



A new approach to the analysis and evaluation of the research output of countries and institutions

Domingo Docampo, Jean-Jacques Bessoule

► To cite this version:

Domingo Docampo, Jean-Jacques Bessoule. A new approach to the analysis and evaluation of the research output of countries and institutions. *Scientometrics*, 2019. hal-02104439

HAL Id: hal-02104439

<https://hal.science/hal-02104439>

Submitted on 19 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A new approach to the analysis and evaluation of the research output of countries and institutions

Domingo Docampo⁽¹⁾ and Jean-Jacques Bessoule^(2,*)

Short Title: The scientific production of countries and institutions.

⁽¹⁾University of Vigo, atlantTIC Research Center for Communications Technologies, Campus Universitario, 36310 Vigo, Spain.

⁽²⁾CNRS-Univ. Bordeaux, Laboratoire de Biogenèse Membranaire, UMR 5200, Villenave d'Ornon, France.

* Corresponding author: J.-J. Bessoule, UMR 5200 CNRS – Univ. Bordeaux, Bâtiment A3, INRA Bordeaux Aquitaine, 71 Avenue Edouard Bourlaux, CS 20032, 33140 Villenave d'Ornon

e-mail address : jean-jacques.bessoule@u-bordeaux.fr; Tel: (33) 5 57 12 25 70

Key Words: Research Output; Bibliometric indicators; Countries; Institutions; Publications; Citations.

Abstract

A plethora of bibliometric indicators is available nowadays to gauge research performance. The spectrum of bibliometric based measures is very broad, from purely size-dependent indicators (e.g. raw counts of scientific contributions and/or citations) up to size-independent measures (e.g. citations per paper, publications or citations per researcher), through a number of indicators that effectively combine quantitative and qualitative features (e.g. the h-index). In this paper we present a straightforward procedure to evaluate the scientific contribution of territories and institutions that combines size-dependent and scale-free measures. We have analysed in the paper the scientific production of 189 countries in the period 2006-2015. Our approach enables effective global and field-related comparative analyses of the scientific productions of countries and academic/research institutions. Furthermore, the procedure helps to identifying strengths and weaknesses of a given country or institution, by tracking variations of performance ratios across research fields. Moreover, by using a straightforward wealth-index, we show how research performance measures are highly associated with the wealth of countries and territories. Given the simplicity of the methods introduced in this paper and the fact that their results are easily understandable by non-specialists, we believe they could become a useful tool for the assessment of the research output of countries and institutions.

Acknowledgments

We thank Dominique Dunon-Bluteau and Daniel Egret for initiating the human link between the authors. We are grateful to Paul Gouguet, Amélie Bernard and Pierre Madre for critical reading of the manuscript. The work of D. Docampo was supported by the European Regional Development Fund (ERDF) and the Galician Regional Government under an agreement for funding the Atlantic Research Center for Information and Communication Technologies (AtlantTIC).

1.Introduction

For several years, many powerful indicators have been suggested to evaluate individual research production, Schreiber et al. (2012). Wildgaard et al. (2014) reviewed 108 of them. Some of those bibliometric indicators have been used to define academic productivity in terms of citation rates not only of individuals but also of whole departments (Garner et al., 2018). Arguably, the h-index (Hirsh, 2005) and the g-index (Egghe, 2006) are the most widely used today to assess individual scientific productions. Several indicators also exist for journals, such as the well-known Impact factor and Scimago Journal Rank (Leydesdorff, 2009). Still, a fair research question of interest for academics, policymakers and the public at large, is how to evaluate the research output of a country or institution. A number of studies have tackled the scientific production of countries. The European Commission (2013) issued a report covering the scientific production profiles of some 42 countries based on a selected set of bibliometric indicators including raw citation and publication counts along with some normalized global and relative impact measures. The report acknowledges the difficulties associated with the use of multiple indicators to determine the relative position of countries without a well-structured ranking mechanism in place. In an attempt to overcome those difficulties, Nejati and Hosseini Jenab (2010) introduced two-dimensional plots of normalized publications per population and citations per paper, instead of raw counts of citations and publications. The plots enabled the authors to carry out a comparative analysis of the scientific production in the Natural Sciences of 50 countries. Hosseini Jenab (2016) later applied the two-dimensional plots to analyse the ranking positions and track the temporal evolution of countries in two specific fields: physics and astronomy. Aksnes et al. (2012) studied the effect on relative citation indicators when citations to documents are fractionalised among the authoring countries and found that relative citation indicators are lower when fractionalised counting is used. Bornmann and Leydesdorf (2013) analysed normalized citation impact values of six large countries along a 30-years period, and found statistically significant trends mostly associated with the internationalization of research in some large continental European countries. Mueller (2016) reports on the use of a machine learning procedure to test a model based upon sixteen macro-level predictors representing “aspects of the research and economic system, of the political conditions, and of structural and cultural attributes of countries” to forecast countries research output across scientific disciplines.

