




Article

Ranking the Performance of Universities: The Role of Sustainability

Christoph Burmann ¹, Fernando García ² , Francisco Guijarro ^{3,*}  and Javier Oliver ^{2,*} 

¹ Markstones Institute of Marketing, Branding & Technology, Universität Bremen, 28359 Bremen, Germany; c.burmann@uni-bremen.de

² Department of Economics and Social Sciences, Universitat Politècnica de València, 46022 Valencia, Spain; fergarga@esp.upv.es

³ Research Institute of Pure and Applied Mathematics, Universitat Politècnica de València, 46022 Valencia, Spain

* Correspondence: fraguima@upvnet.upv.es (F.G.); jaolmun@ade.upv.es (J.O.)

Abstract: University rankings assess the performance of universities in various fields and aggregate that performance into a single value. In this way, the aggregate performance of universities can be easily compared. The importance of rankings is evident, as they often guide the policy of Higher Education Institutions. The most prestigious multi-criteria rankings use indicators related to teaching and research. However, many stakeholders are now demanding a greater commitment to sustainable development from universities, and it is therefore necessary to include sustainability criteria in university rankings. The development of multi-criteria rankings is subject to numerous criticisms, including the subjectivity of the decision makers when assigning weights to the criteria. In this paper we propose a methodology based on goal programming that allows objective, transparent and reproducible weighting of the criteria. Moreover, it avoids the problems associated with the existence of correlated criteria. The methodology is applied to a sample of 718 universities, using 11 criteria obtained from two prestigious university rankings covering sustainability, teaching and research. A sensitivity analysis is carried out to assess the robustness of the results obtained. This analysis shows how the weights of the criteria and the universities' rank change depending on the λ parameter of the goal programming model, which is the only parameter set by the decision maker.

Keywords: ranking; higher education institutions; sustainability; goal programming; multicriteria decision-making



Citation: Burmann, C.; García, F.; Guijarro, F.; Oliver, J. Ranking the Performance of Universities: The Role of Sustainability. *Sustainability* **2021**, *13*, 13286. <https://doi.org/10.3390/su132313286>

Academic Editor:
Jesús Granados-Sánchez

Received: 1 October 2021

Accepted: 27 November 2021

Published: 30 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The process of economic globalization in recent decades has had an enormous impact on our societies and organizations. Higher Education Institutions (HEIs) have not been unaffected by this evolution and are now subject to increased international competition and social scrutiny. In this context, information is needed to enable stakeholders to assess and compare the performance of HEIs globally. One of the most popular instruments are university rankings, which allow for a simple and quick comparison of HEIs on the basis of selected variables [1,2]. The development of these multi-criteria rankings has experienced strong growth in recent years, which has made them an object of analysis by academia. The functions performed by university rankings are multiple and are aimed at meeting the information demands of different stakeholders. They serve to guide prospective students, assess the overall situation of universities, improve competition in the areas assessed in the rankings, project a good image of universities and improve the satisfaction of the university community [3,4]. They can also be used to aid decision makers and facilitate university policies and the allocation of financial resources [5,6].

The development of university rankings can be approached from different perspectives. In order to unify the procedures, the Berlin Principles on Ranking of Higher Education

Institutions were published [7]. These principles are rather generic, but they address important issues that need to be considered in the development of the rankings. Any ranking should clearly define the purpose and goals of the ranking, the design and weighting of the indicators used, the process to collect and process the data, and the way ranking results are presented.

Currently, most global university rankings assess the performance of HEIs in relation to two fundamental aspects: teaching and research [8]. The fact that the rankings value these two aspects of university activity is logical, as these have traditionally been the two main functions of HEIs. On the other hand, the main international rankings value the research aspect much more highly than the teaching aspect and, frequently, what is really being measured is the prestige of the universities [9,10].

Global university rankings have been criticized for a number of reasons [11–13]. One of the main reasons for criticism is to analyze only aspects related to research and teaching [14]. This point is especially relevant if we recognize the importance of rankings in providing information to stakeholders and as a force for promoting specific university policies [1]. In this sense, in a context in which concern about climate change is growing, universities must lead the process of change required by society [15–17]. Therefore, it seems reasonable to incorporate the measurement of environmental performance as a criterion in the elaboration of the rankings. The relationship between universities and the environment is manifold. Universities are like small cities, whose management has repercussions on aspects such as the transport of thousands of students and employees, energy consumption and waste management [18]. They can be an example of environmental management for other public administrations and companies. Moreover, they can promote sustainability culture through numerous actions such as the inclusion of sustainability in curricula, or the promotion of research and transfer of environmental issues [17,19,20].

In addition to analyzing only research and teaching, the global university rankings are criticized for the methodology used, especially the weighting of the criteria [21,22]. Generally, this process is very subjective and the methodology does not make explicit who has decided the weighting of each of the criteria in the final weighting or how it has been calculated [23], so the results obtained are not reproducible [24]. As a result, the rankings differ substantially in their orderings, although the top places are often occupied by the same universities [25]. This lack of transparency undermines the credibility of the rankings and limits their effectiveness in achieving the purposes they are intended to serve.

In the light of the above criticism, the aim of this paper is to present a methodology to develop university rankings by applying goal-based programming that includes both traditional criteria related to research and teaching, as well as sustainability criteria. This multi-criteria methodology allows for a transparent and objective weighting of the different criteria [26] and at the same time is easily reproducible [27]. In this way, this paper contributes to the growing literature on sustainable university management and the development of HEI rankings.

The remainder of the paper is structured as follows. Section 2 presents the literature review related to the assessment of the environmental performance of universities. Section 3 describes the criteria used for the elaboration of HEI rankings under the criteria of research, teaching and sustainability. Section 4 is devoted to the presentation of the methodology to elaborate the multi-criteria ranking based on goal programming (GP). Section 5 describes the database used in the elaboration of the rankings. Section 6 analyses the rankings obtained and finally, Section 7 presents the main conclusions of the work.

2. Assessment of the Environmental Performance of Universities

The inclusion of sustainability performance criteria in the HEI rankings can be a determining factor as a catalyst for action. It allows for measuring progress in promoting sustainability in different aspects, increases transparency and is a means for universities to communicate their commitment to environmental goals. The importance of HEIs as promoters of sustainable development was already highlighted in the Declaration on

the Human Environment in 1972. Since then, numerous policy statements, charters and declarations have been issued dealing with HEIs sustainability. Among the latest examples we can mention the United Nations Higher Education Sustainability Initiative (HESI), People's Sustainability Treaty on Higher Education, Copernicus Charta 2.0, and the G8 University Summit: Statement of Action [28]. The aim of these documents is to promote the commitment of universities to the goals of sustainable development and to facilitate the process of integrating sustainable development into the different activities carried out by HEIs [2,29]. The aim is not only to reduce the environmental impact of universities as operating institutions, but also to turn them into promoters of social change. In this context, universities must introduce sustainable management in aspects such as infrastructure management, energy consumption, waste treatment and water consumption. They must also consider indirect aspects, such as the transport used by students and staff. Furthermore, universities should encourage research and teaching in the area of sustainability, raise awareness among students and staff of the importance of sustainable practices and lead the change towards a more sustainable society [30,31]

