A New Approach to the QS University Ranking Using the Composite I-Distance Indicator: Uncertainty and Sensitivity Analyses

Marina Dobrota
Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, Belgrade 11000, Serbia. E-mail: dobrota.marina@fon.bg.ac.rs

Milica Bulajic
Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, Belgrade 11000, Serbia. E-mail: bulajic.milica@fon.bg.ac.rs

Lutz Bornmann
Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Hofgartenstr. 8, Munich D-80539, Germany. E-mail: bornmann@gv.mpg.de

Veljko Jeremic
University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, Belgrade 11000, Serbia. E-mail: jeremic.veljko@fon.bg.ac.rs

Some major concerns of universities are to provide quality in higher education and enhance global competitiveness, thus ensuring a high global rank and an excellent performance evaluation. This article examines the Quacquarelli Symonds (QS) World University Ranking methodology, pointing to a drawback of using subjective, possibly biased, weightings to build a composite indicator (QS scores). We propose an alternative approach to creating QS scores, which is referred to as the composite I-distance indicator (CIDI) methodology. The main contribution is the proposal of a composite indicator weights correction based on the CIDI methodology. It leads to the improved stability and reduced uncertainty of the QS ranking system. The CIDI methodology is also applicable to other university rankings by proposing a specific statistical approach to creating a composite indicator.

Introduction

In terms of globalization, competitiveness is one of the major goals universities strive to achieve. Thus, a crucial concern is how to assure quality in higher education and enhance global competitiveness through a variety of policies and actions (Cutright, 2013). Given the fact that more and more national accreditation bodies are developing ranking systems, the subject has been widely considered and discussed, and criticism has been raised in the quality assurance community given systems' different aims and approaches (Jovanovic, Jeremic, Savic, Bulajic, & Martic, 2012). The Academic Ranking of World Universities list (ARWU, 2013) has been the most cited ranking list, while the Times Higher Education World University Rankings (THE, 2013) and the Quacquarelli Symonds World University Rankings (QS, 2013) follow closely behind. Many questions about their relevance have been raised (Collyer, 2013; Hou, 2012; Huang, 2012; Jeremic, Bulajic, Martic, & Radojicic, 2011; Pusser & Marginson, 2013; Safon, 2013). Altbach (2006), for example, states that rankings are benchmarks of excellence for the public, and they help academic institutions demonstrate differences among them, which leads to differentiated goals and missions in academic systems. Yet, he still questions these rankings and the parameters they measure and indicates that the use of inadequate metrics to obtain rankings is a highly debated issue.

Because everyone wants a world-class university (Altbach, 2003), rankings have heightened competition between institutions and nations. In doing so, "rankings elevate and fetishize particular conceptualizations of status,

Received February 20, 2014; revised May 5, 2014; accepted May 27, 2014

 • Published online 13 February 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23355

creating a social norm against which all institutions are measured that quietly insinuates itself into public discourse" (Hazelkorn, 2011, p. 24). Chang and Osborn (2005) argue that rankings create powerful images which like advertisements offer a simple "picture" through which consumers, that is, parents and students, can "see" an institution. Although higher education has always been competitive, "rankings make perceptions of prestige and quality explicit" (Hazelkorn, 2011, p. 24; van Vught, 2008).

The starting point for the ranking of universities is thought to be the early 1980s, when *U.S. News and World Report* published the ranking list of American universities (Almgren, 2009). However, ranking of higher education institutions had started much earlier. They were classified in 1870s in the United States (Almgren, 2009). Recent developments in this area (see www.excellencemapping.net) visualize the scientific performance of institutions as ranking lists and on maps (Bornmann, Stefaner, de Moya Anegón, & Mutz, 2014a, 2014b). Despite many similarities, there are some obvious differences among today's popular university rankings. According to Clarke (2002), the conceptualizations of quality that underpin most ranking efforts can be organized into three categories: student achievements, faculty accomplishments, and institutional academic resources. ARWU lists focus on faculty accomplishments by creating the academic ranking, and it consists of a set of indicators describing, among others, the research output (papers published in *Nature* and *Science*, and papers indexed in the *Science Citation Index—Expanded* and *Social Science Citation Index*, and faculty quality [highly cited researchers]). Similarly to ARWU, the ranking of the Centre for Science and Technology Studies (CWTS) in Leiden (CWTS, 2013) is also based on bibliometric data, with lists focusing on the citation impact of institutions. The Leiden ranking excludes from the analysis: journals that are not published in English, journals without a strong international scope, those with a small number of references to other journals in the Web of Science database, and arts and humanities publications (Jeremic, Zornic, Jovanovic-Milenkovic, Markovic, & Radojicic, 2013b; Waltman et al., 2012). CWTS also reports size-independent indicators (average statistics per publication, such as a university's average number of citations per publication), which is a great advantage because it enables comparisons between smaller and larger universities (CWTS, 2013). The QS ranking, on the other hand, focuses on broad areas of interest to prospective students—research, teaching, employability, and internationalization—so it is not as much affected by bibliometric preferences.

According to Hazelkorn (2011), the arrival of the ARWU, THE, and the QS were remarkably well timed and auspicious because global rankings were a product whose time had come. Global rankings have raised the competitive bar and heaped pressure on institutions and systems—becoming the driver of and rationale for significant restructuring and the means by which success and failure are gauged (Hazelkorn, 2011, 2013a). By highlighting reputational advantage, rankings have affected all higher education institutions.

Because higher education plays a fundamental role in creating competitive advantage in a market environment, Hazelkorn (2011) insists that performance does matter and that it is measurable from many different aspects of education. This article studies the aspects that the QS rankings use for evaluating and measuring universities' performances. The QS compares the world's top 800 institutions across broad areas that are of interest to prospective students—research, teaching, employability, and internationalization (QS, 2013)—which are compiled using six criteria: academic reputation, employer reputation, student-to-faculty ratio, citations per faculty, international faculty ratio, and international student ratio. The purpose of the QS ranking has been to recognize universities as the multifaceted organizations that they are and provide a global comparison of their success against the goal of remaining or becoming world class (QS, 2013). They strive to meet the needs of a broad set of stakeholders including students, parents, academics, university staff, and employers, providing (a) a comparative tool to help prospective international students to shortlist potential universities on a global scale and (b) a rudimentary map of the international higher education landscape, which can be used as a component in both institutional and governmental strategic decision making (QS, 2013).

In an attempt to improve their positions in the rankings, universities are strongly tempted to improve their performance specifically in the areas that are measured by ranking indicators (Rauhvargers, 2011). The fact is that publishing in international academic journals has today become predominant mode of a scientific research output. The large weighting factor given to the university's research performance in many different ranking methodologies, such as ARWU, NTU (NTU ranking, 2013), and CWTS Leiden, has led to severe criticism among universities. Jacso (2012) for example points out that "assessment based on Google Scholar instead of Scopus or Web of Science. . .would have serious implications on the nationwide rating and ranking of the research/publishing performance of colleges and universities" (p. 326). Bhattacharjee (2011) observed that during the past couple of years, there have been more than 60 top-ranked researchers from different scientific disciplines, all on the ISI's (Institute for Scientific Information, now Thomson Reuters) highly cited list, who signed a part-time employment arrangement with a university that offered money in exchange for adding their affiliation to a researcher's name on the ISI's list of highly cited researchers. Significant weighting of the university's research performance and the number of cited papers are shortcomings resulting from the very nature of the previously mentioned ranking lists. Despite its other disadvantages, the QS ranking diminishes these deficiencies. Yet, it was criticized for overemphasizing universities' reputations.