But, to evaluate the research output of a country or institution, the most widely used indicators are those found in databases such as Web of Science/Incites. For instance, the number of outputs, the number of citations per publication, the number of papers published in the 25% of journals with the highest impact factor for a given research field (Q1), or the number of papers in the Top-10% of the most cited papers for a given research field (Top-10) (see Clarivate Analytics 2918a). In addition, the number of highly cited papers (HCP), selected

by year and field via the use of citation thresholds based upon the distribution of citations, is also a very informative indicator of both individual and institutional impact (see Clarivate Analytics 2018b).

However, a quick look at the list of countries and territories worldwide, along with the data we can gather from the Web of Science shows that evaluating the research output is not such a straightforward task. Take for instance the case of the Belize which, among 189 countries/territories analysed in this paper, would be ranked 144th if only the number (N) of Web of Science documents were included, but would be ranked 1st in the ratio citations/paper (C/N).

It is widely accepted in the bibliometric community that publication and citation measures refer to quantifiable features of research performance in a statistically reliable manner when sufficiently large and preferably longitudinal data sets are available for analysis (Glanzel et al., 2016). It is apparent that the special case of Belize is connected with the scarcity of publications that prevents from the smoothing of ratios, thus resulting in statistically unlikely data. Our goal is not to solve the conundrum of the quantity-quality debate, but rather to help in making comparisons and associations when the datasets under analysis are sufficiently large, e.g. comparing the research outputs in the period 2006-2015 of France (N = 808,198 documents, C/N = 17.75) and Japan (N = 962,931, C/N = 13.48), which have a different respective rank if N or C/N is considered. Hence, we aimed to define a bibliometric indicator at macro-levels that combines size-dependent measures, such as the number of documents or citations, with size-independent parameters usually based upon ratios between the number of citations and documents.

An indicator based upon the product of a quantitative and a qualitative parameter was introduced many years ago by Lindsey to assess the scientific production of individual researchers in the Social Sciences (Lindsey 1976, 1978). Glanzel and Moed (2002) pointed to the lack of interpretability in Lindsey's approach as one reason why this indicator has found no application. However, in spite of the fact that his papers were not widely cited, it must be noted that Lindsey's indicator was shown to be useful to make the h-index sensitive to hyper-cited articles (Tahira et al., 2014), and to improve cluster analysis of citation history (Luzar et al., 1992). In 2010, Prathap revisited the Lindsey's indicator defined as $CQ = \sqrt{C^3/N}$. For Prathap, "every citation is actually a paper that cites the publication and has the same dimensions as the h-index (or N)". "Thus, the total received citations C which sums over N has the dimension of area i.e., h^2 or N^2 ". Since "CQ does not have the dimensionality of h", Prathap brought it back to the dimensionality of h by introducing a transformation leading to the p-index = $\sqrt[3]{C^2/N}$ (Prathap, 2010a). Prathap used this index to rank countries (Prathap, 2010a) or individuals such as the most prolific economists (Prathap, 2010b), while Gupta (2010) used it to rank the Top 50 Indian Universities. Yan et al., (2016) analyzed the correlation of 28 indicators with the h-index in *Scientometrics*, *Journal of Informetrics*, an *Information Processing and Management*. For the p-index, these correlation coefficients were 0.865, 0.889 and 0.819 respectively. Nevertheless, in the present study, the number of citations for Top-189 countries appears linearly correlated with the number of publications

(Figure 1). Similarly, in Prathap (2010a, 2011), it appears that, at least for Western countries, the C/N ratio did not greatly vary as a function of N . In addition, in the same database used in Prathap (2010a)¹, countries as China for example no longer appeared as outliers in 2015, and it can be shown that the number of citations C received by the N papers published in 2015 by the Top-189 countries (*i.e.* countries with $N \geq 20$, excluding Gibraltar and French territories: Guadeloupe, Martinique, French Polynesia and French Guiana) is $C = 5.94N - 1375$, with $R^2 = 0.977$ (data not shown).