Despite the fact that these declarations contain important guidelines to guide the action of universities in achieving sustainable development and fostering a more sustainable society, none of them is useful at an operational level, i.e., there are no precise instructions on exactly how universities should act in each of the different areas involved in sustainable development. In response to this need, numerous assessment tools have emerged, especially in the last two decades. Some authors identified three categories of assessment tools based on their approaches: accounts assessment, narrative assessment and indicator-based assessment [32]. After the analysis of the strengths and weaknesses of the different approaches, the author concluded that indicator-based assessments offer higher levels of transparency, consistency and usefulness for decision-making. Moreover, indicator-based assessments have an overall higher performance and are more easily measurable and comparable than the other two approaches. It is therefore not surprising that in recent years numerous proposals for sustainability assessment tools for HEIs using the indicator-based approach have emerged. The main proposals have been compared and analyzed by different studies [15,18,19,28]. It is worth noting that there are notable differences between them in terms of purpose, scope and function. Moreover, assessment tools vary also regarding the weighting methods for indicators, flexibility and access to information [28,33]. Some of the sustainability assessment tools in HEIs have been proposed by researchers, for example the Adaptable Model for Assessing Sustainability in Higher Education [34], the Graphical Assessment of Sustainability in Universities GASU [35], the Graz Model for Integrative Development GMID [6,36], the Modifiable Campus-Wide Appraisal Model MOCAM [23], the Sustainable University Model SUM [37], the University Environmental Management System UEMS [30] or the Uncertainty-based quantitative assessment of sustainability for HEIs [17]. Other proposals have been made by universities, organizations and companies, among them the Assessment System for Sustainable Campus ASSC [38], the Sustainability Assessment Questionnaire SAQ [39], the Unit-Based Sustainability Assessment Tool USAT [40], or the Sustainability Leadership Scorecard [41].

Some of the methodological proposals allow for the elaboration of ranking tables, such as the Times Higher Education Impact University Ranking [42], People and Planet University League (P&P) [43], the Sustainability Tracking, Assessment and Rating System (STARS) [44] and the GreenMetric World University Ranking [45]. In this paper, we will use the data collected and used for the elaboration of the UI GreenMetric World University Ranking (GreenMetric). This international ranking assesses HEIs sustainability performance around the globe and is an initiative of Universitas Indonesia. According to Ragazzi and Ghidini [46], this ranking lays a good foundation for the incorporation of the principle of sustainability within the HEIs and reflects the need to quantify the efforts towards sustainability. Several authors have used this ranking in their research. Some authors evaluate the implementation and results of the ranking [47], others focus on conceptual issues surrounding the meaning of sustainability [48], others used this

ranking to quantify the contribution of universities to sustainability [49], others assess the sustainability related performance of Indian HEIs [18], and others analyze the individual indicators employed to obtain the ranking [50]. Other authors also use the GreenMetric ranking as the benchmark and conclude that there is low relation between universities' academic and sustainability performance [8].

3. Criteria Employed in the Ranking of Universities

In recent years, social pressure for a firm commitment to sustainability has grown. Universities have not been oblivious to this development. The work carried out by universities places them in a privileged position to disseminate and promote sustainable behavior on and off campus. In this context, it is important to have a tool that allows to track the progress of universities in their sustainable management and their promotion of sustainability in their teaching and research activity. It is also of interest to be able to compare the situation of HEIs at a global level and to give visibility to those universities with the best performance. To achieve these objectives, university rankings are a very useful tool. The aim of the present work is to draw up a ranking that combines the traditional criteria in the field of research and teaching with sustainability criteria. The aim is to develop a ranking that can influence the policies of HEIs and that is necessary to promote sustainable development and contribute to the fight against climate change. Although some researchers have already pointed out the importance of combining research, teaching and sustainability criteria in the development of university rankings [8], few have made methodological proposals [6]. As expressed by most of the literature on sustainability in higher education, the concept of sustainability includes not just management/campus operations and community engagement but teaching and research activities [51,52]. That means, assessing sustainability implies including teaching and research criteria. On the contrary, assessing the teaching and research performance of universities does not necessarily require the inclusion of sustainability criteria. This is the fact for most traditional rankings based on teaching and research performance.

In order to draw up the ranking, this work uses criteria related to teaching, research and sustainability that are already used in university rankings of recognized prestige. Specifically, to capture the teaching and research areas according to traditional criteria, we include the criteria of the Times Higher Education World University Ranking (THE), one of the global benchmark rankings [53] which has been employed in many research studies [2,12,27]. Criteria to assess sustainability performance are obtained from the GreenMetric ranking [46]. The criteria used are listed in Table 1.

Table 1. Criteria employed in the multicriteria university ranking.

Criterion	Definition	Ranking	Weight in the Ranking
Setting and infrastructure (S&I)	Gives information regarding university policy towards green environment	GreenMetric	15%
Energy and climate change (E&C)	This criterion is concerned with the use of energy efficient appliances, energy use, renewable energy policy, green building, climate change adaptation and greenhouse gas emission reductions policy	GreenMetric	21%
Waste	Focuses on waste treatment and recycling activities	GreenMetric	18%
Water	This criterion deals with water use, conservation and recycling	GreenMetric	10%
Transportation	Assesses the transportation policy of universities, including limitation of motor vehicles in the campus, shuttle services, parking area, and pedestrian path policy	GreenMetric	18%

Table 1. Cont.

Criterion	Definition	Ranking	Weight in the Ranking
Education and research (E&R)	Assesses the role played by universities creating the new generation concern with sustainability issues	GreenMetric	18%
Teaching	Assesses the learning environment by means of a reputation survey, staff-to-student ratio, doctorate-to-bachelor's-ratio, doctorates-awarded-to-academic-staff ratio and institutional income	THE	30%
Research	Measures the volume, income and reputation of the research performed by universities by means of a reputation survey, the research income and research productivity	THE	30%
Citations	Research influence, which is quantified capturing the average number of times a university's published work is cited by scholars globally	THE	30%
International outlook	This criterion is made up by following indicators: Proportion of international students, proportion of international staff and international collaboration	THE	7.5%
Industry Income	Measures knowledge transfer as research income from industry due to inventions, innovations and consultancy	THE	2.5%

Table 1 shows the 11 criteria that will be used in the elaboration of the multi-criteria ranking that considers the performance of universities in the areas of sustainability, teaching and research. The GreenMetric ranking provides six criteria: Setting and Infrastructure, Energy and Climate Change, Waste, Water, Transportation and Education and research. The weighting of each criterion in the GreenMetric ranking is as shown in Table 1. The sum of the weightings is 100% in the GreenMetric ranking. THE uses five criteria: Teaching, Research, Citations, International Outlook and Industry income. The weights associated with each criterion are shown in Table 1. All of them will be used in the elaboration of our ranking and they also add 100.

It can be stated that in the GreenMetric ranking, sustainability also covers the areas of teaching and research. However, the THE ranking uses traditional criteria unrelated to sustainability to measure the performance of universities in these areas.