Another negative aspect, according to Bookstein, Seidler, Fieder, and Winckler (2010), who criticized THE ranking, is its obvious statistical instability. The authors found that most universities (about 95%) remain in the top 200 from year to

DOI: 10.1002/asi

year, but their positions are often unstable. The ranking has also been criticized: The THE's ranking on "international schools," according to Usher (2014), comprehensively fails the "fall-down-laughing test," as he calls it: "In no world would sane academics agree that Abdulaziz and Maynooth are more international than Harvard." This indicates a clear need for stabilizing the ranking positions and reducing the uncertainty of results, which is one of the main goals of this article.

Despite continuous dispute about the validity of the choice of indicators and their weightings, rankings have acquired legitimacy because their methodologies appear to be statistically rigorous, and various producers willingly engage with critics and occasionally make modifications (Hazelkorn, 2011). Education has become a top priority in emerging economies, and many strive to form partnerships with world-leading universities (Luxbacher, 2013). For example, the Brazilian government's national scholarship program, Science Without Borders, aims to send 100,000 students and researchers to some of the world's best institutions, and according to Luxbacher (2013), these partner universities were selected based on their position in the QS and THE rankings. Clearly, one cannot pretend that ranking does not matter! Because of its global character and scientific approach, the QS ranking is one of the most influential ranking systems. Still, relatively high analytical weightings on reputational surveys have exposed the QS World University Rankings to criticisms over the years (Huang, 2012), and the question about the stability of its performance measurements still remains. A criticism of rankings generally, is the fact that rankings are formed based on a subjective, possible biased weighting system that provides an unreliable ranking framework (Hazelkorn, 2011, 2013b). This is often a subject of controversy because many questions on the weighting system have been raised. Huang (2012) has provided a comprehensive discussion of the indicators and weightings adopted in the QS survey. For example, by emphasizing peer reviews, the QS Rankings regarded reputation as efficiently representative of the performance of a university. However, peer reviews might only reflect the reputation of the university rather than the actual performance, and the questionnaire respondents might merely enumerate internationally renowned universities (Huang, 2012). Because this issue is of great importance (Abramo, Cicero, & D'Angelo, 2013; Bornmann, Mutz, & Daniel, 2013; Tayyab & Boyce, 2013; Waltman et al., 2012) but has weaknesses, in this article we attempt to overcome some of these weaknesses and make certain improvements by proposing an alternative methodology for the QS ranking, that is, a composite indicator based on the I-distance method.

When measuring the performance of a multicriteria phenomenon, a certain composite indicator is formed. The stability of the composite indicator ensures the safety of the observed multicriteria phenomenon. The importance of securing the safety of complex systems has been recognized by various risk analysts in industrial and nonindustrial sectors (Arndt, Acion, Caspers, & Blood, 2013; Guttorp & Kim, 2013; Keung, Kocaguneli, & Menzies, 2013; Mahsuli & Haukaas, 2013; Monferini et al., 2013; Saisana, D'Hombres, & Saltelli, 2011; Wainwright, Finsterle, Zhou, & Birkholzer, 2013). The selection of an appropriate methodology is crucial to any attempt to capture and summarize the interactions among the individual indicators included in a composite indicator or a ranking system (Saisana & D'Hombres, 2008; Saisana & Tarantola, 2002).

Many authors stress the need for an explicit conceptual framework for the index and the usefulness of a multivariate analysis prior to the aggregation of individual indicators to a composite indicator (Saisana & D'Hombres, 2008). Paruolo, Saisana, and Saltelli (2013) consider that composite indicators aggregate individual variables with the aim of capturing certain relevant, yet maybe latent, dimensions of reality. Composite indicators are often constructed (Bruggemann & Patil, 2011) and adopted by many institutions, both for specific purposes and for providing a measurement basis for issues of interest. Saltelli et al. (2008, p. 238) characterize the question of composite indicators as follows: "Composite indicators tend to sit between advocacy (when they are used to draw attention to an issue) and analysis (when they are used to capture complex multidimensional phenomena)." Results and values of composite indicators significantly depend on the indicator weights and are therefore often the subject of controversy (Paruolo et al., 2013; Saltelli, 2007).

To justify the modeling of a composite indicator, it is important to verify its quality. Sensitivity analysis is the study of how uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input (Saltelli, Tarantola, Campolongo, & Ratto, 2004). A related practice is uncertainty analysis, which focuses more on quantifying uncertainty in a model output. Ideally, uncertainty and sensitivity analyses should be run in tandem, with uncertainty analysis proceeding in current practice (Saltelli et al., 2008). Sensitivity analysis can serve a number of useful purposes in the economy of modeling (Saltelli et al., 2008). Uncertainty quantification studies, which may include sensitivity and uncertainty analyses, are essential in risk assessments, and in this context, sensitivity analysis can be used to select the most important parameters or to reduce the number of parameters to be included in the risk assessment (Wainwright et al., 2013).

In this regard, different studies have drawn common conclusions that multicriteria methodology definitions suffer from a ranking instability syndrome (Jovanovic et al., 2012; Keung et al., 2013); that is, different researchers offer conflicting rankings as to what is "best" (Myrtveit, Stensrud, & Shepperd, 2005; Shepperd & Kadoda, 2001). According to Keung et al. (2013), given different historical data sets, different sets of best ranking methods exist under various different situations.

This article primarily proposes an advanced methodology of ranking universities through the possible improvement of

DOI: 10.1002/asi

TABLE 1. The Quacquarelli Symonds (QS) ranking indicators and their respective weightings (QS, 2013).

QS ranking indicators	Weights
Academic reputation (AR)	40%
Employer reputation (ER)	10%
Student-to-faculty ratio (FS)	20%
Citations per faculty (CPF)	20%
International faculty ratio (IF)	5%
International student ratio (IS)	5%

the QS ranking. The article is organized as follows. QS ranking gives a review of the QS ranking methodology. The third section describes the fundamental concept of our proposed methodology and the I-distance method our methodology is based on as well as the concept of uncertainty and sensitivity analyses. The results are given in the fourth section, along with the finding that the proposed methodology is much more stable and steadier than the original QS ranking system. The final section provides concluding remarks.

QS Ranking

Since 2004, when it was first compiled, the QS World University Rankings have expanded to rank more and more universities. In the 2013–2014 edition, 800 universities are ranked. However, far more (more than 2,000) are assessed. The top 400 universities are given individual ranking positions, and the other universities are placed within groups, starting from 401–410, up to group 701 + . The rankings compare the world's top 800 institutions across broad areas that are of interest to prospective students: research, teaching, employability, and international outlook. They are compiled using six criteria, as given in Table 1 (QS, 2013).