In the present paper, we try to overcome the shortcomings associated with the use of multiple bibliometric measures by proposing a procedure to evaluate a global research output of countries and institutions through the use of a straightforward measure, Global Research Output index (GRO-index), that is computed as the geometric mean of a quantitative and a qualitative parameter. The index can also be computed for papers within a specific research field, GROr. Both indexes, along with relative indexes, normalized to world absolute measures, Relative Research Output index (RRO-index) and corresponding field-specific RROr, will be introduced in Section 2. Cabrerizo et al., (2010) had also proposed an index based on the geometric average of a quantitative (the h -index) and a qualitative parameter (the m -index) for bibliometric studies at a micro (*i.e.* individual) level. These authors underlined the fact that the use of geometric averages displays several advantages: “it is easy to compute, it is easily understandable in geometric terms, it is not influenced by extremely higher values, and thus, it obtains a value which fuses the information provided by the aggregated values in a more balanced way than other aggregation operators” (Cabrerizo et al., 2010). Although these authors also make use of geometric means, they need to compute two complex indicators first, h and m , whereas our procedure deals directly with simple counts of publications and citations.

We have analyzed institutions and countries, and we further show that for countries and institutions, the sums of the GROr- and RROr- indexes computed research field by research field are very efficient linear predictors of the GRO- and RRO- indexes when all fields are included. That enabled us to analyse the research output of the Top-56 countries research field by research field, and to highlight how the research outputs of the various countries are greatly related to geographical and historical contexts.

Additionally, Abramo and D’Angelo (2016) warned about taking size-independent citation indicators per se as indicators of research performance, unless they are placed in context through accompanying measures of expenditures on research and or researchers. The use of ratios may be shown to violate basic economic reasoning accepted facts such as “that the better performer under parity of resources is the actor who produces more; or under parity of output, the better is the one who uses fewer resources” (Abramo and D’Angelo, 2016). Hence, we would also like to connect bibliometric information with appropriate measures of expenditures or wealth, to shed light on the relative performance and efficiency of countries and territories. We further compared the relation of GRO-index with the usual bibliometric

¹ <https://www.scimagojr.com/countryrank.php>

indicators provided by Web of Science/Incites, as well as with the p-index. Some advantages of the GRO-index will be highlighted in Section 3.

The paper is structured as follows. Section 2 describes the materials and methods used in our investigation. In Section 3 we present and discuss our results. Our conclusions are then presented in section 4 to close the paper.

2. Material and Methods

2.1 Data

Raw bibliometric data for our analysis were extracted from the InCites platform, provided by Clarivate Analytics Philadelphia, Pennsylvania, USA (2017), under the Essential Science Indicators scheme, including articles and reviews from Science Citation Index Expanded and Social Science Citation Index. The dataset does not include publications from Arts & Humanities, Conference Proceedings Citation Index, or Book Citation Index. Our raw data then comprises articles and reviews published between 2006 and 2015. Only countries showing more than 50 documents in that period were analysed. In addition, England, Northern Ireland, Wales and Scotland were removed from the analyses to avoid redundancy with the data from the United Kingdom. For the same reasons, results for Netherlands Antilles and French territories (New Caledonia, Reunion, French Guiana, French Polynesia, Guadeloupe and Martinique) were not included. Nevertheless, it must be noted that following all analyses, these 11 entities did never appear as outliers, and results would be extremely similar if considered. Subsequently to the treatment of our data, a total of 189 countries were analysed. We also analysed up to 4,556 Institutions as Table 1 shows. To avoid redundancy, the 76 University Systems were not included

Financial data were gathered from the International Monetary Fund's World Economic Outlook (IMF, 2017) on 2018, January 4th. Gross domestic products at current prices (GDP values) are based upon GDP in national currency converted to U.S. dollars using market exchange rates (yearly average). For the present analysis, the variable GDP corresponds to the average of the values of GDP along the period 2006-2015. Gross domestic product based on purchasing-power-parity per capita, GDP at current international dollars (PPC) were expressed in GDP in PPP dollars per person. Data are derived by dividing GDP in PPP dollars by total population. For the present analysis, the variable PPC corresponds to the average of the values of GDP (PPP) per capita along the period 2006-2015. From the two financial variables collected, a wealth index, WTH, has been composed as the geometric mean of GDP and PPC. The (ISO alpha-3) three letter codes used to designate the countries that appear in some of the plots later in the paper are shown in Table 2.

2.2 Rationale

To evaluate the research output of a country or institution, we consider two quantitative parameters, namely N, number of articles or reviews published between 2006 and 2015, and

C, number of citations accrued to those papers between 2006 and 2017. A plot of C vs N in a double logarithmic scale is shown in Figure 1. Since the value of the slope in Figure 1 is very close to 1, C and N appear to be linearly correlated. Any of these parameters could, therefore, be *a priori* chosen as the quantitative parameter to evaluate research throughput. Nevertheless, in order not to favour one parameter over the other, the geometric mean of N and C will be considered as the quantitative parameter of the scientific production of a country or institution: $Q = \sqrt{NC}$.