4. Methodology

To compile a ranking, it is necessary to select the criteria to be considered and to assign a weighting to each of them. Both steps are critical and have a great influence on the ranking table. While it is true that the most popular university rankings do adequately describe the criteria they use, there is very little transparency regarding the calculation of the weightings. Prestigious rankings such as Academic Ranking of World Universities (ARWU); Times Higher Education World University Ranking (THE); Quacquarelli Symonds (QS); THE Impact; Sustainability Tracking, Assessment and Rating System (STARS) or IU GreenMetric World University Ranking do not provide enough information about their methodology to obtain the weightings of the different indicators and criteria [5,23]. In cases where the methodology employed to obtain the weights is explained, expert opinion and AHP methodology are generally used, as in the proposals of some researchers [6,20,23,35,54]. When experts are asked, the weights depend on the selection of the experts and is subject to their subjectivity, which introduces a bias in the ranking tables. Finally, the methodologies commonly used in the elaboration of rankings first define dimensions or areas. Then, each criterion or indicator is assigned to a single dimension. First, the importance of the different dimensions is weighted. Second, each criterion is assigned a weight within that dimension, which indirectly implies a weighting in the overall weighting. This way of proceeding assumes that each indicator is associated with only one dimension. However, it is possible

that an indicator is actually related to two or more dimensions. This situation makes it impossible to assume that the different dimensions are independent and makes the correct calculation of indicator weights more complex.

In the following, a multi-criteria goal programming model is proposed that allows solving all these problems simultaneously. The advantageous characteristics of the GP model have led it to be used in different studies. For example, some authors use it to develop a ranking of Spanish saving Banks based on economic and financial variables [55]; others use rank microfinance institutions [56]; others compare sustainable development in the EU-28 countries [57]; and others use rank European companies on their social responsibility [26].

The proposed GP methodology allows for objective rankings without the need for expert decision-makers except for the selection of criteria or indicators. It is a transparent and easily reproducible methodology. In addition, it eliminates the need to create dimensions that encompass the different criteria, so that this prior step is eliminated and it is not necessary to assume the independence of the dimensions. This overcomes the criticisms of other methodologies discussed in the literature.

The proposed GP models allow two different perspectives to be adopted in the elaboration of university rankings. On the one hand, greater weight can be given to criteria that show a greater relationship with the rest of the criteria. On the other hand, greater relevance can be given to criteria that show a singular or independent behavior from the rest of the criteria. Being able to adopt these two perspectives is particularly important in a ranking such as the one we are about to draw up, which involves criteria that quantify HEI activities that may or may not be related, such as education, research or environmental commitment.

GP is a multicriteria technique originally proposed by Charnes and Cooper [58] in which all functions, which may be linear and/or nonlinear and may use continuous and discrete variables, are transformed into goals [59]. Decision-makers are then concerned minimizing the non-achievement of goals [60] and the aim of GP is to minimize the deviations between the achievement goals and their aspiration levels. GP is a realistic approach to many real-world situations where it is not possible to maximize a previously defined utility function and decision-makers try to achieve a set of targets as closely as possible [61]. In this context, GP is in line with the “satisfactory” philosophy [62], as it makes it possible to find optimum solutions in a simplified context or find good enough solutions in a more complex and realistic environment.

The basic formulation of GP is expressed as in Model (1)

$$\min \sum_{i=1}^n |f_i(X) - g_i| \text{ s.t. } X \in F \text{ (} F \text{ is the feasible set)} \quad (1)$$

where x is a vector of decision variables, $f_i(X)$ is usually a linear function of the i -th goal, and g_i is the aspiration level.

The purpose of this paper is to obtain a ranking of Universities which considers the different criteria which have been introduced in the previous section. Therefore, a single measure for universities’ performance must be obtained out of those criteria which will be used to rank the universities. In order to obtain the multicriteria performance of the HEIs, none of the traditional dimensions employed in other rankings are required (i.e., research, education, sustainability). The multicriteria performance of the Universities is obtained as a linear function of the different criteria considered as inputs as expressed in (2):

$$perU_i = \sum_{j=1}^d w_j crit_{ij} \quad (2)$$

where $perU_i$ is the multicriteria performance of the i -th university, $crit_{ij}$ stands for the normalized value of the j -th performance criterion of the i -th university and w_j is the weight of the j -th performance criterion. Our goal is to transparently and objectively determine the w_j weights that are assigned to the different performance criteria. Only

then is it possible to construct a ranking table which is easily reproducible and avoids the criticisms received by other ranking methodologies.

To achieve our goal, we propose different GP models. The first model is known as the weighted goal programming model (WGP). This model assigns the weights to the different criteria by maximizing the similarity between the resulting multicriteria performance and the individual performance criteria. The general WGP model can be expressed as:

$$\min \sum_{i=1}^n \sum_{j=1}^d (\alpha_j n_{ij} + \beta_j p_{ij}) \quad (3)$$

$$s.t. \sum_{j=1}^d (w_j crit_{ij}) + n_{ik} - p_{ik} = crit_{ik} \quad i = 1 \dots n, k = 1, \dots, d \quad (4)$$

$$\sum_{j=1}^d w_j = 1 \quad (5)$$

$$\sum_{j=1}^d w_j crit_{ij} = perU_i \quad i = 1 \dots n \quad (6)$$

$$\sum_{i=1}^n (n_{ij} + p_{ij}) = D_j \quad j = 1 \dots d \quad (7)$$

$$\sum_{j=1}^d D_j = Z \quad (8)$$

The variables in the WGP model must all be positive. The negative (n_{ij}) and positive (p_{ij}) deviations from goals quantify the differences between the observed performance of the i -th university in the j -th criterion and the multicriteria performance estimated by the WGP model for the j -th criterion. In order to capture these situations, coefficients α_j and β_j are employed. α_j takes the value 1 if n_{ij} is unwanted and the value 0 otherwise. $\beta_j = 1$ if p_{ij} is unwanted, otherwise $\beta_j = 0$.

The weight calculated for the j -th criterion is w_j . The weights are computed by minimizing the difference between the estimated multicriteria performance and the performance value of each criterion. In the WGP model the weights of the criteria are obtained without the participation of experts. Experts are only needed to select the criteria that serve as inputs in the model. There is also no need to allocate the different criteria into several areas or dimensions, which weight must also be obtained. The second constraint is that the sum of the weights assigned to each criterion must be one. That is why the deviation variables n_{ij} and p_{ij} are minimized in the objective function. The third constraint shows how the multicriteria performance of the i -th university is obtained. It is the addition of the weighted performance of the i -th-university in all the assessed criteria. The fourth constraint shows that D_j quantifies the difference between the j -th criterion and the estimated multicriteria university performance. Finally, Z is the addition of the estimated overall disagreement. Low Z values mean that the multicriteria performance is in line with the performance of most individual criteria. This will be the case when most criteria are similar to each other. On the contrary, high Z values mean that there are big differences between the multicriteria performance and the performance of the individual criteria. This situation occurs when some criteria are less correlated or independent to the other. When this is the case, the results obtained by the WGP model may be poor, as the objective function seeks for a single multicriteria performance which is aligned with all the criteria employed. Therefore, conflicting criteria for which the improvement of one criterion leads to the worsening of another criterion represent a problem in the WGP model.