QS Rankings adopted reputational survey (academic peer review and employer review) and quantitative indicators (citations per faculty, faculty-student ratio, international faculty, and international students), each weighting 50%. Five fields were reviewed, including, Arts & Humanities, Life Sciences & Biomedicine, Social Sciences, Natural Sciences, and Technology. Rankings of each field are determined by the reputational survey results calculated from questionnaires, which are divided into academic peer review and employer review (Huang, 2012).

Academic reputation is measured through a global survey, asking academics about the best work within their field of expertise. In 2013 the rankings drew on more than 62,000 responses from academics worldwide, collated over 3 years. Only participants' most recent responses were used. They were asked to select 30 universities (excluding their own university) they regarded as the best in the field to which they were affiliated. Regional weightings were applied to counter any discrepancies in response rates. The employer reputation indicator is also based on a global survey, this time taking in 27,900 responses. The survey asks graduate employers to identify the universities that in their view produce the best graduates. It is unique among international university rankings and gives a better sense of how universities are viewed in the job market (QS, 2013).

Student-to-faculty ratio is a simple measure of the number of academic staff employed for every student enrolled. In the absence of an international standard by which to measure teaching quality, it provides insight into which universities are well equipped to provide small-size classes and a good degree of individual supervision. The citation per faculty indicator assesses universities' research output. The QS Rankings collect this information using Scopus, the world's largest database of research abstracts and citations. The latest 5-year data are used, and the total citation count is assessed in relation to the number of academic faculty members at the university (QS, 2013).

International faculty ratio and international student ratio focus on assessing how international a university is, by measuring the proportion of international students and faculty members in relation to overall numbers. Although a highly international student or faculty body is not a measure of quality itself, there is a clear correlation between international intake and success in other areas such as academic reputation and research citation impact. Universities that combine high scores in the round with an international outlook tend to be those that have successfully turned themselves into international centers of excellence (QS, 2013).

The performance ranking of top universities is announced annually. Our research is based on the data from the last five published ranking years within which we have calculated the proposed weights, defined the composite I-distance indicator methodology and performed the appropriate uncertainty and sensitivity analyses.

Methodology

The purpose of our methodology is to calculate the weights of abovementioned indicators based on the data from past years. Weights are based on the correlations between the results of the I-distance values and input indicators. Weights are therefore altered in an appropriate way in order to depict the importance of input variables.

I-Distance Method

The common case with different ranking methods is that possible biases and subjectivity can affect the measurements and evaluation. This problem can be overcome to some extent using the I-distance method, a metric distance in an n-dimensional space (Dobrota, Jeremic, & Markovic, 2012; Jeremic et al., 2011, 2012; Jeremic, Jovanovic-Milenkovic, Martic, & Radojicic, 2013a; Radojicic & Jeremic, 2012). It was originally defined by Ivanovic (1973) and Ivanovic and Fanchette (1973), who devised this method to rank countries according to their level of development based on several indicators, where the main issue was how to use all of them in order to calculate a single synthetic indicator, which will thereafter represent the rank.

DOI: 10.1002/asi

For a selected set of variables $X^T = (X_1, X_2, \ldots X_k)$ chosen to characterize the entities, the I-distance between the two entities $e_r = (x_{1r}, x_{2r}, \ldots x_{kr})$ and $e_s = (x_{1s}, x_{2s}, \ldots x_{ks})$ is defined as

$$D(r,s)=\sum_{i=1}^{k}\frac{|d_i(r,s)|}{\sigma_i}\prod_{j=1}^{i-1}\left(1-r_{ji.12\ldots j-1}\right) \tag{1}$$

where $d_i(r, s)$ is the distance between the values of variable X_i for e_r and e_s, for example, the discriminate effect,

$$d_i(r,s)=x_{ir}-x_{is} \quad i\in\{1,\ldots k\} \tag{2}$$

σ_i the standard deviation of X_i, and $r_{ji.12\ldots j-1}$ is a partial coefficient of the correlation between X_i and X_j, $(j < i)$ (Ivanovic, 1973, 1977).

To overcome the problem of negative coefficient of partial correlation, which can occur when it is not possible to achieve the same direction of movement for all variables in all sets, it is appropriate to use the square I-distance. It is given as:

$$D^2(r,s)=\sum_{i=1}^{k}\frac{d_i^2(r,s)}{\sigma_i^2}\prod_{j=1}^{i-1}\left(1-r_{ji.12\ldots j-1}^2\right) \tag{3}$$

The I-distance measurement is based on calculating the mutual distances between the entities being processed, whereupon they are compared with one another so as to create a rank (Isljamovic, Jeremic, Petrovic, & Radojicic, in press). It is necessary to fix one entity as the reference in the observing set using the I-distance methodology (Seke et al., 2013). The ranking of entities in the set is based on the calculated distance from the reference.

Composite I-Distance Indicator Methodology

In order to create a more stable ranking methodology we modified the weights given by the QS ranking. The process of establishing adequate weights is described in detail. The proposed methodology is referred to as the composite I-distance indicator (CIDI) methodology.

We calculated I-distance values and created I-distance ranks for 5 consecutive years. Subsequently, we examined the stability of each of the compounding indicators by calculating the correlations between the I-distance values and input indicators. Correlations are used because of the special feature of the I-distance method: It is able to present the relevance of input indicators. Instead of assigning subjective weights to input indicators, as it is done within the QS ranking, the I-distance method defines which of the input indicators are most important for the ranking process by putting them into a specific order of importance according to these correlations. Furthermore, because CIDI measures a vector distance, we used Spearman's correlations in the process.

The next step in the proposed methodology was calculating the new weights for each of the compounding indicators, which are based on the appropriate Spearman correlations. Weights are formed by weighting the empirical Spearman correlations: Values of correlations are divided by the sum of correlations. The final sum equals 1, thus forming a novel appropriate weighting system:

$$w_i=\frac{r_i}{\sum_{j=1}^{k}r_j} \tag{4}$$

where r_i $(i = 1, \ldots k)$ is a Spearman correlation between i-th input variable and I-distance value.

Final weights represent the means of acquired values that we call the CIDI. Thus, instead of subjectively defining the values of weights, our proposal is based on a methodological and statistical concept defined by the I-distance method.

Uncertainty and Sensitivity Analyses

QS ranking. To measure the stability and permanence of the official QS ranking system, we have performed uncertainty and sensitivity analyses. The sensitivity of the QS scores is based on the relative contribution of the indicators in the observed period, from the last 5 ranking years, published since 2008. The analysis of the relative contribution of the indicators to each university score can provide useful information as to whether some indicators dominate the overall scores (Saisana & D'Hombres, 2008).

The relative contribution is estimated as a proportion of an indicator score multiplied by the appropriate weight with regard to the overall university score. These contributions have been calculated for each annual ranking list. Subsequently, we calculated the average contributions and their standard deviations. Using the Monte Carlo simulation method, we simulated the score results 100 times, each time recording the results. Furthermore, according to uncertainty and sensitivity methodology (Saisana & D'Hombres, 2008), we counted the ranks for all universities, thus measuring the amount of uncertainty of the official ranking results.