The second parameter we use to evaluate the Global Research Output (GRO) of a country or institution is a qualitative measure, i.e. number of citations per publication, $q = C/N$. Now, to combine the quantitative and qualitative measures, Q and q, we take the geometric mean of both parameters, producing the GRO-index $= \sqrt{Q \cdot q} = (C^3/N)^{1/4}$. The p-index, $\sqrt[3]{C^2/N} = \sqrt[6]{C^2(C^2/N^2)}$, can be also regarded as the geometric mean between a quantitative parameter $(C^{2/3})$ and a qualitative one $(C/N)^{2/3}$, but GRO appears more balanced since it aggregates both the number of publications and the number of citations within the quantitative parameter, thus giving equal importance to both C and N in the computation of the index.

To analyse the research output constrained to a research field r, we use a slight variation of the GRO-index. We call this field-based index GRO_r. The computation is the same as the one used for the GRO-index but for the fact that only papers within the research field are counted. To analyse the research output within the same country or institution across research fields, we use a slight variation of the GRO_r-index. We call this normalized field-based index RRO_r. It is calculated as the GRO_r-index, but normalizing the qualitative parameter for the institution or country x and research field r, the ratio $qr_x = Cr_x/Nr_x$, by the qualitative parameter for the world and same research field r, the ratio $qr_w = Cr_w/Nr_w$. An RRO-index, all fields included, can be also computed. It is straightforward to prove that RRO is related with GRO as Equation 1 shows, where $q_w = C_w/N_w$, for all the papers, N_w , produced in the world, and the citations, C_w , accrued to those papers:

$$RRO = q_w^{-1/2} GRO \quad (1)$$

To gain more insight into these indexes, it can be noted that in the case of a product sold in the market, we compute the total value (V) of the production by multiplying the number of units produced (quantity) by their price (quality). By analogy, an indicator of the value of the research output of a country or institution could be based upon quantitative and qualitative parameters related with their bibliometric output. Moreover, following up on the marketplace analogy, suppose that company A produces high-quality bikes and sells them at twice the worldwide average selling price of bikes. Another company, B, produces low-quality cars and sells them at half the worldwide average selling price of cars. Because of the relative prices of bikes and cars, for the same number of units sold, company A will have a lower market value of production than B, in spite of the higher quality of the units produced by A. Similarly, the most commonly used qualitative measure of the research output of countries or institutions, C/N , may greatly vary depending on the research field r; it would, therefore, be

very useful to introduce as a field-related quality parameter the ratio of institutional or country Cr_x/Nr_x values to world Cr_w/Nr_w measures by research field.

3. Results and Discussion

3.1 Ranking of Institutions and Countries

Table 3 shows, all fields included and research field by research field, the total number of publications in the world for the period 2006-2015, N_w , and the number of citations they received, C_w . We then computed $Q_w = \sqrt{N_w C_w}$, $q_w = \frac{C_w}{N_w}$, and the GRO-indexes for the world, all fields included, (GRO_w) , as well as research field by research field, (GRO_{r_w}) .

From the results shown in Table 3, it is apparent that the qualitative parameter qr_w shows a great deal of variability, from 6.12 (Mathematics) up to 33.87 (Molecular Biology & Genetics). Next, in Table 4, we show the results of the computation of the GRO-index for high performance research institutions. Table 4 include results for 34 institutions worldwide with a GRO-index in excess of 3,000

Figure 2 shows the GRO-index of all the countries (log scale) ordered from highest to lowest. The name and the GRO-index of the Top-56 countries are shown in Table 5. Table 5 also shows the values of a wealth-related indicator that shall be introduced and explained in Section 3.2. As can be seen in Figure 2, if we exclude the highest performer (USA) the Top-189 countries can be classified into four groups. The first three groups would share the characteristic of showing a linear relationship between $\ln(GRO)$ and ranking position. The remaining group of 35 countries displaying the lowest GRO values cannot be “linearized”.

The three groups that can be linearized are defined with respect to GRO values as follows: $GRO > 4,000$, $1,000 < GRO < 3,500$, and $150 < GRO < 850$. The boundary between the first and the second group is clear since countries with a GRO-index between 4,000 and 3,500 cannot be found in the list. The boundary between the second and the third group is also clear, as there is no country with a GRO-index between 1,000 and 850. The number of countries within the three groups is 16, 38, and 99, respectively. It is worth highlighting that the scattering of the countries into the three groups obeys a Bradford’s type law, as it should be expected.

3.2 Sums of the GROr-and RROr- indexes: efficient linear predictors of the GRO- and RRO-indexes.

In addition to GRO- and RRO-, GROr- and RROr- indexes can be used to highlight specific strengths within a given institution or country, by computing both indexes within a given research field r . However, the sums of the RROr- and GROr- indexes obtained for all the Research Fields are not equal to RRO- and GRO- indexes due to the non-linearity in the computation of the indexes. For instance, if we use the last column of Table 3 (values of the GRO-index as well as all the field-related GROr-indexes for the world) we can see that $GRO_w = 0.231 \text{sum}(GRO_{r_w})$.