The MINMAX GP model or Chebyshev GP model is able to cope with the problem of discordant and even opposite indicators [60]. This model minimizes the maximum difference between the multicriteria performance and the unicriterion performances.

$$\min D \quad (9)$$

$$s.t. \sum_{j=1}^d (w_j crit_{ij}) + n_{ik} - p_{ik} = crit_{ik} \quad i = 1 \dots n, k = 1, \dots, d \quad (10)$$

$$\sum_{i=1}^n (\alpha_{ij} n_{ij} + \beta_{ij} p_{ij}) \leq D_j \quad j = 1 \dots d \quad (11)$$

$$\sum_{j=1}^d w_j = 1 \quad (12)$$

$$\sum_{j=1}^d w_j \text{crit}_{ij} = \text{per}U_i, i = 1 \dots n \quad (13)$$

$$\sum_{i=1}^n (n_{ij} + p_{ij}) = D_j, j = 1 \dots d \quad (14)$$

$$\sum_{j=1}^d D_j = Z \quad (15)$$

All variables in the model have been introduced already in the WGP model, except D , which represents the maximum deviation between the multicriteria performance and the unicriterion performance, that is, the performance of each criterion. There are two differences between the WGP model and the MINMAX model. The first difference is the objective function. The second difference is a new constraint in the MINMAX model, which calculates the value of D as the supremum of the sum of deviations for each criterion j . As mentioned by [26], the solutions of both models represent extreme cases of contrasting strategies. The WGP model fosters the general consensus between single criterion performances, whereas the MINMAX GP model overweights conflicting criteria performances.

The extended GP model [60] offers a compromise between both models. An additional parameter λ is introduced to balance the solutions between the WGP and MINMAX GP models. The λ parameter ranges between 0 and 1. When λ equals 1, the extended GP model obtains the same solutions as the WGP model. If λ is set 0, the model obtains the same solutions as the MINMAX GP model. In fact, both the WGP model and the MINMAX GP model can be considered as special cases of the extended GP model. The EGP model is defined as follows:

$$\min \lambda \sum_{i=1}^n \sum_{j=1}^d (\alpha_j n_{ij} + \beta_j p_{ij}) + (1 - \lambda)D \quad (16)$$

$$\text{s.t. } \sum_{j=1}^d (w_j \text{crit}_{ij}) + n_{ik} - p_{ik} = \text{crit}_{ik}, i = 1 \dots n, k = 1, \dots, d \quad (17)$$

$$\sum_{i=1}^n (\alpha_{ij} n_{ij} + \beta_{ij} p_{ij}) \leq D, j = 1 \dots d \quad (18)$$

$$\sum_{j=1}^d w_j = 1 \quad (19)$$

$$\sum_{j=1}^d w_j \text{crit}_{ij} = \text{per}U_i, i = 1 \dots n \quad (20)$$

$$\sum_{i=1}^n (n_{ij} + p_{ij}) = D_j, j = 1 \dots d \quad (21)$$

$$\sum_{j=1}^d D_j = Z \quad (22)$$

5. Database

To illustrate the development of a multi-criteria ranking of universities that encompasses both the traditional criteria of teaching and research together with sustainability performance by applying the proposed methodology, a database from 2020 of 718 universities from all over the world has been compiled. The selected universities belong simultaneously to THE and GreenMetric rankings. This is because the information regarding the criteria used is obtained from these rankings and the method proposed for the ranking can only work with complete information. In the case of missing information from a university regarding a criterion, that university could not be included in the ranking. The criteria used are those described in Table 1. The descriptive statistical analysis of these indicators is shown in Table 2, which includes the minimum, maximum, range, median, mean and standard deviation.

Table 2. Descriptive analysis of the criteria used in the elaboration of the multi-criteria ranking.

	Min	Max	Range	Median	Mean	Std.Dev
Setting and Infrastructure	0	1450	1450	825	802.54	287.98
Energy and Climate Change	50	1800	1750	900	914.96	323.88
Waste	0	1800	1800	900	876.92	423.03
Water	0	1000	1000	412.5	419.6	228.16
Transportation	0	1700	1700	825	848.71	316.39
Education and Research	0	1800	1800	925	942.51	354.87
Teaching	11.2	90.5	79.3	21.3	25.3	11.41
Research	7.2	99.6	92.4	17	21.37	14.5
Citations	2.1	100	97.9	38.3	43.22	27.33
Industry Income	34.4	100	65.6	39.15	46.69	16.75
International Outlook	14.2	99.1	84.9	41.95	44.9	19.79

In addition to the analysis in Table 2, it is also interesting to perform a correlation analysis (see Table 3). In general terms, the criteria used are not highly correlated. The only exception is the high correlation between the criterion Research and the criterion Teaching (0.87), both included in THE. On the other hand, the correlation between the criteria from the GreenMetric ranking is generally higher than the correlation between these criteria and those used in the THE ranking, and vice versa. Although both rankings include criteria to account for the performance in teaching and research, (Education and Research in the case of GreenMetric and the two criteria Teaching and Research in the case of THE) the criteria of the different rankings do not overlap. That means, they are measuring different realities and therefore their correlation is very low. This conclusion is obvious, especially when comparing how the criteria are defined (see Table 1). This fact suggests that universities can be grouped into two blocks: those that focus their policy on improving their sustainability performance and those that focus mainly on teaching and research aspects. This does not imply that they abandon the other dimensions, but rather that one of these dimensions stands out from the others. Furthermore, Table 3 shows that some criteria have a correlation of less than 0.1 with other criteria, indicating that they are independent in relation to these criteria. This is the case for Industry Income, which is independent of Citations and International Outlook, with a correlation coefficient of 0.02 and 0.05, respectively, and Citations and Setting and Infrastructure (0.02).

Table 3. Correlation analysis.

	S&I	E&C	Waste	Water	Transp.	E&R	Teaching	Research	Citations	Industry Income	Int. Outlook
S&I	1										
E&C	0.36	1									
Waste	0.45	0.60	1								
Water	0.42	0.63	0.59	1							
Transport.	0.52	0.60	0.66	0.57	1						
E&R	0.43	0.63	0.68	0.53	0.67	1					
Teaching	0.21	0.28	0.37	0.37	0.33	0.32	1				
Research	0.21	0.34	0.42	0.43	0.37	0.34	0.87	1			
Citations	0.02	0.24	0.38	0.27	0.22	0.28	0.41	0.52	1		
Industry_Income	0.15	0.11	0.16	0.15	0.21	0.14	0.30	0.42	0.02	1	
Int. Outlook	0.06	0.32	0.35	0.33	0.33	0.38	0.35	0.45	0.57	0.05	1

In order to operate with the collected data and implement the target programming model, a 0–1 normalization must be applied to all values. The purpose of this normalization is to avoid that the weights assigned to the criteria are biased due to the fact that some

criteria have much higher absolute values than the rest. The normalized value of the criteria is calculated as follows:

$$crit_{ij}^* = (crit_{ij} - crit_j^{min}) / (crit_j^{max} - crit_j^{min}) \quad (23)$$

where $crit_{ij}^*$ is the normalized value of the j-th criterion in the i-th university; $crit_j^{max}$ is the maximum value of the j-th criterion; and $crit_j^{min}$ is the minimum value of the j-th criterion.

6. Results and Discussion

With the database described in the previous section, 500 multi-criteria university rankings have been produced by applying the EGP model. The other models described, the WGP model and the MINMAX GP model, are nothing but special cases of the EGP model, when λ takes the value 1 and 0, respectively. Obviously, changing the value of λ affects the weighting of the criteria used and, therefore, the performance of the universities and their position in the ranking table. The model has been run for 500 equally spaced λ values, between 0 and 1. In this way we have obtained 500 rankings, each with its particular criteria weights.