CIDI. Proposed weights are acquired from the last 5-year data set, from 2008 to 2013. Based on these values we were able to calculate mean values, and standard deviations, which are used as an input for the Monte Carlo simulation. Using this method, we simulated the score results 100 times, each time recording the results. Furthermore, similarly to measuring the sensitivity of the original results (QS, 2013; Saisana & D'Hombres, 2008), we counted the ranks for all universities, thus measuring the amount of uncertainty of the CIDI ranking results.

Results

CIDI

The first step in our research implies calculating I-distance values for the QS ranking. The inputs in calculating

DOI: 10.1002/asi

TABLE 2. Weightings of input Quacquarelli Symonds (QS) ranking indicators based on I-distance methodology.

QS ranking indicators	2008	2009	2011	2012	2013	Mean	*SD*
Academic reputation (AR)	0.201	0.189	0.202	0.189	0.212	0.199	0.009776
Employer reputation (ER)	0.194	0.217	0.171	0.211	0.195	0.198	0.017842
Student-to-faculty ratio (FS)	0.151	0.150	0.161	0.144	0.135	0.148	0.009700
Citations per faculty (CPF)	0.142	0.126	0.132	0.130	0.148	0.136	0.009310
International faculty ratio (IF)	0.146	0.146	0.155	0.142	0.140	0.146	0.006052
International student ratio (IS)	0.165	0.172	0.179	0.184	0.171	0.173	0.007175

TABLE 3. Differences in composite I-distance indicator (CIDI) weights and original Quacquarelli Symonds (QS) ranking weights.

QS ranking indicators	QS ranking weights	CIDI weights
Academic reputation (AR)	40%	19.9%
Employer reputation (ER)	10%	19.8%
Student-to-faculty ratio (FS)	20%	14.8%
Citations per faculty (CPF)	20%	13.6%
International faculty ratio (IF)	5%	14.6%
International student ratio (IS)	5%	17.3%

I-distance values are six indicators which constitute the QS ranking. The relevance of I-distance ranking is presented and elaborated in a number of papers (Dobrota et al., 2012; Jeremic et al., 2012, 2013a; Jovanovic et al., 2012). Thus, we were able to gain I-distance values for 5 consecutive years. Subsequently, we calculated the correlations between I-distance values and the whole set of input indicators (Dobrota et al., 2012; Jeremic et al., 2012) for each referred year.

Because the results appeared to be quite stable, and there were no large oscillations between the correlations, we calculated the new weights for each compounding indicator based on the appropriate correlations of these items. The proposed weights are calculated by dividing the appropriate correlations with the sum of correlations, providing the sum equals 1. Thus we obtained the appropriate weights for input indicators, as presented in Table 2. Proposed weights are mean values of weights calculated for the whole period.

Table 3 compares the weights proposed by the QS ranking and CIDI weights. The largest differences are with indicators academic reputation, employer reputation, and international student ratio. Academic reputation is assigned 40% weight according to the official QS ranking, whereas our method calculates the share of this indicator to be 19.9%, giving it a lower significance in the overall ranking score. This finding is in accordance with the views of some authors who have criticized this indicator and the fact that it has been given such a heavy weight (Huang, 2012; Liu, 2013).

On the other hand, employer reputation is weighted at 10% according to the official QS ranking, whereas our method calculates the share of this indicator to be 19.8%, thus reinforcing its significance. Similarly, international student ratio is weighted at 5% according to the official QS ranking, and 17.3% according to the CIDI methodology.

Table 4 presents the results of our research, giving the CIDI scores, CIDI ranks, and the comparison of our scores to the official QS ranking scores. The results are shown for 40 (10%) top ranked universities (the whole set of the results for 392 universities is available upon request).

Many differences between the QS and CIDI scores and ranks are visible in Table 4. The Massachusetts Institute of Technology (MIT) is ranked first according to both QS and CIDI. Imperial College London is ranked second according to CIDI, whereas it is ranked fifth according to the QS; this is probably because of the fact that this university scores 100 in employer reputation, which holds it back in the QS score because of its 10% weighting, whereas the 19.8% weighting by the CIDI allows it to break through; similarly it scores 99.9 in the international student ratio that weights 17.3% by the CIDI and only 5% by the original QS weighting. Generally, the top six universities remain the same according to both ranking methods, with slight changes in positions.

The first significant difference in ranking positions is found in the Ecole Polytechnique Fédérale de Lausanne (EPFL), which is ranked ninth according to the CIDI and 19th according to the QS. Along with ETH Zurich (Swiss Federal Institute of Technology), which is ranked seventh by the CIDI and 12th by the QS, the CIDI has enhanced their overall ranking positions. This may be attributed to the fact that the QS has been criticized for favoring universities from English-speaking countries (Huang, 2012; Li, Shankar, & Tang, 2011; Safon, 2013; Soh, 2013; Taylor & Braddock, 2007). In fact, among the 30 top ranked universities, only these two are not from English-speaking countries (Switzerland).

Official QS Ranking Uncertainty and Sensitivity

We have carried out a thorough uncertainty and sensitivity analyses of the official QS ranking under a plurality of scenarios in which we activate simultaneously sources of uncertainty. As elaborated before, the uncertainty and sensitivity of the QS scores is based on the relative contribution of the indicators, which is estimated as the proportion of an indicator score multiplied by the respective weight with regard to the overall university score. Using the Monte Carlo simulation method, we have simulated the score results 100 times. The results show that, on average, the shares of the

DOI: 10.1002/asi

TABLE 4. Composite I-distance indicator (CIDI) scores, composite I-distance indicator ranks, and comparison with official Quacquarelli Symonds (QS) ranking scores and ranks for 2013; 40 top ranked universities (sorted in ascending order by the CIDI rank).

Institution	QS score	QS rank	CIDI score	CIDI rank
Massachusetts Institute of Technology (MIT)	100.00	1	98.97	1
Imperial College London	98.80	5	98.90	2
University College London (UCL)	98.90	4	98.45	3
University of Oxford	98.70	6	98.15	4
University of Cambridge	99.00	3	98.02	5
Harvard University	99.20	2	96.49	6
ETH Zurich (Swiss Federal Institute of Technology)	94.30	12	95.12	7
Yale University	96.50	8	92.94	8
Ecole Polytechnique Fédérale de Lausanne (EPFL)	90.90	19	92.86	9
National University of Singapore (NUS)	89.40	24	91.97	10
King's College London (KCL)	90.90	20	91.89	11
University of Edinburgh	91.30	18	91.85	12
Stanford University	96.80	7	91.46	13
University of Toronto	91.30	17	90.99	14
University of Hong Kong (HKU)	88.60	26	90.94	15
California Institute of Technology (CALTECH)	96.10	11	90.56	16
University of Chicago	96.20	9	90.55	17
Australian National University (ANU)	88.50	27	89.94	18
Princeton University	96.10	10	89.46	19
McGill University	90.60	21	89.28	20
The Hong Kong University of Science and Technology (HKUST)	84.40	34	88.12	21
Cornell University	92.50	15	87.83	22
University of California, Berkeley (UCB)	89.00	25	87.49	23
The University of Melbourne	86.00	31	87.11	24
The University of Manchester	85.20	33	87.08	25
University of Bristol	86.60	30	86.30	26
Nanyang Technological University (NTU)	81.10	41	86.24	27
University of Pennsylvania	93.80	13	85.65	28
The University of Sydney	82.90	38	85.56	29
The University of New South Wales (UNSW)	78.80	52	84.17	30
Ecole Polytechnique Paristech	81.10	42	84.16	31
The University of Queensland (UQ)	80.90	43	83.48	32
Johns Hopkins University	92.10	16	83.47	33
Columbia University	93.60	14	83.06	34
The Chinese University of Hong Kong (CUHK)	82.30	39	82.83	35
London School of Economics and Political Science (LSE)	73.90	68	82.19	36
The University of Warwick	74.50	64	81.38	37
University of St Andrews	71.60	83	79.33	38
University of Glasgow	78.90	51	79.04	39
Monash University	73.70	69	78.82	40