However, to add up all the GROr field indexes together makes sense, since there is no overlapping among the 22 ESI research fields, the relevant question being whether that sum over the 22 ESI fields constitutes a reasonable approximation of the global GRO-index. To explore the answer to this important question we have collected values from a large number of institutions. Academic institutions are in general comprehensive and thus constitute natural candidates to check the validity of the approximation. We analysed 2,766 institutions classified by InCites as “Academic”, a group composed mainly by universities. A strong ($R^2 = 0.983$) linear relationship between GRO and sum (GROr) was observed. The slope of the linear regression line, **0.249**, is very close to the value of the world ratio, $GRO_w/\text{sum}(GROr_w) = 0.231$. It is, therefore, apparent that in spite of the inherent nonlinearities of the GRO-indexes, they can be reasonably approximated for academic institutions by adding GROr values across research fields. However, for other institutions, although still very high, the correlation coefficient was a little lower: $R^2 = 0.938$, and the slope a little higher (**0.289**), prompting us to further investigate the reasons behind these differences.

By considering institutions focused in just one major research field, the more specialised the institution is, the higher its $\text{max}(GROr/\text{sum}(GROr))$ becomes. We thus took $SGr = \text{max}(GROr/\text{sum}(GROr))$ as a specialization index of an institution, and analysed the variation of the $GRO/\text{sum}(GROr)$ ratio as a function of SGr . It appears that the ratio increases with SGr and that the slope of the straight line is lower for academic institutions (Figure 3a) than for “other” institutions (Figure 3b). Overall, academic institutions are by and large more comprehensive and thus show much lower specialization indexes, SGr , than institutions classified as “other”. It is worth noticing that if we used the sum of the two highest ratios $GROr/\text{sum}(GROr)$ instead of the highest one only, the conclusions would be the same (data not shown). We may safely conclude that differences in the correlation coefficients and in the slopes of the straight lines $GRO/\text{sum}(GROr)$ were observed because academic institutions are globally less specialised than “other” Institutions. In fact, only 5.8% of Academic institutions display a specialisation index, SGr , higher than 0.3 compared to 39.6% for “other” institutions, while 107 “other” Institutions (7.3%) display a specialisation index higher than 0.5, compared to 21 for Academic Institutions, (0.76%). Moreover, 8 of those 21 academic institutions could be regarded as “Research Institutes”, and as such could, therefore, have been classified as “other” institutions. These eight institutions are the following: Institute of Physics of the Azerbaijan National Academy of Sciences; Instituto de Fisica Corpuscular; Yerevan Physics Institute; Institut des Hautes Etudes Scientifiques; Institut d'Optique Graduate School – Dublin; National Research Nuclear University -Moscow Engineering Physics Institute, Dublin Institute for Advanced Studies; National Centre for Physics - Pakistan. The same comment applies to another two institutions, namely the European Southern Observatory and the Warsaw University Observatory. In addition, it is not surprising to find among these 21 institutions, very specialised entities such as Ufa State Aviation Technical University, National Research Nuclear University, Paris School of Economics and Princeton Plasma Physics Laboratory.

Such an analysis of the correlation between **GRO** and **sum(GROr)** was repeated for the sample of 189 countries worldwide. In this case, the correlation reaches 0.999 (Figure 4a). The slope of the straight line (0.225) is very close to the $\text{GRO}_w/\text{sum}(\text{GROr}_w)$ ratio (0.231) calculated from results shown in Table 3. It is also close to the slope determined for academic institutions (0.249). Interestingly, for the 189 countries, the same correlation coefficient ($R^2 = 0.999$) and a similar slope of the straight line (0.219) are observed by plotting RRO as a function of **sum(RROr)** (Figure 4b). Given that the sums of the GROr and RROr are very efficient linear predictors of the GRO- and RRO- index respectively, research field scores constitute a share of the index and thus enable to profile the research output of countries. Moreover, given that $\text{RRO}/\text{GRO} = (\text{qw})^{-1/2}$ and, for countries, $\text{GRO} = a \text{ sum}(\text{GROr})$, and $\text{RRO} = b \text{ sum}(\text{RROr})$, it appears that any one of these indexes is sufficient to analyse their research output. However, to compare two very specialised institutions focusing on two research fields displaying very different q_{rw} values (as for example Mathematics and Molecular Biology & Genetics, see Table 3), the RRO-index rather than the GRO-index should be used.