First the weights assigned to the different criteria are analyzed. Figure 1 shows the boxplot representation of the weights assigned to the selected criteria. Analyzing the median value, there is no criterion with a weighting higher than 20%, i.e., there is no criterion that clearly dominates the rest. We can group the criteria into four groups, according to their median weighting: median weight between 20% and 15% (Research and Water), between 15% and 10% (Transportation), between 10% and 5% (Citations, Educations and Research; Energy and Climate Change; Industry Income; International Outlook; Waste) and weight below 5% (Setting and Infrastructure; Teaching). It is necessary to underline that the fact that a criterion is assigned a low weighting does not necessarily imply that it is not important when assessing the performance of the universities. This is because criteria may be correlated with each other, so that a lower weighting of one criterion may imply a higher weighting of a correlated criterion. A low weight may also be obtained if there is little dispersion in the values of one indicator. In this case, the criterion is not valid to discriminate among universities and therefore it will receive a low weight.

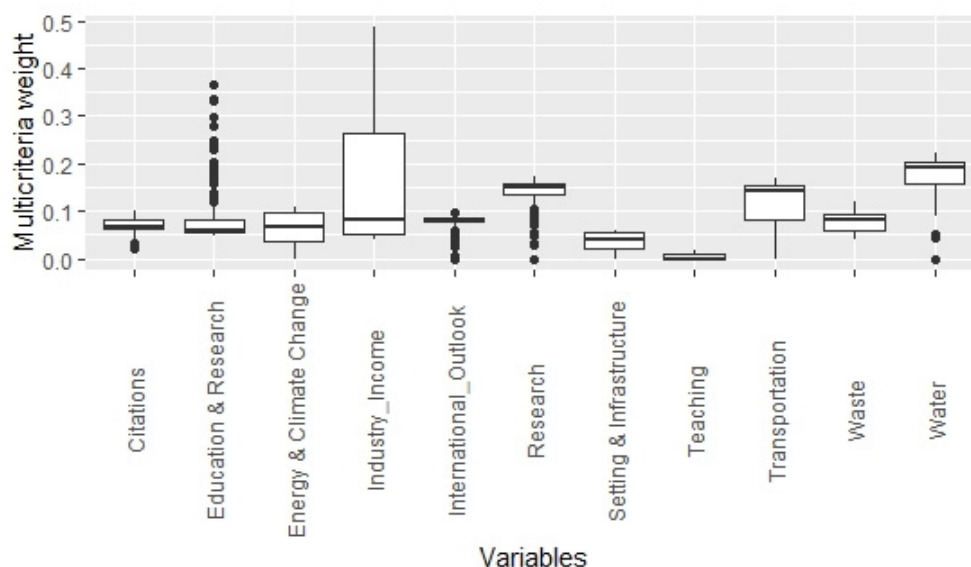


Figure 1. Weights assigned to the criteria for 500 different λ values.

It is important to underline that the weights have been obtained objectively, without the involvement of experts, who may have subjective and potentially discordant opinions. Moreover, the weights of the criteria have been calculated directly, without the need to

create dimensions grouping the different criteria in order to facilitate the assignment of weights.

Besides the analysis of the median value of the weights, it is also interesting to study how the weights of the different criteria change as the λ -value increases from 0 to 1. Figure 2 shows that for low λ -values the weight of sustainability criteria dominate, while increasing λ -values result in higher weights for the traditional criteria (teaching and research, as traditionally measured). In fact, there is a tradeoff between traditional and sustainability criteria when allocating weights.

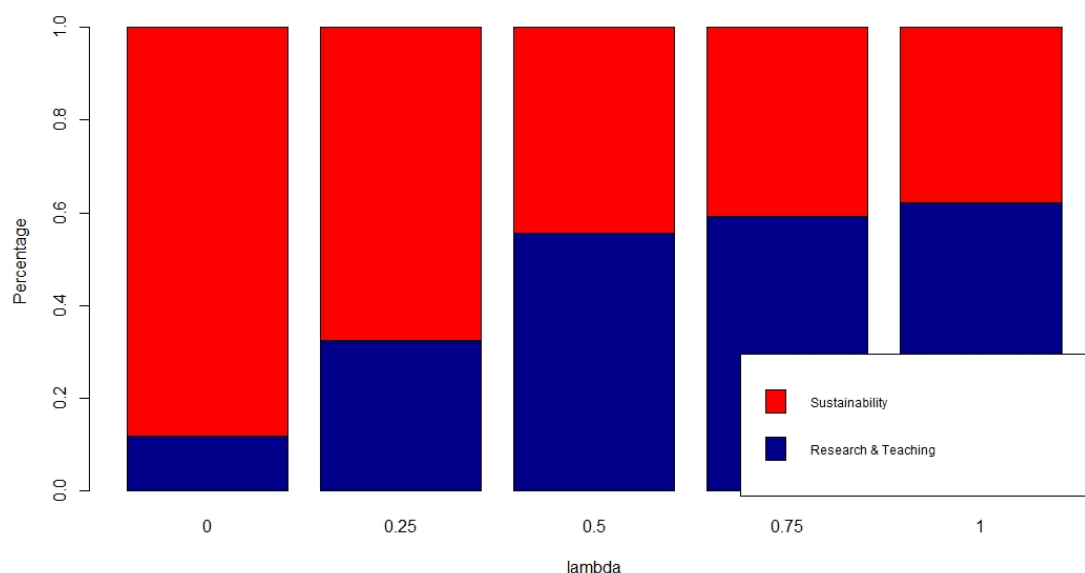


Figure 2. Weights assigned to traditional and sustainability criteria for selected λ -values.

The different values of λ , by influencing the weighting of the criteria, modify the position of the universities within the multi-criteria ranking. Figure 3 shows the top 50 universities if the universities are ordered according to their median position in the 500 rankings. Wageningen University and Research has the best median performance, followed by the University of Groningen. The median performance value is the one used to rank the universities: a higher multi-criteria performance value implies a better position in the ranking. What can be seen in Figure 3 is that there is a clear leader in the ranking, Wageningen University and Research, followed by 4 solid universities heading the ranking: University of Groningen, University of California Davis, Delft University of Technology and Georgia Institute of Technology, which all have a very similar performance. Looking at the median value of the performance, it can also be concluded that the differences are minimal between universities close to each other in the ranking.

Finally, it is interesting to analyze how different λ values impact the performance obtained by the universities, which, in turn, affect their position in the ranking. Figure 4 shows the performance of the top 50 universities according to their median position in the 500 rankings and how this performance changes for different λ values. Again, Wageningen University and Research outstands with very high scores regardless the λ value, so it always leads the ranking. For most universities, the position in the ranking can greatly vary depending on the λ values assigned. It becomes evident that some universities get a much better score when λ is near zero, that is, when sustainability criteria have the most weight (see Figure 2), and low scores for higher λ values. For example, this is the case for Asia University Taiwan and Istanbul Technical University. Other universities receive a better scoring when the λ values are near 1, i.e., when the traditional teaching and research criteria have more weight. This shows how important it is in any ranking methodology to disclose which criteria are employed but also how the weight is assigned to the selected criteria.