TABLE 5. Shares of input Quacquarelli Symonds (QS) ranking indicators.

QS ranking indicators	2008	2009	2011	2012	2013	Mean	*SD*
Academic reputation (AR)	0.392	0.410	0.417	0.412	0.414	0.409	0.009655
Employer reputation (ER)	0.101	0.100	0.056	0.094	0.105	0.091	0.020283
Student-to-faculty ratio (FS)	0.201	0.193	0.247	0.211	0.203	0.211	0.021136
Citations per faculty (CPF)	0.203	0.197	0.172	0.184	0.184	0.188	0.012038
International faculty ratio (IF)	0.051	0.049	0.052	0.050	0.048	0.050	0.001556
International student ratio (IS)	0.051	0.050	0.056	0.049	0.046	0.051	0.003735

QS indicators in the overall score are not consistent with the corresponding weights proposed by the QS ranking system in all cases. These shares are presented in Table 5.

For example, citations per faculty was assigned a 20% weight, yet the relative average share of this indicator in the overall score is lower (18.8%). On the other hand, student-to-faculty ratio has an average contribution of 21.1%, which is higher than the 20% weight assigned to this indicator. For the remaining indicators the average contributions are closer to the actual weights. According to these results, an analysis

TABLE 6. Uncertainty and sensitivity of Quacquarelli Symonds (QS) ranks (20 first ranked universities).

Institution	1–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45
Massachusetts Institute of Technology (MIT)	72	22	4	2					
Imperial College London	63	29	4	4					
University of Cambridge	57	34	8	1					
University College London (UCL)	47	41	9	3					
Harvard University	53	31	11	3	2				
University of Oxford	54	29	13	1	3				
Yale University	33	36	19	9	2	1			
Stanford University	29	32	25	11	3				
University of Chicago	22	43	23	10	2				
Princeton University	26	34	22	13	4	1			
California Institute of Technology (CALTECH)	15	44	23	12	4	2			
ETH Zurich (Swiss Federal Institute of Technology)	6	24	39	23	5	2	1		
University of Pennsylvania	8	20	35	25	8	4			
Columbia University	7	21	33	21	14	4			
Johns Hopkins University	1	11	29	31	16	11	1		
Cornell University	3	6	24	29	27	7	4		
University of Toronto	2	8	21	28	20	20	1		
Ecole Polytechnique Fédérale de Lausanne (EPFL)		6	19	33	24	26	1	1	
University of Edinburgh		6	23	22	32	12	5		
King's College London (KCL)	1	7	18	34	19	14	5	1	1

of the shares is of added value, because the relative contribution of each indicator to the QS score is not quite reflected in the weight linked to the indicator.

The frequency matrix of the university ranks based on the QS ranking for the 20 first ranked universities is given in Table 6.

MIT, Imperial College London, University of Cambridge, University College London (UCL), and Harvard University are averagely ranked as top five universities according to the original QS ranking. According to our Monte Carlo simulation results, MIT is ranked in positions 1 to 5 in about 70% of simulations, in positions 6 to 10 in about 20% of simulations, and in other simulations it spreads even to position 20. The others have taken top five positions in less than 65% of the simulations. Harvard University, which is ranked second in 2013, is in positions 1 to 5 in only about 50% of cases; in about 30% in positions 6 to 10; and in other simulations it even spreads to position 25. As we move forwards, the impact of the assumptions on the rank becomes even more pronounced, as shown in Table 6. For example, the University of Toronto, which is ranked 17th by the QS ranking in 2013, has a very uncertain position: it could be ranked anywhere between the 1st and 35th place. Moving forwards to 392th place, results are even more widely spread, which is evident in Figure 1.

According to these results we can conclude that universities are medium or highly sensitive to the methodological assumptions in the QS ranking. For highly sensitive universities, the space of inference of a university's rank is too wide to draw any meaningful conclusions. According to Saisana and D'Hombres (2008), common to most of the "highly sensitive" universities is the fact that they are neither very good nor very bad in the majority of the input indicators but somewhere in between.

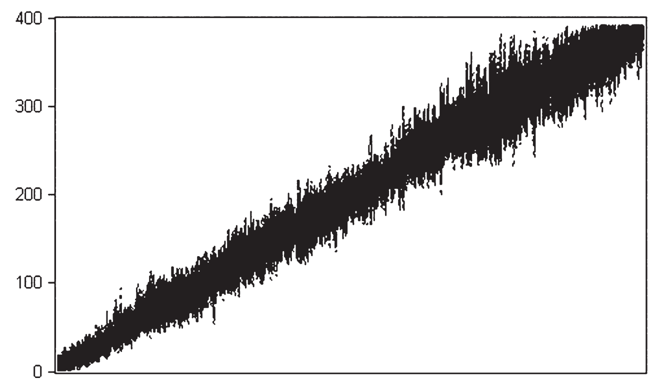

FIG. 1. Uncertainty and sensitivity of Quacquarelli Symonds ranks.

Composite I-Distance Indicator Uncertainty and Sensitivity

For comparison purposes, we performed the same analyses on the CIDI. In this case, uncertainty and sensitivity analyses were based on the proposed weights, acquired from the data. Using the Monte Carlo simulation method, we have simulated the score results 100 times. The frequency matrix of the university ranks based on the CIDI for the 20 first-ranked universities is given in Table 7.

In this case, the results of the Monte Carlo simulation have been evidently improved. Massachusetts Institute of Technology (MIT), which is ranked first according to both CIDI and QS scores, has taken the position 1 to 5 in 80% of simulations, whereas in other cases is in positions 6 to 10. Along with Imperial College London (ranked in positions 1 to 5 in 85% of simulations), UCL, University of Cambridge, and University of Oxford, it is among the top five averagely

DOI: 10.1002/asi

TABLE 7. Uncertainty and sensitivity of composite I-distance indicator (CIDI) ranks (20 first ranked universities).

