ranking for scientists is not made by using different databases and metrics (the details of which will be discussed in the later parts of this study), and this situation emerges as a gap in literature. Perhaps each database may not need the others because it sees itself as the most superior and sufficient. However, this is not the right approach because no database is 100% comprehensive and encompassing.

In recent years, lists of the best scientists have been announced by prestigious institutions and these achievements are followed with interest by universities. The 'Best Scientists Rankings' published by Elsevier BV is a popular ranking referred to in this context (Ioannidis, 2024). The 'Highly Cited Researchers' (HCR) category published by Clarivate is a popular indicator of distinguished individual researchers and is based on the number of WOS citations received by WOS (Web of Science)-based publications (Docampo and Cram, 2019). On the other hand, there are also metrics that rank individual performance by filtering according to the user's wishes. The embedded, dynamic, and well-structured autonomous ranking obtained by 'InCites' using the WOS database is an example of this (Markusova et al., 2023). The best scientist rankings of the AD Scientific Index can be shown as another example (AD Scientific Index, n.d.). There are two important problems here: the world's best scientists are being evaluated with only a few criteria from a database. However, it is more accurate to make an integrated performance measurement with indicator metrics obtained from different databases. For example, there are scientists who have publications scanned in Scopus but not in WOS, and their names may not be mentioned in WOS metrics. This should not mean that their performance is zero. Second, most universities naturally do not have enough or any of the world's best scientists. They may demand to see the performance of their current scientists, but they cannot collect integrated performance measurement information sufficiently. Therefore, they have difficulty in developing an insightful policy because they do not know the current performance comprehensively, accurately, and sufficiently. In the current limited metrics, some use only the results of metrics such as WOS, some only Scopus, and some only Google h-index. However, a global academic performance ranking can be made with criteria such as the number of WOS publications in the highest impact article class, the number of WOS Q1 articles, the number of WOS citations with high impact value, CNCI scores, Scopus FWCI score, the number of Scopus articles and citations, and the number of Google Scholar citations, each of which is important, valuable and partially includes shortcomings according to its own weight.

Indeed, the list of ten principles to guide research evaluation published in the 'Leiden Manifesto (LM)' in 2015 for research metrics that are widely accepted on a global scale also supports our proposed argument. LM includes a set of 10 principles that recommend more responsible and fair use of bibliometric or scientometric methods (h-index, impact factor, and website display indicators such as metrics) in scientific research performance evaluations. LM addresses the problems caused by the incorrect or excessive use of metrics in scientific performance evaluations and provides guidance for a more balanced and ethical use of these metrics. According to LM, first, differences in publication and citation practices should be taken into account according to the field. Citation rates can vary greatly between different disciplines; for example, the impact factors of the best journals in finance may be very different from the impact factors of the best journals in oncology. This feature suggested by LM is also present in the metrics called CNCI provided by InCites and FWCI provided by Scopus. Unfortunately, this feature is not used by raters as an indicator of either individual performance or university rankings. Again, according to LM, it is necessary to avoid unnecessary concreteness and false certainty. For example, when looking at a specific scientist, a low citation rate may lead the researcher to assume low research quality; this implies causality from correlation. Using more than one robust indicator can reduce inappropriate concreteness. Finally, according to LM, it is useful to measure performance according to the researcher's research missions and it is also necessary to regularly review and update the indicators (Hicks et al., 2015).

As a solution technique for such problems, it is logical to use the MCDM methodology, which can handle many indicators simultaneously and with different weights and, moreover, has more than 200 types. In addition, it was mentioned above that it is necessary to regularly review and update the indicators according to LM. We know that WOS, Scopus, google.scholar, and especially InCites update their data regularly. The CNCI score, which is perhaps closer to the ideal metric described by LM, is updated monthly by InCites by pulling WOS data. The Stanford list based on Scopus data can make the CNCI metrics list of scientists based on field. Instead of giving priority to high-impact journals, research conducted in a specific field or area should also be allowed to be published in relevant local research journals. In this respect, when measuring researcher performance, not only WOS or Scopus data but also google.scholar publication citations can be evaluated. Again, LM recommends that data collection and analytical processes be kept open, transparent and simple. It can be said that WOS, Scopus and google.scholar citation and publication metrics try to provide sufficient attention in this regard (scholar.google.com/; clarivate.com/highly-cited-researchers/; incites.clarivate.com/#/analysis/0/person ; scopus.com/search/form.uri#basic).

In our age, a large amount of data is collected, and hidden and interesting patterns, trends, and relationships are discovered through data processing, data analytics, and data mining. Thus, raw data is ultimately transformed into useful information for the decision-maker. Moreover, new insights are reached by reproducing and distilling data from useful information. Artificial intelligence and machine learning applications can easily apply many quantitative theorems that were previously in the scientific sense and were passive, on data. In this respect, decision support systems are becoming even stronger with the accompaniment of artificial intelligence. Likewise, these systems are gradually progressing toward becoming autonomous decision-makers. Therefore, databases containing data on researcher performance, metrics, or scientific indicators can be pulled and simultaneously integrated, and holistic performance can be presented to users autonomously with the help of the MCDM methodology. Stakeholders such as researchers, reviewers, editors, university administrations, public administrators, ranking producers, funders, project evaluators, associate professorship, and doctoral committees, who are among the information users, can evaluate integrated performance to make more robust and reliable decisions.

Current university rankings may not be exciting enough to motivate scientists. However, scientists being aware of their position in rankings showing the general and regional, national, or global performance of universities in their fields, learning and understanding the current situation will certainly increase their motivation and insight. Moreover, university administrations can also develop interesting and attractive policies for scientists with high scientific research performance. These policies have been effective in some universities such as Shanghai Jiaotong, Tsinghua, and Zhejiang in China, which have lower opportunities but can compete with universities with very high budgets such as Harvard, Stanford, and MIT (URAP, 2023). Likewise, some scientists can receive more WOS-based Q1 articles and citations than some universities on their own, and this type of extraordinary success should be well understood and evaluated by those concerned. For example, individuals in some small study groups can enter lists such as Stanford's best scientists in the world, and on the contrary, not a single scientist from some universities can enter these lists institutionally. This situation urgently warns and points out the need to better understand the value of quality scientists and to develop policies accordingly.

On the other hand, there are some methodological nuances in calculating performance rankings. As in world university rankings, weighted aggregation methods are mostly used in individual research performance. Although this method seems logical on the surface, it also has some problems. In recent years, some fundamental problems have been identified in this regard. For example, problems such as false precision, weight inconsistencies, mutual compensation between criteria, lack or redundancy of indicators, and rank reversal can be listed. These problems can cast doubt on the reliability of ranking results (Soh, 2017). In this respect, it is useful to benefit from the MCDM (Multi-Criteria Decision Making) methodology family, which has more than 200 members. The methodologists of this method have been addressing the problems in question for almost half a century and offering solutions. For example, contemporary MCDM methodologies have brought many suggestions to solve and improve MCDM problems such as fair weighting, sensitivity analysis, compensability, and rank reversal (Baydaş et al., 2024). The main purpose of this study is to offer suggestions to scientific ranking and metric producers to reveal the individual performance of scientists with MCDM methodologies, which are a decision-support system tool.