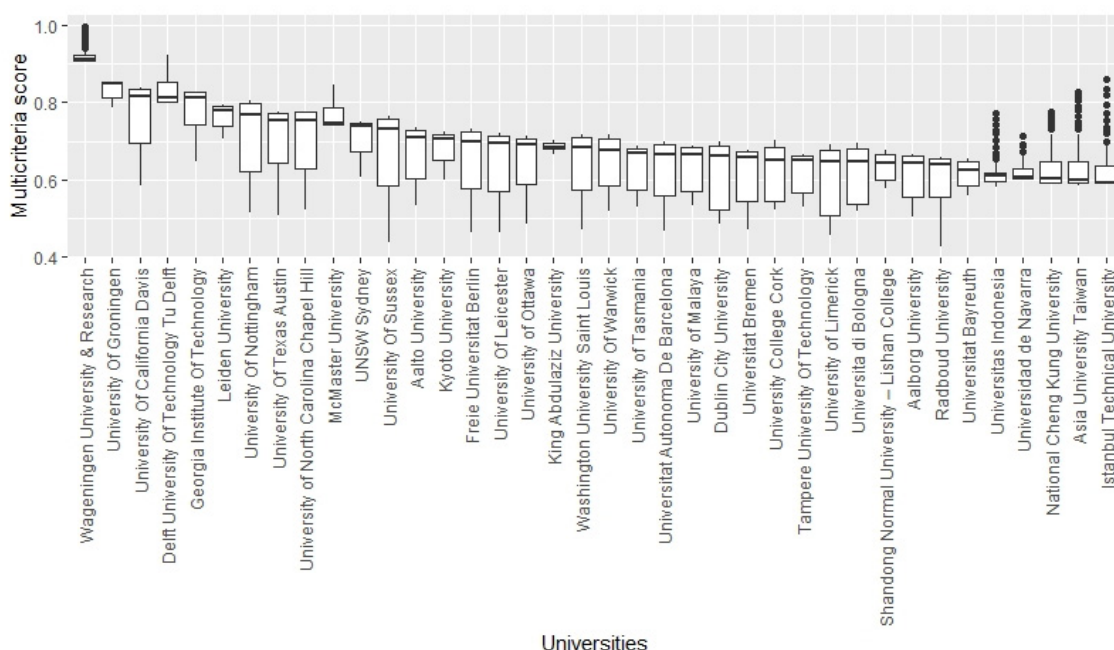


Figure 3. Top 50 universities in the multicriteria rankings obtained for 500 different λ values.

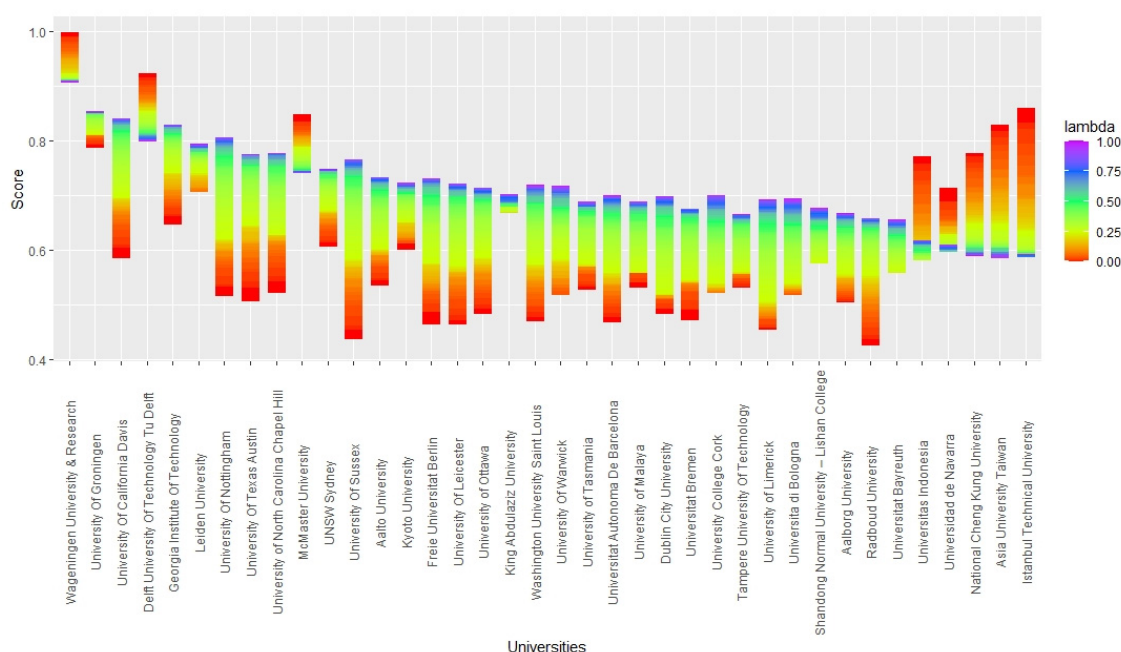


Figure 4. Comparison of the rank obtained by the Top 50 universities in our study, in the UI GreenMetric World University Ranking and in the Times Higher Education World University Ranking.

Looking at Figure 4, we can identify some universities which are particularly strong regarding sustainability criteria, such as Universitas Indonesia or National Cheng Kung University, among others. This is probably related with some strategic decisions by the universities to promote sustainability policies. Other universities are especially strong in teaching and research, such as the University of California Davis or the University of Nottingham, but get poor scores in sustainability criteria. This is probably also due to political decisions and those universities are now starting to focus on sustainability increasing actions. Finally, some universities, such as Wageningen University and Research, University of Groeningen and Delft University of Technology Tu Delft, are both outstanding regarding teaching and research and sustainability. Interestingly, all three universities are

located in the Netherlands, which is a country with a long tradition in teaching and research and where population is very aware of sustainability problems.

7. Conclusions

University rankings are an instrument that allows stakeholders to evaluate and compare the performance of universities in various fields. They are also a powerful instrument for guiding university performance and promoting policies. In fact, many national governments and many universities globally aim to improve their position in international university rankings in order to enhance their prestige. Currently, most university rankings measure university performance from a multi-criteria perspective, which encompasses two aspects of university activity: teaching and research. However, this vision of university's mission, limited to these two areas, may be incomplete. Indeed, there are many stakeholders who believe that universities should promote sustainable development and lead the fight against climate change in the evolution towards a sustainable society. The relationship between universities and sustainable development is multiple and encompasses aspects such as the environmental management of universities, research and technology transfer or the design of curricula that awaken a commitment to sustainability in students. For all these reasons, it seems reasonable to develop university rankings that include the sustainable performance of universities, together with the traditional performance in teaching and research.

The development of rankings has been the subject of much criticism. From a methodological point of view, the process of selecting criteria, the allocation of weights and the lack of transparency are particularly criticized. In this regard, criteria are generally grouped into dimensions and their importance is weighted against the other criteria within that dimension. A weighting is then assigned to each dimension, and in this indirect way the final weighting of each criterion in the ranking is established. With this procedure, a criterion can only belong to one dimension, which is not always the case. Then, it is not reasonable to assume independence between dimensions, which makes it even more difficult to calculate the weights of the criteria.