Institution	1–5	6–10	11–15	16–20	21–25	25–30	31–35	36–40	41–45
Imperial College London	85	15							
Massachusetts Institute of Technology (MIT)	80	20							
University College London (UCL)	83	14	3						
University of Cambridge	75	19	6						
University of Oxford	70	26	4						
Harvard University	45	45	10						
ETH Zurich (Swiss Federal Institute of Technology)	29	54	12	5					
Yale University	4	45	34	14	3				
King's College London (KCL)	8	46	27	15	4				
University of Toronto	5	26	41	24	4				
Ecole Polytechnique Fédérale de Lausanne (EPFL)	4	29	38	19	8	2			
University of Edinburgh	1	32	33	25	21				
National University of Singapore (NUS)	2	28	38	23	6	3			
Stanford University	4	20	33	21	16	6			
University of Hong Kong (HKU)	4	20	28	33	11	4			
University of Chicago		19	36	26	16	3			
Princeton University		16	29	31	14	10			
Australian National University (ANU)		6	29	35	19	8	3		
California Institute of Technology (CALTECH)		9	20	33	22	15	1		
McGill University		6	21	28	29	12	4		

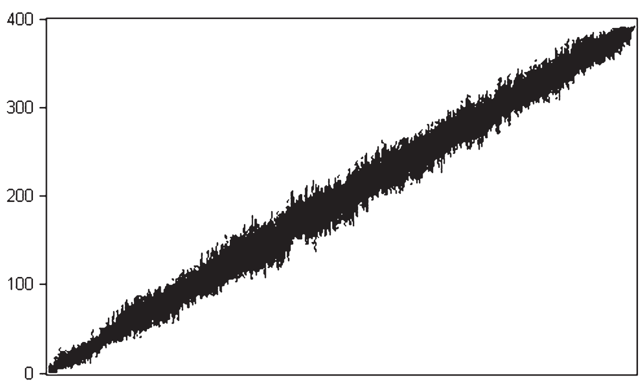

FIG. 2. Uncertainty and sensitivity of CIDI ranks.

ranked universities. All of these universities have taken positions from 1 to 5 in more than 70% of simulations. Harvard University, which is ranked sixth according to the CIDI, has taken positions 1 to 5 in 45%, 6 to 10 in 45%, and 11 to 15 in 10% of simulations. As we move downwards, we can clearly notice much greater stability with the CIDI ranking system than within the official QS ranking results. If we move forward to 392th place, we notice the greater stability of the CIDI ranking system that we have proposed, as shown in Figure 2.

We can conclude that all universities have a low or, in the worst case, medium degree of sensitivity to the CIDI methodological assumptions. Thus, we have been able to propose a more stable methodology that decreases the entropy of the ranking system.

As mentioned previously, our proposal provides methodologically and statistically justified weights which are derived from the correlations calculated for previous periods. In addition, the proposal provides a different perspective on the importance of each input variable, and a correction in the weighting factor for each input indicator. Moreover, not only does our proposal give a more accurate result, but it is also more stable than the official QS ranking. If we calculate the weights according to the CIDI methodology, we are able to achieve a more stable result and to decrease the entropy of the ranking system.

Discussion and Conclusion

Despite shortcomings, biases, and drawbacks, rankings are accepted by stakeholders and the wider public because of their simplicity. Thus, university rankings are not going to disappear; indeed, the number of rankings is expected to increase although they will become more specialized (Marginson, 2011). University leaders and administrators are gaining experience by working with rankings, making this issue the subject of debate in many meetings and events held over the last few years (Rauhvargers, 2014).

The QS World University Rankings is one of the most influential ranking systems. It attempts to provide a comparative tool to help students shortlist potential universities (QS, 2013). However, it does have its limitations. For example, subjective weightings on academic reputation surveys have brought about a large amount of distrust over the years (Avery, Glickman, Hoxby, & Metrick, 2013), especially because the ranking relies too heavily on reputation indicators derived from expert opinions (Huang, 2012; Safon, 2013; Taylor & Braddock, 2007). According to De Witte and Hudrlikova (2013, p. 341) this indicator may be a mere "symptom" of excellence: It favors world-renowned institutions and does not represent current research performance. They emphasize that the ranking rather reflects

DOI: 10.1002/asi

reputational factors and not necessarily the quality or performance of institutions, and although rankings should take into account the diversity among institutions, the current reputation surveys only reinforce the existing reputation and prestige of particular universities. Rankings such as the QS, which are totally or partially based on academic, recruiter or student opinions, are beginning to lose touch with changes in quality over time, and, instead, exclusively reflect changes in the prestige of institutions (Bowman & Bastedo, 2011).

The ranking was also criticized for the low response rates of the review surveys and for a general lack of methodological transparency (Jons & Hoyler, 2013; Safon, 2013). Furthermore, these surveys commonly have regional bias, as the majority of the received questionnaires come from English-speaking countries, clearly favoring their universities (Huang, 2012; Safon, 2013; Soh, 2013; Taylor & Braddock, 2007). Universities in English-speaking countries have the advantage of being able to recruit both native and nonnative English-speaking academics from around the world (Li et al., 2011; Safon, 2013).

By proposing the CIDI for measuring the performance of world universities, we have attempted to improve a measuring and ranking system. Important improvement is reflected in the objectiveness of our methodology. According to Soh (2013, 2014) a certain discrepancy can be found between nominal (assigned) weights and attained (actual) indicator weights. ". . . when a ranker specifies the weights for the constituent indicators to calculate the overall scores (on which the universities are ranked), rank users assume the weights are maintained as specified. In the event that the assigned (nominal) weights and the actual (attained) weights differ, rank users are not getting what they have been promised and are thereby misinformed" (Soh, 2014, p. 2). "Choosing the most appropriate weights is a crucial step in the aggregation of outcome variables as they have obviously a significant effect on the overall composite indicator and the university's ranking. There is, however, inevitable subjectivity in setting weights. In a priori selected weights, the ranking favors universities for which the weights 'fit best' " (De Witte & Hudrlikova, 2013, p. 342). By defining the CIDI methodology we have overcome the disadvantage of subjectively assigned weights to the set of input indicators. Instead of using these weights, we have established the weighting system based on a multivariate statistically and methodologically grounded course.

As a future research directions, we suggest incorporating a greater number of Monte Carlo simulations (1,000, or even 10,000), which will consequently lead to more impartial results. Furthermore, even if the results do show an improvement regarding the stability and reducing their uncertainty, there is still plenty of room for refining the methodology to obtain even more stable and less uncertain results. As some authors (Bookstein et al., 2010) questioned the stability of ranking lists, the improvements that CIDI has made concerning the significant decrease in the uncertainty of universities' ranking positions are welcome, even if not necessary.

As Rauhvargers (2014) claims: Rankings are here to stay. He states that even if academics are aware that the results of rankings are biased and cannot satisfactorily measure institutional quality, they also recognize that an impressive position in the rankings can be a key factor in securing additional resources, recruiting more students and attracting strong partner institutions. So instead of emphasizing their flaws, a more effective path might be to revise and improve them. The proposed weighting system in forming the CIDI is far more stable, and this was shown by the differences between Tables 6 and 7, as well as differences between Figures 1 and 2. The uncertainty and sensitivity analyses of the official QS rankings show that this methodology is affected with high degree of sensitivity to the university ranks. Previous revisions of the QS ranking have generated different rankings over the years (Aguillo, Bar-Ilan, Levene, & Priego, 2010; Jons & Hoyler, 2013). In comparison, the CIDI reveals a lower degree of uncertainty and sensitivity to the plurality of the methodological scenarios, thus allowing CIDI ranks to be more reliable.