After the introduction above, the literature research will be given in the second section below, information about the material and method will be given in the third section. In the fourth section, the analysis findings will be discussed in the discussion section and finally, some insights will be conveyed to the readers in the conclusion section.

2. Literature Review

Although the performance of universities is measured annually by various rating agencies, there are not enough ranking producers or practitioners regarding individual performance. On the other hand, while literature is more focused on global university performance, global researcher performance rankings have not been sufficiently addressed. Here, researchers are not considered as a unit but in terms of their academic contribution to the institution. However, it is understood from the indicators that each researcher has a special story and that some researchers perform better than the institution. The best, worst, or mediocre institutions perform almost the same every year. And this is not very interesting and exciting and may not even have significant news value. On the other hand, it is possible for researchers with good performance to work in institutions that are consistently mediocre. In other words, individual performance can be quite flexible, and researchers can surpass their peers and take their place in the best ranking lists in the short or long term with hard work.

A common framework regarding what should be the universal research performance indicators of scientists and their measurement with MCDM has not been proposed in the past. There is also no clear set of indicators agreed upon regarding research performance criteria. However, research performance is usually analyzed in two dimensions. The first is productivity, which depends on the publication output (number). The second is impact (citation), which is an indicator of the quality of publications (Maral, 2024). However, there are many types that address various aspects of these two indicators, and the large number of them necessitates addressing the subject with MCDM methodology. Here, we will touch upon some limited studies conducted on researcher performance using the MCDM methodology. In their study, Li et al. (2014), wanted to test the applicability of multi-criteria group decision analysis methodology for the evaluation of academic research outputs.

In 2012, the publication and citation numbers of 20 researchers (6 professors, 5 associate professors, and 9 senior assistants) from Sichuan and Tsinghua Universities taken from the Thomson Reuters ISI WOS database were evaluated. In the study, 8 evaluation criteria were determined (6 of them are subjective evaluation criteria). The importance levels were determined by the professors. An intuitionistic fuzzy weighted average operator was used in a fuzzy environment to determine the criteria weights. A fuzzy distance-based method was also developed to determine the weights of the evaluators. In the study, crisp and fuzzy ratings were combined using the revised TOPSIS method and scientists were ranked with this method. This structure was an approach that considered more criteria by aiming for maximum group consensus in evaluations. Tuan et al. (2020), developed a TOPSIS MCDM model integrated with fuzzy AHP and Neutrosophic to evaluate the research productivity of four faculty members. The data of faculty members at the Faculty of Economics and Business, Vietnam National University were evaluated. Fuzzy AHP was used as the weighting method for the criteria of number of publications, publication quality, number of books, graduate student advisors, and research grants received as a project leader. As a result, it was shown that this integrated MCDM approach is effective and applicable. Ardil (2021), ranked seven researchers working in similar research areas according to four evaluation criteria, namely the number of annual and career-long publications, citations, and scholar index (h-index) using the TOPSIS method. Since the dataset covers more journals and publication types compared to other databases, it includes cumulative impact and five-year impact indicators obtained from the Google Scholar database, which can often find more citation references. Sensitivity analysis was also performed, and it was emphasized that the proposed MCDM method yielded reliable results.

Here, it is necessary to evaluate the issue separately in the context of existing popular ranking metrics based on databases. We first look at the ‘Stanford University Scientists Ranking’ published by Elsevier BV (Ioannidis, 2024): We want to examine studies that evaluate the global research performance of scientists in the context of popular rankings. First, we look at the “Stanford University Scientists Ranking” published by Elsevier in 2024. This ranking highlights the top 2% of researchers worldwide. The list includes normalized data based on various indicators such as citation counts and h-index. The Stanford/Elsevier ranking is prepared with data obtained from Scopus and covers 22 scientific fields and 174 subfields. Performance indicators include criteria such as the number of articles, citations, h-index, and co-authorship-adjusted hm-index. The evaluation is done in two ways: Career-long performance and the impact of the last year. In addition, self-citations and other citations are evaluated separately. All scientists with at least 5 Scopus articles are included in this ranking. The ranking includes the top 100,000 scientists by c-score or the top 2% in the subfield. Career-long data includes citations up to the end of the previous year, while recent-year data includes citations within the last calendar year. This study is based on Scopus data provided through Elsevier’s ICSR Lab, and the computational outputs are updated each autumn.

If we evaluate, first of all, this list has been important and useful in terms of being one of the first comprehensive and systematic attempts based on scientific indicators. The second positive aspect is that it has an objective ranking measured on a multi-criteria plane, since the performance information of scientists is directly taken from the Scopus database (Ioannidis et al., 2020). The disadvantage is that it only uses Scopus scientific publication and citation data. Although the scope and context are slightly different, there may be scientific publications and citations that are not in Scopus but scanned in Google Scholar. In other words, it would be fairer to include scientific publications and citations that are not scanned in both Scopus and WOS in the measurement, even if a small weight coefficient is given. In addition, this metric does not evaluate successful and high-quality articles in the Q1 quarter. Another aspect that is open to criticism is that this systematic does not give enough importance to the criteria weights. Finally, another issue we suggest is that an indicator related to field-based citation impact can be added to this list. The FWCI score, which is actually a Scopus-based score (calculated in a broader context), can be dynamically (updated) added to the individual's overall score. And it can also be suggested that the final ranking list be updated more quickly for smaller time periods rather than annually.

The second popular ranking list we focused on is the “Highly Cited Researchers” ranking list published by Clarivate, which is entirely based on citations. And in this respect, it is open to criticism. However, we can say that Claravative has closed this limited measurement gap with its embedded, dynamic (updatable), filtered, and very well-structured autonomous ranking obtained by using InCites’ WOS database. Here (in InCites), thanks to the large number of filtering features, hundreds of different rankings can be obtained autonomously for researchers. In addition, the number of criteria can be increased. However, the lack of an integrated performance system here as in Scopus (i.e., the scattered nature of single criterion data) directs information users to make individual calculations for total performance. On the other hand, it is advantageous that it is quite practical and easy to use. (https://clarivate.com/highly-cited-researchers/; https://incites.clarivate.com/#/analysis/0/person)