This paper presents a methodology for the elaboration of multi-criteria rankings based on GP that addresses the above-mentioned criticisms. With this methodology, rankings can be obtained objectively, without the need for experts with subjective opinions and views that may not coincide. In addition, experts may have problems assigning weights to unrelated dimensions. There is also no need to group criteria into dimensions and the weights of the criteria are calculated directly. When applying the proposed methodology, the decision maker only has to set the value of the λ parameter, then the weights of the different criteria are calculated automatically. The methodology is transparent and easily reproducible. To weight the criteria, the proposed GP method, the EGP model, allows the decision-maker to favor more or less criteria that show similar behavior to the other criteria, through the λ parameter. Different values of λ imply different weightings of the criteria of the model and, therefore, different values of the performance of the universities and different positions in the ranking table. In this sense, it should be noted that the proposed methodology limits the subjectivity of the decision-maker in the selection of the λ value.

Once the methodology has been presented, it is applied to the elaboration of a multi-criteria ranking of universities that includes sustainability, teaching and research criteria. The sample of universities consists of 718 universities included in the Times Higher Education World University Ranking and UI GreenMetric World University Ranking. The 11 criteria employed are those used in both rankings. The paper performs a sensitivity analysis to check how different values of λ affect the weight of the criteria and the position of the universities in the ranking. For this purpose, 500 λ values are used. The result shows that, for the selected sample, the weights of the criteria significantly vary depending on the λ -values. There is a tradeoff between traditional criteria (teaching and research) and sustainability criteria. The changes in the weight of the criteria have a major impact on the ranking of universities. This fact underlines the importance of the determination of

weightings in the ranking tables and the importance of transparent methods. Otherwise, the prestige and usefulness of the rankings may be questioned.

Author Contributions: C.B., conceptualization, supervision; F.G. (Fernando García), conceptualization, data curation, writing—original draft, supervision; F.G. (Francisco Guijarro), conceptualization, investigation, data curation, methodology; J.O., investigation, data curation, formal analysis, editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Heffernan, T.A.; Heffernan, A. Language games: University responses to ranking metrics. *High. Educ. Q.* **2017**, *72*, 29–39. [CrossRef]
2. Johnes, J. University Rankings: What Do They Really Show? *Scientometrics* **2018**, *115*, 585–606. [CrossRef]
3. Alves, H. The measurement of perceived value in higher education: A unidimensional approach. *Serv. Ind. J.* **2011**, *31*, 1943–1960. [CrossRef]
4. Brown, R.M.; Mazzarol, T.W. The importance of institutional image to student satisfaction and loyalty within higher education. *High. Educ.* **2008**, *58*, 81–95. [CrossRef]
5. García, F. International university rankings as indicators for the quality of the Spanish universities. *Financ. Mark. Valuat.* **2020**, *6*, 69–84. [CrossRef]
6. Lukman, R.; Krajnc, D.; Glavič, P. University ranking using research, educational and environmental indicators. *J. Clean. Prod.* **2010**, *18*, 619–628. [CrossRef]
7. Berlin Principles on Ranking of Higher Education Institutions-IHEP. Available online: <https://www.ihep.org/publication/berlin-principles-on-ranking-of-higher-education-institutions/> (accessed on 23 November 2021).
8. Muñoz-Suárez, M.; Guadalajara, N.; Osca, J. A Comparative Analysis between Global University Rankings and Environmental Sustainability of Universities. *Sustainability* **2020**, *12*, 5759. [CrossRef]
9. Bowman, N.A.; Bastedo, M.N. Anchoring effects in world university rankings: Exploring biases in reputation scores. *High. Educ.* **2010**, *61*, 431–444. [CrossRef]
10. Safón, V.; Docampo, D. Analyzing the impact of reputational bias on global university rankings based on objective research performance data: The case of the Shanghai Ranking (ARWU). *Scientometrics* **2020**, *125*, 2199–2227. [CrossRef]
11. Jódar, L.; De La Poza, E. How and Why the Metric Management Model Is Unsustainable: The Case of Spanish Universities from 2005 to 2020. *Sustainability* **2020**, *12*, 6064. [CrossRef]
12. Lim, M.A. The building of weak expertise: The work of global university rankers. *High. Educ.* **2017**, *75*, 415–430. [CrossRef]
13. Olcay, G.A.; Bulu, M. Is measuring the knowledge creation of universities possible? A review of university rankings. *Technol. Forecast. Soc. Chang.* **2017**, *123*, 153–160. [CrossRef]
14. Uslu, B. A Path for Ranking Success: What Does the Expanded Indicator-Set of International University Rankings Suggest? *High. Educ.* **2020**, *80*, 949–972. [CrossRef]
15. Findler, F.; Schönherr, N.; Lozano, R.; Stacherl, B. Assessing the Impacts of Higher Education Institutions on Sustainable Development—An Analysis of Tools and Indicators. *Sustainability* **2018**, *11*, 59. [CrossRef]
16. Kappo-Abidemi, C.; Kanayo, O. Higher education institutions and corporate social responsibility: Triple bottomline as a conceptual framework for community development. *Entrep. Sustain. Issues* **2020**, *8*, 1103–1119. [CrossRef]
17. Waheed, B.; Khan, F.I.; Veitch, B.; Hawboldt, K. Uncertainty-based quantitative assessment of sustainability for higher education institutions. *J. Clean. Prod.* **2011**, *19*, 720–732. [CrossRef]
18. Parvez, N.; Agrawal, A. Assessment of sustainable development in technical higher education institutes of India. *J. Clean. Prod.* **2019**, *214*, 975–994. [CrossRef]
19. Caeiro, S.S.; Sandoval-Hamón, L.A.; Martins, R.; Bayas Aldaz, C.E. Sustainability Assessment and Benchmarking in Higher Education Institutions—A Critical Reflection. *Sustainability* **2020**, *12*, 543. [CrossRef]
20. Ozdemir, Y.; Kaya, S.K.; Turhan, E. A scale to measure sustainable campus services in higher education: “Sustainable Service Quality”. *J. Clean. Prod.* **2019**, *245*, 118839. [CrossRef]
21. Bougnol, M.-L.; Dulá, J.H. Technical pitfalls in university rankings. *High. Educ.* **2014**, *69*, 859–866. [CrossRef]
22. Soh, K. The seven deadly sins of world university ranking: A summary from several papers. *J. High. Educ. Policy Manag.* **2016**, *39*, 104–115. [CrossRef]