References

Abramo, G., Cicero, T., & D'Angelo, C.A. (2013). The impact of unproductive and top researchers on overall university research performance. Journal of Informetrics, 7(1), 166–175. doi: 10.1016/j.joi.2012.10.006

Aguillo, I.F., Bar-Ilan, J., Levene, M., & Priego, J.L.O. (2010). Comparing university rankings. Scientometrics, 85(1), 243–256. doi: 10.1007/s11192-010-0190-z

Almgren, E. (2009). Ranking of universities and higher education institutions for student information purposes? Swedish National Agency for Higher Education: Stockholm. Retrieved from https://www.hsv.se/download/18.3127d9ec12496ad89237ffe857/0927R.pdf

Altbach, P.G. (2003). Introduction to higher education theme issue on the academic profession in central and Eastern Europe. Higher Education, 45(4), 389. doi: 10.1023/A:1023949808905

Altbach, P.G. (2006). The dilemmas of ranking. International Higher Education, 42, 2–3.

Arndt, S., Acion, L., Caspers, K., & Blood, P. (2013). How reliable are county and regional health rankings? Prevention Science, 14(5), 497–502. doi: 10.1007/s11121-012-0320-3

ARWU (2013). Academic ranking of world universities. Retrieved from http://www.shanghairanking.com/ARWU2013.html

Avery, C.N., Glickman, M.E., Hoxby, C.M., & Metrick, A. (2013). A revealed preference ranking of U.S. colleges and universities. The Quarterly Journal of Economics, 128(1), 425–467. doi: 10.1093/qje/qjs043

Bhattacharjee, Y. (2011). Citation impact. Saudi universities offer cash in exchange for academic prestige. Science, 334(6061), 1344–1345.

Bookstein, F.L., Seidler, H., Fieder, M., & Winckler, G. (2010). Too much noise in the times higher education rankings. Scientometrics, 85(1), 295–299. doi: 10.1007/s11192-010-0189-5

Bornmann, L., Mutz, R., & Daniel, H.D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden ranking 2011/2012. Journal of the American Society for Information Science and Technology, 64(8), 1649–1658. doi: 10.1002/asi.22857

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014a). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualization of results from multi-level models. Online Information Review, 38(1), 43–58.

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014b). What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. Journal of Informetrics, 8(3), 581–593.

Bowman, N.A., & Bastedo, M.N. (2011). Anchoring effects in World University Rankings: Exploring biases in reputation scores. High Education, 61(4), 431–444. doi: 10.1007/s10734-010-9339-1
Bruggemann, R., & Patil, G.P. (2011). Ranking and prioritization for multi-indicator systems. Dordrecht, The Netherlands: Springer. ISBN:978-1-4419-8477-7.
Chang, G.C., & Osborn, J.R. (2005). Spectacular colleges and spectacular rankings: The "US News" rankings of American "best" colleges. Journal of Consumer Culture, 5(3), 338–364. doi: 10.1177/1469540505056794
Clarke, M. (2002). Some guidelines for academic quality rankings. Higher Education in Europe, 27(4), 443–459.
Collyer, F. (2013). The production of scholarly knowledge in the global market arena: University ranking systems, prestige and power. Critical Studies in Education, 54(3), 245–259. doi: 10.1080/17508487.2013.788049
Cutright, M. (2013). University rankings: Theoretical basis, methodology, and impacts on global higher education. Review of Higher Education, 37(1), 103–104.
CWTS (2013). Centre for Science and Technology Studies Leiden Ranking 2013. Retrieved from http://www.leidenranking.com/
De Witte, K., & Hudrlikova, L. (2013). What about excellence in teaching? A benevolent ranking of universities. Scientometrics, 96(1), 337–364.
Dobrota, M., Jeremic, V., & Markovic, A. (2012). A new perspective on the ICT development index. Information Development, 28(4), 271–280. doi: 10.1177/0266666912446497
Guttorp, P., & Kim, T.Y. (2013). Uncertainty in ranking the hottest years of U.S. surface temperatures. Journal of Climate, 26(17), 6323–6328. doi: 10.1175/JCLI-D-12-00760.1
Hazelkorn, E. (2011). Rankings and the reshaping of higher education: The battle for world-class excellence. Houndsmills, UK: Palgrave MacMillan.
Hazelkorn, E. (2013a). How rankings are reshaping higher education. In V. Climent, F. Michavila, & M. Ripolles (Eds.), Los rankings universitarios: Mitos y realidades. Madrid: Universitat Jaume I y Editorial Tecnos.
Hazelkorn, E. (2013b). World-class universities or world class systems? Ranking and higher education policy choices. Keynote address, UNESCO Global Forum on Rankings and Accountability in Higher Education, May 16–17, Paris, France.
Hou, A.Y.C. (2012). Impact of excellence programs on Taiwan higher education in terms of quality assurance and academic excellence, examining the conflicting role of Taiwan's accrediting agencies. Asia Pacific Education Review, 13(1), 77–88. doi: 10.1007/s12564-011-9181-x
Huang, M.H. (2012). Opening the black box of QS World University Rankings. Research Evaluation, 21(1), 71–78. doi: 10.1093/reseval/rvr003
Isljamovic, S., Jeremic, V., Petrovic, N., & Radojicic, Z. (in press). Colouring the socio-economic development into green: I-distance framework for countries' welfare evaluation. Quality & Quantity. doi: 10.1007/s11135-014-0012-0
Ivanovic, B. (1973). A method of establishing a list of development indicators. Paris, France: United Nations Educational, Scientific and Cultural Organization.
Ivanovic, B. (1977). Classification theory. Belgrade, Serbia: Institute for Industrial Economic.
Ivanovic, B., & Fanchette, S. (1973). Grouping and ranking of 30 countries of Sub-Saharan Africa, two distance-based methods compared. Paris, France: United Nations Educational, Scientific and Cultural Organization.
Jacso, P. (2012). Grim tales about the impact factor and the h-index in the Web of Science and the Journal Citation Reports databases: Reflections on Vanclay's criticism. Scientometrics, 92(2), 325–354. doi: 10.1007/s11192-012-0769-7
Jeremic, V., Bulajic, M., Martic, M., & Radojicic, Z. (2011). A fresh approach to evaluating the Academic Ranking of World Universities. Scientometrics, 87(3), 587–596. doi: 10.1007/s11192-011-0361-6
Jeremic, V., Bulajic, M., Martic, M., Markovic, A., Savic, G., Jeremic, D., & Radojicic, Z. (2012). An evaluation of European countries health systems through distance based analysis. Hippokratia, 16(2), 175–179.
Jeremic, V., Jovanovic-Milenkovic, M., Martic, M., & Radojicic, Z. (2013a). Excellence with leadership: The crown indicator of SCImago institutions rankings IBER report. El Profesional de la Informacion, 22(5), 474–480.
Jeremic, V., Zornic, N., Jovanovic-Milenkovic, M., Markovic, A., & Radojicic, Z. (2013b). Proportion of collaborative publications as an ultimate indicator of Leiden 2013 world best universities rankings. Management, 18(68), 5–12. doi: 10.7595/management.fon.2013.0020
Jons, H., & Hoyler, M. (2013). Global geographies of higher education: The perspective of World University Rankings. Geoforum; Journal of Physical, Human, and Regional Geosciences, 46(3), 45–59. doi: 10.1016/j.geoforum.2012.12.014
Jovanovic, M., Jeremic, V., Savic, G., Bulajic, M., & Martic, M. (2012). How does the normalization of data affects the ARWU ranking? Scientometrics, 93(2), 319–327. doi: 10.1007/s11192-012-0674-0
Keung, J., Kocaguneli, E., & Menzies, T. (2013). Finding conclusion stability for selecting the best effort predictor in software effort estimation. Automated Software Engineering, 20(4), 543–567. doi: 10.1007/s10515-012-0108-5
Li, M., Shankar, S., & Tang, K.K. (2011). Why does the USA dominate university league tables? Studies in Higher Education, 36(8), 923–937. doi: 10.1080/03075079.2010.482981
Liu, N.C. (2013). The academic ranking of world universities and its future direction. In M. Marope, P.J. Wells, & E. Hazelkorn (Eds.), Rankings and accountability in higher education: Uses and misuses (pp. 23–39). Paris: UNESCO.
Luxbacher, G. (2013). World University Rankings: How much influence do they really have? The Guardian. Retrieved from http://www.theguardian.com/higher-education-network/blog/2013/sep/10/university-rankings-influence-government-policy
Mahsuli, M., & Haukaas, T. (2013). Sensitivity measures for optimal mitigation of risk and reduction of model uncertainty. Reliability Engineering & System Safety, 117, 9–20. doi: 10.1016/j.ress.2013.03.011
Marginson, S. (2011). Higher education and public good. Higher Education Quarterly, 65(4), 411–433. doi: 10.1111/j.1468-2273.2011.00496.x
Monferini, A., Konstandinidou, M., Nivolianitou, Z., Weber, S., Kontogiannis, T., Kafka, P., Demichela, M. (2013). A compound methodology to assess the impact of human and organizational factors impact on the risk level of hazardous industrial plants. Reliability Engineering & System Safety, 119, 280–289. doi: 10.1016/j.ress.2013.04.012
Myrtveit, I., Stensrud, E., & Shepperd, M. (2005). Reliability and validity in comparative studies of software prediction models. IEEE Transactions on Software Engineering, 31(5), 380–391.
NTU ranking (2013). Performance ranking of scientific papers for world universities. National Taiwan University Ranking. Retrieved from http://nturanking.lis.ntu.edu.tw/Default.aspx
Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and rankings: Voodoo or science? Journal of the Royal Statistical Society. Series A, (Statistics in Society), 176(3), 609–634. doi: 10.1111/j.1467-985X.2012.01059.x
Pusser, B., & Marginson, S. (2013). University rankings in critical perspective. The Journal of Higher Education, 84(4), 544–568.
QS (2013). Quacquarelli Symonds (QS) World University Rankings. Retrieved from http://www.qs.com/ranking.html
Radojicic, Z., & Jeremic, V. (2012). Quantity or quality: What matters more in ranking higher education institutions? Current Science, 103(2), 158–162.
Rauhvargers, A. (2011). Global university rankings and their impact. Brussels, Belgium: European University Association.
Rauhvargers, A. (2014). Where are the global rankings leading us? An analysis of recent methodological changes and new developments. European Journal of Education, 49(1), 29–44. doi: 10.1111/ejed.12066
Safon, V. (2013). What do global university rankings really measure? The search for the X factor and the X entity. Scientometrics, 97(2), 223–244. doi: 10.1007/s11192-013-0986-8
Saisana, M., & D'Hombres, B. (2008). Higher education rankings: Robustness issues and critical assessment. How much confidence can we have in