According to the ‘Leiden Manifesto’, first of all, differences in publication and citation practices should be taken into account by field. Citation rates can vary greatly across disciplines, for example, the impact factors of the top journals in finance may be very different from those of the top journals in oncology (Hicks et al., 2015). This insightful feature suggested by LM is found in the metrics CNCI provided by InCites and FWCI provided by Scopus. The researcher CNCI (Category Normalised Citation Impact) score embedded in InCites has gained increasing popularity in recent years due to its depth, meaning and usefulness. The (CNCI) indicator puts citations in good context. After normalising the citation counts by year, document type and subject category, it allows reasonable comparisons to be made based on the impact of articles published within (and outside) a given subject area. This CNCI score, the detailed calculation of which is shown in the methodology section, is an impact factor where the citation counts are normalized based on the world average value of 1.00. It compares the times cited for an item with the expected (average) number of citations for the same publication type, publication year and subject area or other items in the journal. It is calculated by dividing the actual number of citations by the expected citation rate. The fact that this score is embedded in InCites (the citation data is updated autonomously with a one-month delay, compared to WOS) is another positive aspect. Considering the unifying aspect and depth of CNCI, whether it should be evaluated as a single criterion, or a composite criterion is a separate topic of discussion. Because it focuses on measuring the field-based efficiency or impact of citations rather than the number of citations of the researcher. It is understood that it is a very unique and robust criterion in terms of revealing the real performance, effectiveness and efficiency of the researcher in his/her field (Potter et al., 2024).

The disadvantage of the CNCI metric is that for a researcher, this score can be excessive and misleading due to the high number of citations that come with a small number of publications. When it comes to publications by a single person, CNCI values can be inflated due to a single highly cited article. InCites also accepts this situation and warns researchers. An important warning here is that for this metric, which is based on the WOS citation system, to be meaningful, the user must take at least a two-year period as a basis for success. For example, a one-year CNCI score may not be very meaningful in terms of the impact of the citation, which is confirmed by the statements on the institutional InCites website (incites.help.clarivate.com). At this point, we can make a small suggestion: It would be fairer to add the number of WOS scientific publications to the CNCI score by standardizing it to a certain extent. Because it is likely and common for someone who publishes many WOS publications to have a low CNCI score compared to someone who only produces one article. Another point is that the CNCI score is multifaceted. CNCI value can be evaluated on a per-person, country, institution and article basis. On the other hand, our final suggestion to researchers is that Scopus publication and citation counts, along with WOS citation counts, and Google h-index data (albeit with a different weighting coefficient) can be added to the integrated researcher performance together with CNCI.

We can say that the 'AD Scientific Index' (AD Scientific Index, n.d.), another research performance ranking metric for scientists, is gradually approaching the publication of the rankings of almost all scientists in a very comprehensive manner. In this respect, it can be said that they have achieved a commendable scope. However, its most important shortcoming is that it only bases itself on Google H-index data equally. The AD Scientific systematic for rankings is based on three main criteria: H-index, i10 index and number of citations. It is known that the H index measures research productivity at the researcher level (calculated by determining the number of H publications cited at least H times), while the i10 index measures the number of academic publications produced by the researcher and received at least ten citations. The third indicator adopted by this ranking is the total number of citations. These indicators were developed by Google Scholar (Al-Hagree et al., 2023).

Indeed, Web of Science and Scopus are often considered the largest sources of information (bibliographic database) in today's academic world. In this respect, performance criteria such as the researcher's WOS-Q1 publication count, WOS citation count, and Scopus publication and citation count are important in distinguishing a quality researcher. We can easily understand this from the weight coefficient and criteria information of the rating agencies that rank the world's best universities. On the other hand, the widespread belief and demand that research results be published only in journals indexed in WoS and Scopus, and the fact that careers and salaries are often dependent on such publications, can, unfortunately, shift the scientific focus to the quantity of publications rather than quality (Pranckutė, 2021; Birkle, 2020). While databases such as the Web of Science- Science Citation Index (SCI) rank journals according to a hierarchy of prestige, sites such as Scopus, Orchid, and Google Scholar have continued to develop to count citations and compare the work of an academic with others (Hyland, 2023).

When the views mentioned above are considered, when past studies on scientific indicators that are taken as a basis for institutional and individual academic productivity performance on a global scale are scanned, we see that there are both positive and negative criticisms. When we gather some evaluations on CNCI metric, WOS citation, and document, InCites, Scopus citation and document and h-index, which have become quite popular in recent years, it is clearly understood that no metric, citation counting, and document archiving system is sufficient on its own. For example, when only WOS data is used, the low number of WOS journals (and therefore the number of citations) in some fields can be seen as a disadvantage. In addition, there are many relatively prestigious journals that are not scanned in WOS but are scanned in Scopus. In fact, some authors with very high WOS, Scopus citation, and document counts may have quality work in journals scanned in other indexes. Again, there are authors who have no articles in WOS and Scopus-based journals but often have quality articles in other indexes. On the other hand, the WOS/InCites-based CNCI and Scopus-based FWCI, which normalize the number of citations of an individual with the field-based expected citation averages, are also quite refined and insightful indicators. However, it will not be enough to be content with only these two indicators. Because there are weaknesses such as uncertainty about how to evaluate the uncertainty in the time dimension of citations and factors such as the number of documents showing the scores low or showing them excessively high for these metrics. Therefore, there is a gap in the literature regarding both the measurement of individual academic performance and the combination of indicators from different databases. While performance measurement is actually done to know, understand, and evaluate, it is clear that an incorrect or inadequate measurement will lead to unnecessary prejudices and wrong decisions.

This study uses the MCDM methodology to include and combine existing popular indicators and to objectively and precisely determine the weight coefficient of each indicator using the Entropy method. Although such a comprehensive and relatively fair performance support system for individual academic productivity has been indirectly suggested in the literature before, it has not been put into practice.

3. Material and Methods

The data in this study were obtained manually from "https://akademik.yok.gov.tr/AkademikArama/", an open-access source address whose infrastructure is formed by the Yüksek Öğretim Kurumu (Turkish Council of Higher Education) (YÖK Akademik, 2024), which includes the personal pages of all academics in the country where they create their CVs and then transferred to Excel tables. The aim of this study is to draw a valuable decision support system framework and provide insight to those who want to create a global performance measurement metric with artificial intelligence in the future by measuring individual academic performance with MCDM methods. In this study, the productivity performance data of 36 researchers in the field of Quantitative Decision Making Methods in Turkey (who mainly have previous studies on MCDM) whose data are accessible were selected.

The indicators were determined by considering the factors of prestige, prevalence, and diversity as a requirement for choosing the right and appropriate. In order to determine the number of prestigious articles of academicians, some of the indicators (SSCI/SCI, WOS Q1 article, ESCI, Scopus indexed articles) were evaluated separately because their impact and importance were different. For the number of article citations, the number of Google Scholar citations, which include the WOS and Scopus indexed scanning systems and other indexed data other than these, was also included in the evaluation (with a small weight coefficient). On the other hand, we also took the WOS-based CNCI and Scopus-based FWCI scores, which are updatable, popular, and dynamic metrics focusing on the impact value of the researcher's article in the field, which have been discovered in recent years, as a basis for the measurement. CRADIS, a relatively new method from the MCDM methodology family, was selected to obtain the final ranking of each candidate. In addition, TOPSIS, the most widely used method from different schools, and FUCA, a practical and easy method of the outranking school, were used in the comparison. Finally, a new and different sensitivity analysis of recent times was applied for testing the robustness of MCDM methods. The WOS, InCites, Scopus, and Google Scholar profiles of academics were carefully examined, and the criterion values of the alternatives were obtained. Table 1 below shows the methods and indicator criteria used in academic performance research.