3.3 Factor Analysis of GROr- indexes

To better understand the composition of the GRO-index in terms of its Field-constituents, we carried out an exploratory analysis on the set of GROr values for a large number of academic institutions (close to 2,000) worldwide. We found six principal components with an eigenvalue larger than one, which makes it difficult to reduce dimensionality and interpret the results at the same time. To gain more insight into the Field distribution we used a mixed approach. We first carried out a hierarchical cluster analysis over the set of 22 scores of the sample of academic institutions, *i.e.* their GROr values corresponding to the research fields described in Table 3. We found that the 22 fields can be adequately classified in five clusters, as shown in Table 6, in which the names of the clusters try to describe the areas inside. Now, because GRO-indexes can be aggregated, we computed the scores on the five aggregated sets of research fields according to the cluster compositions. We then analysed the data corresponding to the 56 countries shown in Table 5 and carried a principal component analysis to reduce dimensionality on the new five aggregated variables. We used the covariance matrix since data are commensurable. We found two principal components with eigenvalues in excess of 1.0 that account for more than 70% of the variance of the sample.

Figure 5 shows the location of the five variables (corresponding to the five clusters from Table 6) as well as the Top-countries in the two-dimensional principal components plot. It is clear that the first component is associated with the medlife, socsi and agrenv clusters on the right side and the chemateng cluster on the left side. So, the balance between the Natural Sciences and Engineering (excluding Physics as well as Agricultural and Environmental Engineering) with the Medical, Life, and Social Sciences defines this component. The second component appears to be linked to the Physical Sciences, including Space Science and Geophysics.

Except Japan, all Asian and North-African countries are found above the horizontal axis, and with the exception of Thailand, all of them are in the upper-left quadrant, corresponding to Chemistry, Materials Science, Mathematics, and Engineering. By contrast, all countries of former Soviet-Union and satellite countries are found below the horizontal axis, and with the exception of Hungary, Croatia and Estonia, they are plotted in the lower-left quadrant corresponding to Physics, Space Science, and Geosciences. Western countries, Latin American countries, as well as South Africa and Japan were close to the horizontal axis, and rather located at the right side of the vertical axis (Life Sciences plus Human and Social Sciences). Chile, located at the bottom of the lower-left quadrant is an exception. This is due to its strong involvement in Space Sciences: 14% of $\text{sum}(\text{GROr})$, in comparison with all other countries (between 0.26% and 7.1%). This can be easily explained by the presence of many high-end astronomical observatories in this country.

3.4 On the relationship between GRO-indexes and wealth

It is a well-documented fact that some bibliometric measures are highly sensitive to the wealth of countries and territories. Data taken from the last UNESCO Science Report on countries' GDP in 2007 and 2014, and number of publications per country in 2008 and 2014 (UNESCO, 2018) shows the positive association between the number of scientific publications and the GDP of a country or territory. Furthermore, the data shows that the correlation between both measures has increased from 2008 to 2014. Since GRO indexes are based on bibliometric variables, we wanted to explore the relationship between the wealth of countries and their GRO-indexes. As mentioned in Material and Methods, and as done for the research output of countries, the wealth of a country (WTH) was defined as the geometric mean of a quantitative parameter (GDP: Gross Domestic Product in US\$) and a qualitative one (PPC: GDP in PPP US\$ per capita), averaged in both cases over the period 2006-2015. The WTH Values of the Top-56 countries were shown in Table 5. With the exception of Kenya (high GRO/WTH ratio) and Saudi Arabia (low GRO/WTH ratio) which constitute apparent outliers, there is clear evidence supporting a strong linear relationship between $\text{LN}(\text{WTH})$ and $\text{LN}(\text{GRO})$, as Figure 6 shows.

To further checking the relationship between GRO and WTH we carried out the same analysis for the states of the USA, dodging this way spurious causes of correlation due to between-countries differences. Interestingly, with the exception of Maryland (home of the National Institute of Health, Howard Hughes Medical Institute, Johns Hopkins University...) and Massachusetts (Massachusetts Institute of Technology, Harvard and so forth) which appeared as outliers, Figure 7 shows that results for the states of the USA were consistent with the ones shown for countries in Figure 6. It comes as no surprise that the research output of a country/state is by and large commensurate with its wealth; whether wealth is the scientific progress driver or the other way around is a debate. Wealth and knowledge production appear nowadays so intertwined that it is very difficult to answer the question of which causes which, although the modern endogenous growth theory states that the stock of human capital is an endogenous source of technological change which determines the rate of

growth (Romer, 1990), thus solidly linking the production of knowledge with the wealth of nations.

3.5 Why use GRO- related indexes instead of the h-index?

In contrast with countries for which, to our knowledge, no h-indexes were available on the INCITES platform, the study of Institutions provided the opportunity to analyse correlations between the GRO- and h-indexes. By using data from the 4,556 Institutions analysed, it is apparent that the GRO- and h-indexes are highly correlated ($R^2 = 0.980$, not shown). If only data from the 1,205 Institutions displaying an h-index higher than 100 were used, the high correlation still holds ($R^2=0.977$, not shown). Even focusing on the 129 Institutions with a GRO-index higher than 2,000, the correlation is still very significant ($R^2 = 0.932$, Figure 8).