higher education rankings? Ispra, Italy: EUR23487, Joint Research Centre, Publications Office of the European Union. doi: 10.2788/92295

Saisana, M., & Tarantola, S. (2002). State-of-the-art report on current methodologies and practices for composite indicator development. EUR Report 20408 EN, European Commission, JRC-IPSC, Italy.

Saisana, M., D'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. Research Policy, 40(1), 165–177. doi: 10.1016/j.respol.2010.09.003

Saltelli, A. (2007). Composite indicators between analysis and advocacy. Social Indicators Research, 81(1), 65–77.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity analysis in practice: A guide to assessing scientific models. Chichester, West Sussex: John Wiley & Sons, Ltd.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., & Tarantola, S. (2008). Introduction to sensitivity analysis. In A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, S. Tarantola (Eds.), Global sensitivity analysis. The Primer (pp. 1–51). Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/9780470725184.ch1

Seke, K., Petrovic, N., Jeremic, V., Vukmirovic, J., Kilibarda, B., & Martic, M. (2013). Sustainable development and public health: Rating European countries. BMC Public Health, 13(77), 1–7. doi: 10.1186/1471-2458-13-77

Shepperd, M., & Kadoda, G. (2001). Comparing software prediction techniques using simulation. IEEE Transactions on Software Engineering, 27(11), 1014–1022.

Soh, K. (2013). Rectifying an honest error in World University Rankings: A solution to the problem of indicator weight discrepancies. Journal of Higher Education Policy and Management, 35(6), 574–585. doi: 10.1080/1360080X.2013.844670

Soh, K. (2014). Nominal versus attained weights in Universitas 21 ranking. Studies in Higher Education, 39(6), 944–951. doi: 10.1080/03075079.2012.754866; Published online.

Taylor, P., & Braddock, R. (2007). International university ranking systems and the idea of university excellence. Journal of Higher Education Policy and Management, 29(3), 245–260. doi: 10.1080/13600800701457855

Tayyab, S., & Boyce, A.N. (2013). Impact factor versus Q1 class of journals in World University Rankings. Current Science, 104(4), 417–419.

THE (2013). Times Higher Education World University Rankings. Retrieved from http://www.timeshighereducation.co.uk/world-university-rankings/2013-14/world-ranking

Usher, A. (2014). When the Times higher education rankings fail the fall-down-laughing test. Retrieved from http://higheredstrategy.com/when-the-times-higher-education-rankings-fail-the-fall-down-laughing-test/

van Vught, F. (2008). Mission diversity and reputation in higher education. Higher Education Policy, 21(2), 151–174.

Wainwright, H.M., Finsterle, S., Zhou, Q.L., & Birkholzer, J.T. (2013). Modeling the performance of large-scale CO_2 storage systems: A comparison of different sensitivity analysis methods. International Journal of Greenhouse Gas Control, 17, 189–205. doi: 10.1016/j.ijggc.2013.05.007

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., van Eck, N.J., Wouters, P. (2012). The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. Journal of the American Society for Information Science and Technology, 63(12), 2419–2432. doi: 10.1002/asi.22708