Table 1. MCDM Methods, Weighting Technique, Criteria and Robustness Criteria for MCDM Methods Used in This Study

Weighting method	MCDA methods	Criteria	Robustness criterion for MCDM methods
Entropy	CRADIS, TOPSIS and FUCA	SSCI/SCI, q1 article, ESCI and/or SCOPUS article, CNCI impact score, FWCI impact score, WOS citation, Scopus citation, Google Scholar citation	Sensitivity analysis based on normalization

3.1. Decision Criteria

It is common to analyze research performance using descriptive statistical methods, but the problem in this study is different: Selecting, classifying and ranking researchers in the most appropriate, accurate and fair way based on multiple criteria, multiple alternatives and different weight coefficients is a problem here. For this reason, it is more logical to examine research performance using multi-criteria decision making (MCDM) methods, which are included in the field of operations research, rather than statistics.

The ideal goal for researchers is to be able to draw an effective, efficient, universal, valid and common academic performance framework with a correct route. In this sense, different agreed performance criteria should be evaluated simultaneously. However, there is no clear set of indicators agreed upon regarding research performance criteria. However, this situation should not result in the solution of the problem in question being suspended. In fact, it is noticeable that research performance is generally analyzed in two dimensions as a common opinion. The first is productivity, which depends on publication outputs (number). The second is impact (citation), which is an indicator of the quality of publications (Maral, 2024). This situation can make our job easier in choosing indicators. Here, it is noteworthy that there are many types of these two types of indicators. The quality, importance and function of these types address quite different needs and dimensions. Moreover, the degree of agreement on these types of indicators is also relative. While each ranking organization uses a few indicators, we do not come across studies and organizations that use others comprehensively. It is clear that the rankings are based on some indicators in a limited database. It can be said that this study aims to evaluate this literature and application gap. Table-2 shows the description and databases (data sources) of the criteria preferred for this study.

Table 2. Academic Metrics Used in This Study

	Criteria	Definitions	References databases
1.	SSCI (social sciences citation index) and/or SCI (science citation index-expanded) article counts	It refers to the number of SCI/SSCI-based articles of each candidate.	https://incites.clarivate.com/#/analysis/0/person*
2.	Q1 (quarter) article counts	It refers to the total number of Q1 articles scanned in SCI/SSCI for each candidate.	https://incites.clarivate.com/#/analysis/0/person*
3.	ESCI or SCOPUS article counts	It refers to the number of ESCI and Scopus-level articles of each candidate.	https://incites.clarivate.com/#/analysis/0/person* https://www.scopus.com/search
4.	WOS citation counts	Refers to the number of citations each candidate receives.	https://incites.clarivate.com/#/analysis/0/person*
5.	Scopus citation counts	It refers to the number of citations for each candidate according to the Scopus database.	https://www.scopus.com/search
6.	Google Scholar citation counts	It refers to the total number of citations for each candidate according to the Google Scholar database.	https://scholar.google.com/
7.	CNCI Score	It expresses the field-based impact value of each candidate's articles.	https://incites.clarivate.com/#/analysis/0/person*
8.	FWCI Score	It expresses the field-based impact value of each candidate's articles.	https://www.scopus.com/search

Source: incites.clarivate (2024); Scopus (2024); scholar.google (2024).

***Note:** Data used in this analysis is sourced from Clarivate Analytics' InCites Benchmarking & Analytics platform and requires special access.

3.2. Criterion Weighting Method: Entropy

In subjective evaluations regarding the weight coefficient of indicators, the discussion usually does not result in consensus due to the different perspectives of the parties. In this sense, when the parties who are undecided in determining the importance level of the criteria or those who lack sufficient information on the basis of the criteria turn to objective methods, the Entropy method comes up and this can be a suitable alternative. Information entropy, which is a measure of uncertainty, was first put forward by Shannon (1948). Moreover, Shannon information entropy, which is shown by Stewart (2012), as one of the 17 equations that changed the world in one way, was also suggested for determining the weight coefficient of the criteria in MCDM. According to the idea of information entropy, which is widely used in many fields, entropy can be used to measure the amount of useful information provided by the data itself. In other words, it can be said that it is an approach that determines the importance level for each column data series (criteria). If we continue the same path, it can be said that the Entropy weighting method is actually the suggested determination of the amount of objective information about the criteria in the decision matrix. Accordingly, the system works as follows: The smaller the entropy value, the larger the entropy-based weighting coefficient, thus the more information a particular criterion provides (Li et al., 2011; Wu et al., 2011). The basic steps are (Wang et al., 2020):

Step 1. Transform the data by applying sum normalization, then create the initial decision matrix with *m* rows (alternatives) and *n* columns (criteria):

$$F_{ij} = \frac{f_{ij}}{\sum_{k=1}^{m} f_{kj}} \quad i \in \{1,2,\ldots,m\};\ j \in \{1,2,\ldots,n\} \tag{1}$$

Step 2. Entropy is calculated for each criterion column.

$$E_j = -\frac{1}{\ln(m)} \sum_{i=1}^{m} (F_{ij} \ln F_{ij}) \quad j \in \{1,2,\ldots,n\} \tag{2}$$

Step 3. The weight coefficient of each criterion is determined.

$$w_j = \frac{1-E_j}{\sum_{j=1}^{n}(1-E_j)} \quad j \in \{1,2,\ldots,n\} \tag{3}$$

3.3. MCDM Methods

The MCDM methods used in this study are of three types from different schools. The first method is CRADIS, which is the basis of this study and is relatively new and quite efficient (Baydaş et al., 2023). The second is the most widely used method and the third is the simplest and most practical method from the outranking school. In other words, 3 different methods, which are common, popular, and new, were preferred. The studies based on the MCDM methods and weighting techniques preferred for this study are shown in Table 3 below.

Table 3. References to Relevant MCDM Formulations and Weighting Techniques Used in This Study

MCDM Methods	References
CRADIS	Puška et al., 2022; Puška and Stojanović, 2022
FUCA	Wang ve Rangaiah 2017
TOPSIS	Chakraborty, 2022.

3.3.1. CRADIS (Compromise ranking of alternatives from distance to ideal solution) method

The distinctive nuance of the CRADIS method is related to determining the deviations of alternatives from the ideal and anti-ideal solutions. In fact, it can be said that this method is a combination of different steps taken from ARAS, MARCOS, and TOPSIS methods. The CRADIS method is a relatively new approach that produces a unique combination by using these existing methods. In this method, alternatives are basically observed according to general criteria, and ideal solutions representing the maximum value of the ideal solution of the alternative and the minimum value of the alternative are used (Puška et al., 2022; Puška and Stojanović, 2022).