23. Adenle, Y.A.; Chan, E.H.W.; Sun, Y.; Chau, C. Modifiable Campus-Wide Appraisal Model (MOCAM) for Sustainability in Higher Education Institutions. *Sustainability* **2020**, *12*, 6821. [\[CrossRef\]](#)
24. Docampo, D. Reproducibility of the Shanghai academic ranking of world universities results. *Scientometrics* **2012**, *94*, 567–587. [\[CrossRef\]](#)
25. Moed, H.F. A critical comparative analysis of five world university rankings. *Scientometrics* **2016**, *110*, 967–990. [\[CrossRef\]](#)
26. García-Martínez, G.; Guijarro, F.; Poyatos, J.A. Measuring the social responsibility of European companies: A goal programming approach. *Int. Trans. Oper. Res.* **2017**, *26*, 1074–1095. [\[CrossRef\]](#)
27. García, F.; Guijarro, F.; Oliver, J. A Multicriteria Goal Programming Model for Ranking Universities. *Mathematics* **2021**, *9*, 459. [\[CrossRef\]](#)
28. Alghamdi, N.; Heijer, A.D.; De Jonge, H. Assessment tools' indicators for sustainability in universities: An analytical overview. *Int. J. Sustain. High. Educ.* **2017**, *18*, 84–115. [\[CrossRef\]](#)
29. Lozano, R.; Lukman, R.; Lozano, F.J.; Huisingh, D.; Lambrechts, W. Declarations for sustainability in higher education: Becoming better leaders, through addressing the university system. *J. Clean. Prod.* **2013**, *48*, 10–19. [\[CrossRef\]](#)
30. Alshuwaikhat, H.M.; Abubakar, I.R. An integrated approach to achieving campus sustainability: Assessment of the current campus environmental management practices. *J. Clean. Prod.* **2008**, *16*, 1777–1785. [\[CrossRef\]](#)
31. Disterheft, A.; Caeiro, S.; Ramos, M.R.; Azeiteiro, U. Environmental Management Systems (EMS) implementation processes and practices in European higher education institutions—Top-down versus participatory approaches. *J. Clean. Prod.* **2012**, *31*, 80–90. [\[CrossRef\]](#)
32. Lozano, R. Incorporation and institutionalization of SD into universities: Breaking through barriers to change. *J. Clean. Prod.* **2006**, *14*, 787–796. [\[CrossRef\]](#)
33. Shriberg, M. Institutional assessment tools for sustainability in higher education: Strengths, weaknesses, and implications for practice and theory. *High. Educ. Policy* **2002**, *15*, 153–167. [\[CrossRef\]](#)
34. Gómez, F.U.; Sáez-Navarrete, C.; Lioi, S.R.; Marzuca, V.I. Adaptable model for assessing sustainability in higher education. *J. Clean. Prod.* **2015**, *107*, 475–485. [\[CrossRef\]](#)
35. Lozano, R. A tool for a Graphical Assessment of Sustainability in Universities (GASU). *J. Clean. Prod.* **2006**, *14*, 963–972. [\[CrossRef\]](#)
36. Mader, C. Sustainability process assessment on transformative potentials: The Graz Model for Integrative Development. *J. Clean. Prod.* **2013**, *49*, 54–63. [\[CrossRef\]](#)
37. Velazquez, L.; Munguia, N.; Platt, A.; Taddei, J. Sustainable university: What can be the matter? *J. Clean. Prod.* **2006**, *14*, 810–819. [\[CrossRef\]](#)
38. Assessment System for Sustainable Campus—Hokkaido University Sustainable Campus Management Office. Available online: <https://www.osc.hokudai.ac.jp/en/action/assc> (accessed on 23 November 2021).
39. Sustainability Assessment Questionnaire—ULSF. Available online: <https://ulsf.org/sustainability-assessment-questionnaire/> (accessed on 23 November 2021).
40. Unit-Based Sustainability Assessment Tool (USAT Tool). Available online: <https://www.ru.ac.za/elrc/publicationsandresources/unit-basedsustainabilityassessmenttoolusattool/> (accessed on 23 November 2021).
41. Sustainability Leadership Scorecard | EAUC. Available online: https://www.eauc.org.uk/sustainability_leadership_scorecard (accessed on 23 November 2021).
42. Impact Rankings 2021 | Times Higher Education (THE). Available online: https://www.timeshighereducation.com/impactrankings#!/page/0/length/25/sort_by/rank/sort_order/asc/cols/undefined (accessed on 23 November 2021).
43. People & Planet University League Methodology | People & Planet. Available online: <https://peopleandplanet.org/university-league-methodology> (accessed on 23 November 2021).
44. STARS. Sustainability Tracking Assessment & Rating System. Available online: <https://stars.aashe.org/> (accessed on 23 November 2021).
45. UI GreenMetric. Available online: <https://greenmetric.ui.ac.id/what-is-greenmetric/> (accessed on 23 November 2021).
46. Ragazzi, M.; Ghidini, F. Environmental sustainability of universities: Critical analysis of a green ranking. *Energy Procedia* **2017**, *119*, 111–120. [\[CrossRef\]](#)
47. Suwartha, N.; Sari, R.F. Evaluating UI GreenMetric as a tool to support green universities development: Assessment of the year 2011 ranking. *J. Clean. Prod.* **2013**, *61*, 46–53. [\[CrossRef\]](#)
48. Lauder, A.; Sari, R.F.; Suwartha, N.; Tjahjono, G. Critical review of a global campus sustainability ranking: GreenMetric. *J. Clean. Prod.* **2015**, *108*, 852–863. [\[CrossRef\]](#)
49. Puertas, R.; Marti, L. Sustainability in Universities: DEA-GreenMetric. *Sustainability* **2019**, *11*, 3766. [\[CrossRef\]](#)
50. Perchinunno, P.; Cazzolle, M. A clustering approach for classifying universities in a world sustainability ranking. *Environ. Impact Assess. Rev.* **2020**, *85*, 106471. [\[CrossRef\]](#)
51. Higher Education in the World 4: Table of Contents | Guni Network. Available online: <http://www.guninetwork.org/report/higher-education-world-4/documents> (accessed on 26 November 2021).
52. Filho, W.L.; Eustachio, J.H.P.P.; Caldana, A.C.F.; Will, M.; Salvia, A.L.; Rampasso, I.S.; Anholon, R.; Platje, J.; Kovaleva, M. Sustainability Leadership in Higher Education Institutions: An Overview of Challenges. *Sustainability* **2020**, *12*, 3761. [\[CrossRef\]](#)
53. García, F. International university rankings as a quality measure for the Spanish universities. *Financ. Mark. Valuat.* **2019**, *5*, 33–44. [\[CrossRef\]](#)

-
54. Aliyev, R.; Temizkan, H.; Aliyev, R. Fuzzy Analytic Hierarchy Process-Based Multi-Criteria Decision Making for Universities Ranking. *Symmetry* **2020**, *12*, 1351. [[CrossRef](#)]
 55. García, F.; Guijarro, F.; Moya, I. Ranking Spanish savings banks: A multicriteria approach. *Math. Comput. Model.* **2010**, *52*, 1058–1065. [[CrossRef](#)]
 56. Cervelló-Royo, R.; Guijarro, F.; Martinez-Gomez, V. Social Performance considered within the global performance of Microfinance Institutions: A new approach. *Oper. Res.* **2017**, *19*, 737–755. [[CrossRef](#)]
 57. Guijarro, F.; Poyatos, J.A. Designing a Sustainable Development Goal Index through a Goal Programming Model: The Case of EU-28 Countries. *Sustainability* **2018**, *10*, 3167. [[CrossRef](#)]
 58. Charnes, A.; Cooper, W.W. Management Models and Industrial Applications of Linear Programming. *Manag. Sci.* **1957**, *4*, 38–91. [[CrossRef](#)]
 59. Ignizio, J.P.; Romero, C. Goal Programming. *Encycl. Inf. Syst.* **2003**, 489–500. [[CrossRef](#)]
 60. Romero, C. Extended lexicographic goal programming: A unifying approach. *Omega* **2001**, *29*, 63–71. [[CrossRef](#)]
 61. Tamiz, M.; Jones, D.F.; Romero, C. Goal programming for decision making: An overview of the current state-of-the-art. *Eur. J. Oper. Res.* **1998**, *111*, 569–581. [[CrossRef](#)]
 62. Simon, H. A rational decision-making in business organizations. *Am. Econ. Rev.* **1978**, 493–513. [[CrossRef](#)]