The h- and GRO-indexes of institutions display high correlation values, yet the GRO-index exhibits advantages vis-à-vis the h-index in order to analyse the scientific production of an institution or a country, namely:

- (i) Unlike the GRO-index, a given h-index would not increase when new publications are added to the sample, unless the number of their citations exceeds the h-value;
- (ii) The GRO-index is a much simpler and therefore transparent tool than the h-index considering its formula: $(C/N \sqrt{CN})^{1/2} = (C^3/N)^{1/4}$. Its limpidity enables its use to calculate scientific production of institutions and countries over several years, differentiating itself from the complexity of the h-index;
- (iii) Just by adding the number of publications and by adding the number of citations, it is easier to calculate the GRO-index than the h-index of a group of countries or Institutions. The same remark can be made for the analysis of a group of several research fields (e.g. the various research fields addressing Human and Social Sciences);
- (IV) As shown in Figure 9, in comparison with the h-index, the GRO-index appears to be more strongly correlated with other indicators such as Q1, Top-10, and HCP. If we considered the p-index, it appeared that it is even less correlated to these indicators than the h-index. It must be noted that in contrast with the p-, h- and GRO- indexes, Q1, TOP 10% and HCP greatly depend on the amount and quality of the scientific production of other countries/Institutions.

4. Conclusions

In this paper we have introduced a straightforward methodology to analyse the scientific contribution of territories and institutions that combines size-dependent and scale-free measures. Our methodology helps in dealing with the shortcomings associated with the use of multiple bibliometric measures by relying solely on the use of two streamlined indicators of global and field-related research performance, GRO and RRO, which are computed as the geometric mean of a quantitative and a qualitative parameter. We show in the paper how the use of these two measures enables effective global and field-related comparative analyses of the scientific productions of countries and academic/research institutions. However, some limitations related to the data gathered for our analysis should be acknowledged. We have

used information from a particular provider of bibliometric data, Clarivate Analytics. We know that there are a number of scientific publications that are not covered in the Web of Science and the InCites platform. Whether the addition of those publications would have affected the findings and conclusions of the paper is an open question that needs more research on the subject. Furthermore, as pointed by one of the anonymous reviewers of the paper, the GRO index gives equal weighting to qualitative and quantity measures. We could have used instead a quadratic qualitative term and a linear quantitative term. This approach would have led us to GRO^2 , a quadratic expansion of GRO that would amplify country differences. However, this could have the positive side effect of reconciling the measure with the dimension of N, becoming then homogeneous with the size of the country in terms of scientific production. We think that it will be worth exploring in our further work the avenues that this approach contributes to open.

One of the main strengths of the methodology introduced in the present paper comes from the fact that for universities and countries the sum of the GRO- RRO- field indexes over the 22 ESI fields shows an extremely accurate linear fit with the global GRO- RRO- indexes ($R^2 = 0.983$). This somehow surprising additivity property, not generally enjoyed by h-type indexes, helps identifying strengths and weaknesses of a given country or institution by tracking variations of performance ratios across research fields, thus contributing to enlighten our analysis of the higher education sector: universities do not appear so heterogeneous when assessed by the indexes introduced in the present paper.

Thanks to the "market value of industry production" analogy, the GRO-index can be easily explained and understood. This is a critical point, especially to reach a wider audience of non-specialists, including policy makers. The "market value of industry production" analogy clearly shows that research output could be regarded, either as a whole or research field by research field, as an end product. Moreover, we have compared the results of the GRO-Index with a straightforward wealth-index, and have been able to show how research performance measures are highly associated with the wealth of countries and territories. We have shown in our analysis how the research output measured following our methodology is greatly related to geographical- historical- and economic -contexts.

Reasonably combining quantitative and qualitative data is arguably a matter of great interest to inform decision- and policy-making within institutions or whole research systems. We have in our paper put together straightforward bibliometric measures associated with quantity and impact of the scientific production. Given the simplicity of our methodology and the fact that its results can be easily understood by non-specialists, we believe it could become a useful tool for the assessment of the research output of countries and institutions, much in the same way as the impact factor is for journals or the h-index for individuals.