The CRADIS method will be used to rank the materials. The steps of the CRADIS method are explained below (Puška et al., 2022):

Step 1: The first decision matrix is created.

Step 2: Normalization of decision matrix.

$$u_{ij} = \frac{t_{ij}}{max\,(t_{ij})} \tag{4}$$

$$u_{ij} = \frac{min\,(t_{ij})}{t_{ij}} \tag{5}$$

Step 3: The aggravated decision matrices.

$$s_{ij} = u_{ij}.w_{jIN} \tag{6}$$

Step 4: The ideal and anti-ideal solutions are determined with Equations 7 and 8.

$$v_i = maxs_{ij} \tag{7}$$

$$v_{ai} = mins_{ij} \tag{8}$$

Step 5: Deviations from ideal and anti-ideal solutions are computed by Equations 9 and 10.

$$d^+ = max\ v_i - s_{ij} \tag{9}$$

$$d^- = s_{ij} - min\ v_{ai} \tag{10}$$

Step 6: The grades of the deviations for each alternative from anti-ideal and ideal solutions are computed as.

$$o_i^+ = \sum_{j=1}^{n} d^+ \tag{11}$$

$$o_i^- = \sum_{j=1}^{n} d^- \tag{12}$$

Step 7: The utility function for each alternative pertaining to the deviations from the optimal alternatives is computed as.

$$K_i^+ = \frac{o_{opt}^+}{o_i^+} \tag{13}$$

$$K_i^- = \frac{o_{opt}^-}{o_i^-} \tag{14}$$

In Equation 13, o_{opt}^+ denotes the optimal alternative having the smallest distance from the ideal solution. In Equation 14, o_{opt}^- denotes the optimal alternative having the greatest distance from the anti-ideal solution.

Step 8: Ranking alternatives. The average deviation value (Q_i) for each alternative is computed as.

$$Q_i = \frac{K_i^+ + K_i^-}{2} \tag{15}$$

The alternative with the highest Q_i is determined as the best alternative.

3.3.2. TOPSIS (The technique of reaching the order of preference with similarity to the ideal solution)

In this method, the best alternative selected is the point closest to the positive ideal solution and the farthest from the negative ideal solution. This method, which is distance-based and attaches great importance to the ideal values in the criteria, can be called the most adopted type of MCDM (Chakraborty, 2022).

3.3.3. FUCA (Faire un choix adéquat)

The most practical and simple method is the MCDM method. Since it uses the rank value, this method does not require normalization. It is based on a ranking value of the alternatives for each criterion. The first row has the best value (one:1), while the last row (n) is assigned the worst value. Then, the weighted sum of the values is calculated for each solution point and the solution with the smallest total value is the best-selected solution (Wang and Rangaiah, 2017).

4. Application

In the application part of the study, scientific indicators (e.g., number of articles, number of citations, impact factor) obtained from various databases (including Web of Science, Scopus, InCites, and Google Scholar) were used to measure global academic research performance in a multidimensional and realistic way. In this context, the analysis steps of the study can be seen in detail below:

Step 1: Creation of Initial Decision Matrix with Academic Performance Values

In this step, the basic criteria to be used for a comprehensive evaluation of academic performance were selected. In the study, performance indicators that are important in scientific performance literature and frequently used in databases were preferred. The initial decision matrix was created to analyze these determined criteria and their values with MCDM (Multi-Criteria Decision Making) methods on 36 researchers in the field of quantitative decision-making. The performance analysis was made more holistic by considering other indicators such as the number of publications, the number of citations, as well as the field normalized impact factor among the criteria. The weighting coefficient assignment process was carried out using the Entropy method, which is an objective approach, and thus the objective weights of the criteria were calculated.

Step 2: Determining Ranking Results with MCDM Methods

In this step, the created decision matrix was analyzed using methods from three different MCDM schools. These methods used in the study can reach different ranking results because the process steps are different. The MCDM methods selected for analysis provide a more in-depth perspective beyond classical ranking methodologies by technically addressing researcher performance in a multifaceted manner. The ranking results obtained with these methods provide more reliable and valid results compared to rankings made with other single and limited criteria available in the literature.

Step 3: Sensitivity Analysis for Academic Performance Analysis

The Entropy weighting method preferred in the study was used to ensure the consistency of the weights in the decision matrix. In addition, a comprehensive sensitivity analysis was performed to examine the reliability and stability of the ranking results. In the sensitivity analysis, the results obtained with different weighting methods were compared and the reliability and stability of the MCDM methods were confirmed.

Final Step: Obtaining Performance Results

To make the findings of this study more understandable, the performance of each researcher was scored according to the specified criteria. In this way, academic performance findings were transformed into more useful and concrete information for university administrations and other stakeholders. This step aims to increase the effectiveness of the decision support system and provide clearer insights into the current status of researchers.

This comprehensive methodology provides an integrated and innovative model that will contribute to the literature in measuring the individual and institutional performance of researchers. The framework developed in the study is a scientific reference for future ranking systems and will contribute to more accurate and comprehensive decisions in academic performance evaluation processes. The table below shows the weight coefficients determined according to the Entropy method.

Table 4. Weight Coefficients Determined for Each Criterion Column

Criteria	Criterion Direction (Max or Min)	Weightage used (Entropy)
SSCI/SCI articles	Max	0.082169
Q1 articles	Max	0.176963
Score of CNCI	Max	0.116449
WOS citations	Max	0.154856
ESCI-SCOPUS articles	Max	0.085446
FWCI Score	Max	0.181039
Scopus citations	Max	0.140755
Google Scholar citations	Max	0.062323

According to the Entropy results in the table above, it is reasonable that the criteria with the highest weight coefficient are Q1 article, FWCI and WOS citations because they are also important in the literature. Again, it seems appropriate that the criterion with the lowest weight is the Google Scholar citation count because this criterion is a relatively heterogeneous criterion since it evaluates citations from all types of publications.

The normalization techniques used in this study are Max, Min-Max, Vector, and Sum normalization techniques (Wang et al., 2020). The formulas related to these are included in Appendix-1 at the end of the study. The table below shows the ranking results obtained according to the CRADIS method depending on different normalization types.

Table 5. Ranking Results of the CRADIS Method

	Max based	Rank	Min-Max based	Rank	Vector based	Rank	Sum based	Rank
A1	0.420485	5	0.604611	5	0.476263	5	0.503875	5
A2	0.540241	2	0.694738	2	0.579056	3	0.605236	3
A3	0.299445	9	0.512829	9	0.343623	9	0.368962	9
A4	0.264291	12	0.486098	12	0.325193	12	0.358669	11
A5	0.155326	36	0.403355	36	0.214838	36	0.248662	36
A6	0.301831	8	0.514699	8	0.351508	8	0.377417	8
A7	0.179133	29	0.421404	29	0.237666	27	0.269007	27
A8	0.206241	20	0.441942	20	0.250837	24	0.276141	25
A9	0.15991	34	0.406863	34	0.217986	34	0.25094	34
A10	0.181272	27	0.422948	27	0.242685	26	0.274392	26
A11	0.16794	32	0.412914	32	0.223365	32	0.25505	32
A12	0.272288	11	0.491905	11	0.333757	10	0.366923	10
A13	0.394646	6	0.585199	6	0.435843	6	0.456375	6
A14	0.465108	4	0.638164	4	0.523802	4	0.553184	4
A15	0.232086	18	0.461429	18	0.282622	18	0.308381	18
A16	0.2541	15	0.478157	15	0.304093	15	0.331841	14
A17	0.541567	1	0.695495	1	0.591216	2	0.619863	2
A18	0.239704	17	0.467211	17	0.297588	17	0.329638	16
A19	0.256086	14	0.479584	14	0.304924	14	0.331095	15
A20	0.158626	35	0.40588	35	0.217104	35	0.250302	35
A21	0.259524	13	0.482276	13	0.315344	13	0.343868	13
A22	0.198226	23	0.435931	23	0.251808	23	0.282848	24
A23	0.203409	22	0.439669	22	0.260399	21	0.291672	21
A24	0.17926	28	0.421476	28	0.231446	30	0.261456	30
A25	0.175419	31	0.418545	31	0.235969	28	0.266931	28
A26	0.275487	10	0.494423	10	0.329062	11	0.357907	12
A27	0.183765	26	0.424909	26	0.235562	29	0.26543	29
A28	0.303328	7	0.515545	7	0.366987	7	0.398665	7
A29	0.197197	24	0.435064	24	0.255193	22	0.285938	22
A30	0.177485	30	0.420124	30	0.229913	31	0.260112	31
A31	0.186426	25	0.426839	25	0.250271	25	0.284937	23
A32	0.160029	33	0.406917	33	0.219465	33	0.252385	33
A33	0.245913	16	0.472017	16	0.298954	16	0.326993	17
A34	0.206115	21	0.441788	21	0.265265	20	0.297111	19
A35	0.214719	19	0.448306	19	0.266214	19	0.295756	20
A36	0.515005	3	0.675603	3	0.598529	1	0.640643	1

The table above shows that the normalization type can change the best alternative. In fact, Max and Min-Max normalizations put the A17 alternative first, while Vector and Sum put the A36 alternative first. When the standard deviation of the final scores of the alternatives is observed, it is understood that the Min Max method has the lowest deviation. In Table 6 below, the effect of the normalization type on the final results and ranking distribution is better understood from the standard deviation values.

Table 6. Standard Deviation Values of the Final Scores of CRADIS Rankings According to the Normalization Type

	Maximum	**Min-Max**	**Vector**	**Sum**
StDv	0.109138	*0.082668*	0.108805	0.10848

The distribution flexibility or drift level of CRADIS final scores is better seen in Figure 1 below. It is understood once again that CRADIS results based on Min-Max normalization are more stable.

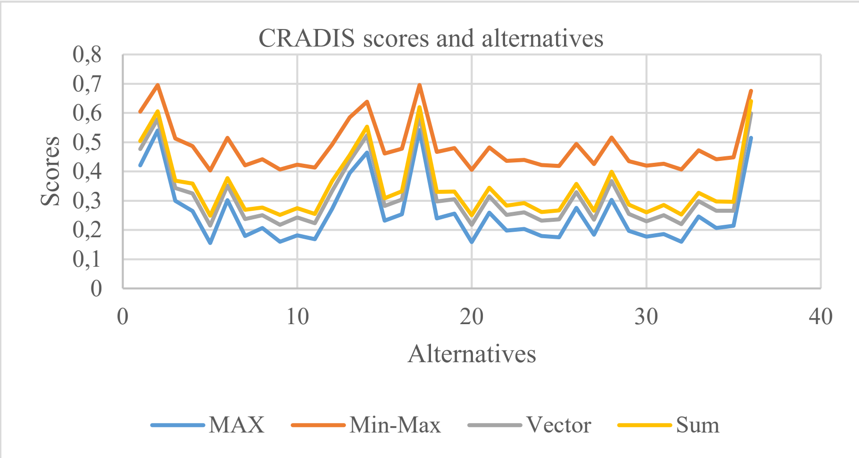

Figure 1. Effect of Normalization Techniques on the Sensitivity of CRADIS Scores

After solving the problem of normalization selection within CRADIS, we compare the results obtained with the min-max method with TOPSIS and FUCA. The final ranking results of three different MCDM methods are compared in Table 7 below.

Table 7. Final Ranking Results of MCDM Methods

Rank	*CRADIS*	*FUCA*	*TOPSIS*	***Rank***	*CRADIS*	*FUCA*	*TOPSIS*
1	*A17*	*A2*	*A36*	***19***	A35	A15	A35
2	*A2*	*A36*	*A17*	***20***	A8	A35	A23
3	*A36*	*A17*	*A2*	***21***	A34	A31	A8
4	A14	A13	A14	***22***	A23	A29	A22
5	A1	A14	A1	***23***	A22	A23	A34
6	A13	A1	A13	***24***	A29	A25	A29
7	A28	A28	A28	***25***	A31	A22	A31
8	A6	A21	A4	***26***	A27	A27	A10
9	A3	A26	A12	***27***	A10	A10	A25
10	A26	A33	A6	***28***	A24	A30	A27
11	A12	A6	A3	***29***	A7	A32	A24
12	A4	A16	A26	***30***	A30	A24	A30
13	A21	A12	A19	***31***	A25	A7	A7
14	A19	A3	A18	***32***	A11	A8	A11
15	A16	A18	A21	***33***	A32	A20	A32
16	A33	A34	A16	***34***	A9	A11	A9
17	A18	A4	A33	***35***	A20	A9	A20
18	A15	A19	A15	***36***	A5	A5	A5

According to this table, it is understood that MCDM methods generally agree but reach different results in choosing the best alternative. However, it is seen that this mostly affects or changes the struggle in the first, second or third place. For the last place, it is understood that there is full agreement in the methods on the A5 alternative.

In the following Table, the relationships between MCDM methods were investigated to examine the ranking consistency. According to the results, CRADIS obtained closer results with TOPSIS.

Table 8. Spearman Correlation Results between MCMD Methods. (Spearman rho is shown in the first row and p-value is shown in the second row)

	CRADIS	*FUCA*
FUCA	*0.942*	
	0.000	
TOPSIS	*0.988*	*0.933*
	0.000	*0.000*

5. Discussion

The results of the study also reveal the accuracy of the Entropy method in weighting the criteria used in academic performance evaluations. According to the Entropy results, the highest weight coefficients of Q1 article count, FWCI (Field-Weighted Citation Impact), and Web of Science (WoS) citation count support the importance of these indicators in the literature. Features such as high-quality publications and field-normalization contribute significantly to the ranking process as they reflect the impact of the researcher more accurately. The fact that the Google Scholar citation count has the lowest weight is an expected finding due to the heterogeneous structure of this criterion and the fact that it covers all types of publications. These results show that not only numerical values but also the quality and focus level of the criteria should be taken into account in academic performance evaluations.

The ranking results obtained by the CRADIS method according to different normalization types clearly reveal the effect of the normalization selection on the final rankings. When the Max, Min-Max, Vector, and Sum normalization techniques are compared, it is seen that the best-performing alternative varies according to the normalization type. The fact that Min-Max normalization has the lowest standard deviation reveals that this technique is more successful in maintaining ranking stability. This situation reveals that the normalization method used in performance evaluation processes is a critical parameter in terms of sensitivity to the evaluation results. The stable results provided by Min-Max normalization are important. Thus, according to the sensitivity findings in this study, it can be said that the min-max modified CRADIS method is recommended in this study instead of the classical max normalization suggested by the authors.

Comparison of the Min Max-based CRADIS results with the other MCDM methods TOPSIS and FUCA shows that although there is a general consistency between the methods, there are differences in determining the best alternative. When the ranking correlation between the three methods is evaluated, it is seen that the CRADIS and TOPSIS methods produce results closer to each other. The high agreement obtained in the Spearman correlation analysis reveals that CRADIS and TOPSIS handle the ranking criteria similarly and can be considered as methods that support each other. However, the differences shown by the methods in determining the best alternative increase the sensitivity of the ranking results, especially in researcher groups with high competition in the first ranks.

Table 9. Final Results for Best Academics

	CRADIS	FUCA	TOPSIS
BEST	A17	A2	A36
2.	A2	A36	A17
3.	A36	A17	A2

6. Conclusion

This study suggests how to find the right route in creating global individual academic performance rankings. It provides a holistic approach to academic performance evaluations by examining the known benefits and limitations of multi-criteria decision-making (MCDM) methods. Weighting analyses using the entropy method ensured that academic performance criteria (e.g. Q1 article, CNCI and FWCI score, WoS citation, etc.), which are also important in the literature, were included in the ranking more effectively. The fact that the CRADIS method produces stable ranking results under different normalization techniques, especially because Min-Max normalization provides the lowest deviation, shows that this method can be used as a reliable tool in sensitive ranking analyses.

The results obtained in the study reveal that MCDM methods can be used as an alternative decision-support tool in academic performance rankings. The high correlation between CRADIS and TOPSIS shows that the accuracy of the ranking results can be increased with these two methods. However, the differences in determining the best alternative of the methods indicate that the ranking struggles among high-performing researchers should be analyzed in more detail.

In conclusion, this study fills an important gap in the literature with its proposed framework for a more accurate and comprehensive measurement of individual academic performance. In particular, multiple normalization techniques used with the CRADIS method offer a different approach to ranking systems and contribute to a more fair and reliable evaluation of academic performance. According to the sensitivity findings in this study, it can be said that instead of the classical max normalization suggested by the authors, the min-max modified CRADIS method is recommended in this study.

In addition, since global rankings consist of large data, autonomous ranking generators that can avoid the complex calculation steps of MCDM methods can also use FUCA, a simple and practical method used in this study, as an alternative. On the other hand, the dynamicity of the framework in this study is intended to be that this framework consists of updatable indicators. In other words, AI generators that adopt such an integrated framework can automatically update the ranking performance of researchers on a year, month, week, and even day basis. For example, Stanford announces the top 2 percent of scientists in the world every October. This proposed rapid framework can be used as a valuable guide for academic ranking organizations and university administrations in their strategic decision-making processes.

Recommendations for Future Researchers

In future research, larger databases and alternative measures can be integrated to strengthen academic performance evaluations. In addition to methods such as entropy, different subjective weighting techniques based on expertise such as AHP and BWM should be tried. Field-specific normalization techniques should be used to take into account the field-based variability of citation rates and long-term performance changes of researchers should be monitored on an annual basis. Developing hybrid MCDM models can strengthen ranking accuracy. Creating a real-time ranking system with dynamic data updates will provide more flexible evaluations. Dynamic, integrated and global performance rankings supported by artificial intelligence can be obtained with indicator values drawn from each database.

Limitations of the Study

This study was conducted with limited data and methods. The normalization techniques and selected indicator types are relatively limited, and therefore the findings should be interpreted by taking this into account.

Availability of data and material: Not applicable.

Competing interests: The authors declare that they have no competing interests.

Funding: Not applicable.

Authors' contributions: All author(s) read and approved the final manuscript.

It is declared that this study is not in a study group that requires ethics committee approval.

References

AD Scientific Index. (2024, September 15). *World Scientist and University Rankings, September 2024*. https://www.adscientificindex.com/

Al-Hagree, S., Mohsen, A. A., Abdulrazzak, F. H., Al-Sanabani, M., Alalayah, K. M., Al-Gaphari, G., ... and Gawbah, H. (2023, 10-11 October). Universities the best performers: a cluster analysis of yemeni universities rankings. *In 2023 3rd International Conference on Emerging Smart Technologies and Applications* (p. 1-9). Taiz, Yemen. https://doi.org/10.1109/eSmarTA59349.2023.10293385

Ardil, C. (2021). Scholar index for research performance evaluation using multiple criteria decision making analysis. *International Journal of Educational and Pedagogical Sciences*, *13*(2), 93-104.

Baydaş, M., Elma, O. E., and Stević, Ž. (2024). Proposal of an innovative MCDA evaluation methodology: knowledge discovery through rank reversal, standard deviation, and relationship with stock return. *Financial Innovation*, *10*(1), 1-35. https://doi.org/10.1186/s40854-023-00526-x

Baydaş, M., Eren, T., Stević, Ž., Starčević, V., and Parlakkaya, R. (2023). Proposal for an objective binary benchmarking framework that validates each other for comparing MCDM methods through data analytics. *PeerJ Computer Science*, *9*,1-24. https://doi.org/10.7717/peerj-cs.1350

Birkle, C., Pendlebury, D. A., Schnell, J., and Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363-376. https://doi.org/10.1162/qss_a_00018

Chakraborty, S. (2022). TOPSIS and Modified TOPSIS: A comparative analysis. *Decision Analytics Journal*, 2(2022), 1-7. https://doi.org/10.1016/j.dajour.2021.100021

Clarivate analytics (Incites). (2024, September 15). *Incites Analysis Person*, *September 2024*. https://incites.clarivate.com/#/analysis/0/person

Clarivate™. (2024, September 15). *Clarivate Highly Cited Researchers, September 2024*. https://clarivate.com/highly-cited-researchers/

Clarivate™. (2024, September 15). *InCites Help Center*, *September 2024*. https://tinyurl.com/yn34arxx

Docampo, D., and Cram, L. (2019). Highly cited researchers: a moving target. *Scientometrics*, *118*(3), 1011-1025. https://doi.org/10.1007/s11192-018-2993-2

Google Academic. (2024, September 15). *Google Academic*, September 2024. https://scholar.google.com/

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., and Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429-431. https://www.nature.com/articles/520429a.pdf

Hyland, K. (2023). Academic publishing and the attention economy. *Journal of English for Academic Purposes*, *64*(2023), 1-10. https://doi.org/10.1016/j.jeap.2023.101253

Ioannidis, J. P. A. (2024). August 2024 data-update for "Updated science-wide author databases of standardized citation indicators", *Elsevier Data Repository*, V7. https://doi.org/10.17632/btchxktzyw.7

Ioannidis, J. P. A., Boyack, K. W., and Baas, J. (2020). Updated science-wide author databases of standardized citation indicators. *PLOS Biology*, *18*(10), 1-3. https://doi.org/10.1371/journal.pbio.3000918

Li, X., Wang, K., Liu, L., Xin, J., Yang, H., and Gao, C. (2011). Application of the entropy weight and TOPSIS method in safety evaluation of coal mines. *Procedia engineering*, *26*(2011), 2085-2091. https://doi.org/10.1016/j.proeng.2011.11.2410

Li, Z., Liechty, M., Xu, J., and Lev, B. (2014). A fuzzy multi-criteria group decision making method for individual research output evaluation with maximum consensus. *Knowledge-Based Systems*, *56*, 253-263. https://doi.org/10.1016/j.knosys.2013.11.018

Maral, M. (2024). Research performance of higher education in OECD countries: A hybrid multi-criteria decision-making approach. *SAGE Open*, *14*(2), 1-20. https://doi.org/10.1177/21582440241257752

Markusova, V.A., Libkind, A.N., Zolotova, A.V., and Kotelnikova, N.A. (2023). Priority areas of scientific cooperation between scientists of Russia, Iran, India, and Turkey: bibliometric analysis according to the InCites database (2011–2021). *Automatic Documentation and Mathematical Linguistics*, *57*(5), 274-283. https://doi.org/10.3103/S0005105523050047

Olcay, G. A., and Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting and Social Change*, *123*, 153-160. https://doi.org/10.1016/j.techfore.2016.03.029

Parker, C. (2000). Performance measurement, *Work Study*, *49*(2), 63-66. https://doi.org/10.1108/00438020010311197

Potter, R. W., Kovač, M., and Adams, J. (2024). Tracking changes in CNCI: the complementarity of standard, collaboration and fractional CNCI in understanding and evaluating research performance. *Scientometrics*, *129*, 1-14. https://doi.org/10.1007/s11192-024-05028-w

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, *9*(1), 1-59. https://doi.org/10.3390/publications9010012

Puška A, Stević Ž., and Pamučar D. (2022). Evaluation and selection of healthcare waste incinerators using extended sustainability criteria and multi-criteria analysis methods. *Environment, Development and Sustainability, 24*(9), 11195-11225. https://doi.org/10.1007/s10668-021-01902-2

Puška A., and Stojanović, I. 2022. Fuzzy multi-criteria analyses on green supplier selection in an agrifood company. *Journal of Intelligent Management Decision* *1*(1),2–16. https://doi.org/10.56578/jimd010102

Scopus (Elsevier). (2024, September 5). *Scopus Search*, *September 2024.* https://www.scopus.com/search/form.uri#basic

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Soh, K. (2017). The seven deadly sins of world university ranking: A summary from several papers. *Journal of Higher Education Policy and Management*, *39*(1), 104-115. http://dx.doi.org/10.1080/1360080X.2016.1254431

Stewart, I. (2012). In pursuit of the unknown: 17 equations that changed the world. https://doi.org/10.1007/s00283-012-9303-2

Szluka, P., Csajbók, E., and Győrffy, B. (2023). Relationship between bibliometric indicators and university ranking positions. *Scientific Reports*, *13*(1), 14193. https://doi.org/10.1038/s41598-023-35306-1

Tuan, N., Hue, T., Lien, L., Thao, T., Quyet, N., Van, L., and Anh, L. (2020). A new integrated MCDM approach for lecturers' research productivity evaluation. *Decision Science Letters*, *9*(3), 355-364. https://doi.org/10.5267/j.dsl.2020.5.001

URAP (2023, October 28). *2023-2024 URAP Türkiye sıralaması basın açıklaması, October 2023*. https://tinyurl.com/2ujta97z

Wang, Z., and Rangaiah, G.P. (2017). Application and analysis of methods for selecting an optimal solution from the Pareto-Optimal front obtained by multiobjective optimization. *Industrial & Engineering Chemistry Research, 56*(2), 560-574. https://doi.org/10.1021/acs.iecr.6b03453

Wang, Z., Parhi, S. S., Rangaiah, G. P., and Jana, A. K. (2020). Analysis of weighting and selection methods for pareto-optimal solutions of multiobjective optimization in chemical engineering applications. *Industrial & Engineering Chemistry Research, 59*(33), 14850-14867. https://doi.org/10.1021/acs.iecr.0c00969

Wu, J., Sun, J., Liang, L., and Zha, Y. (2011). Determination of weights for ultimate cross efficiency using Shannon entropy. *Expert Systems with Applications*, *38*(5), 5162-5165. https://doi.org/10.1016/j.eswa.2010.10.046

YÖK Akademik. (2024, September 23). *Yüksek Öğretim Akademik Arama*, *September, 2024*. https://akademik.yok.gov.tr/AkademikArama/

Appendix

The table below shows the formulas of the normalization methods used in this study.

Appendix-1. Normalization Methods**.**

Method	Calculations
Sum Normalization	$F_{ij} = \frac{f_{ij}}{\sum_{k=1}^{m} f_{kj}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}$
Vector Normalization	$F_{ij} = \frac{f_{ij}}{\sqrt{\sum_{k=1}^{m} {f_{kj}}^2}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}$
Min-Max Normalization	$F_{ij} = \frac{f_{ij} - min_{i\in m} f_{ij}}{max_{i\in m} f_{ij} - min_{i\in m} f_{ij}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}\ for\ benefit\ criteria$ $F_{ij} = \frac{max_{i\in m} f_{ij} - f_{ij}}{max_{i\in m} f_{ij} - min_{i\in m} f_{ij}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}\ for\ cost\ criteria$
Max Normalization	$F_{ij} = \frac{f_{ij}}{max_{i\in m} f_{ij}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}\ for\ benefit\ criteria$ $F_{ij} = \frac{min_{i\in m} f_{ij}}{f_{ij}} \quad i \in \{1,2,\dots,m\};\ j \in \{1,2,\dots,n\}\ for\ cost\ criteria$

Source: Wang et al. (2020).