Note:

Some of the results shown in the present paper have been already made public in the preprint arXiv: 1807.01049

References:

- Abramo, C., & D'Angelo, A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics*, 10(2), 646–651.
- Aksnes, D.W., Schneider, J. W., & Gunnarsson, M. (2012). Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, 6(1), 36–43.
- Bornmann, L., & Leydesdorf, L. (2013). Macro-Indicators of Citation Impacts of Six Prolific Countries: InCites Data and the Statistical Significance of Trends, *PLoS ONE* 8(2): e56768. doi:10.1371/journal.pone.0056768.
- Cabrerizo, F.J., Alonso, S., Herrera-Viedma, E., & Herrera, F. (2010). q^2 -Index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core. *Journal of Informetrics*, 4(1), 23–28.
- Clarivate Analytics (2017). Clarivate Analytics/Web of Sciences/In Cites Essential Science Indicators. <https://clarivate.com/products/essential-science-indicators/>
- Clarivate Analytics (2018a). InCites Indicators Handbook, downloaded from: <http://ipsience-help.thomsonreuters.com/inCites2Live/8980-TRS/version/default/part/AttachmentData/data/InCites-Indicators-Handbook-6%2019.pdf>
- Clarivate Analytics (2018b). ESSENTIAL SCIENCE - HIGHLY CITED PAPERS, <http://archive.sciencewatch.com/about/met/core-hcp/>
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.
- European Commission (2013). Country and Regional Scientific Production Profiles. <https://ec.europa.eu/research/innovation-union/pdf/scientific-production-profiles.pdf>, downloaded from the EC server on November 30, 2018.
- Garner, R.M., Hirsch, J.A., Albuquerque, F.C, & Fargen, K.M. (2018). Bibliometric indices: defining academic productivity and citation rates of researchers, departments and journals. *Journal of Neuro Interventional Surgery*, 10(2), 102-106.
- Glänzel, W., & Moed, H.F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171-193.
- Glänzel, W., Thijs, B., & Debackere, K. (2016). Productivity, performance, efficiency, impact. What do we measure anyway?: Some comments on the paper “A farewell to the MNCS and like size-independent indicators” by Abramo and D’Angelo. *Journal of Informetrics*, 10(2), 658–660.
- Gupta, B.M. (2010) Ranking and performance of Indian Universities, based on publication and citation data. *Indian J Sci Technol.*, 3, 837-843.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA*, 102(46), 16569–16572.
- Hosseini Jenab, S.M. (2016). Two-dimensional mapping of scientific production of nations in the fields of physics and astronomy. *S Afr J Sci.* 2016;112(5/6),Art. #2015-0119, 8 pages. <http://dx.doi.org/10.17159/sajs.2016/20150119>
- IMF (2017). World Economic Outlook (WEO)

- <http://www.imf.org/external/pubs/ft/weo/2017/01/weodata/index.aspx>, downloaded from the IMF webpage on October 17th, 29th018.
- Leydesdorff, L. (2009). How are New Citation-Based Journal Indicators Adding to the Bibliometric Toolbox? *J. Am. Soc. Inf. Sci. Technol.*, 60(7), 1327-1336.
- Lindsey, D. (1976). Distinction, Achievement, and Editorial Board Membership. *American Psychologist*, 31(11), 799-804.
- Lindsey, D. (1978). The Corrected Quality Ratio: A Composite Index of Scientific Contribution to Knowledge. *Social Studies of Science*, 8, 349-354.
- Luzar, V., Dobrić, V., Maričić, S., Pifat, G., & Paventi, J. (1992). A methodology for cluster analysis of citation histories. *Qual. Quant.*, 26, 337-365.
- Mueller, C.E. (2016). Accurate forecast of countries' research output by macro-level indicators. *Scientometrics*, 109 (2), 1307–1328.
- Nejati A., & Hosseini Jenab, S.M. (2010). A two-dimensional approach to evaluate the scientific production of countries (case study: the basic sciences). *Scientometrics*, 84(2), 357–364.
- Prathap, G. (2010a). Is there a place for a mock h-index? *Scientometrics*, 84(1), 153-165.
- Prathap, G. (2010b) The 100 most prolific economists using the p-index. *Scientometrics*, 84(1), 167–172.
- Prathap, G. (2011). Quasity, when quantity has a quality all of its own— toward a theory of performance. *Scientometrics*, 84(2), 555-562.
- Romer, P.M. (1990). Endogenous Technological Change. *The Journal of Political Economy*, 1990, 98(5), S71-S10.
- Schreiber, M., Malesios, C.C., & Psarakis, S. (2012). Exploratory factor analysis for the Hirsch index, 17 h-type variants, and some traditional bibliometric indicators. *Journal of Informetrics*, 6(3), 347– 358.
- Tahira, M., Alias, R.A., Bakri, A., & Shabri A. (2014). h-index, h-type Indices, and the Role of Corrected Quality Ratio. *J. of infosci. theory and practice*, 2(4), 20-30.
- UNESCO (2015), UNESCO Science Report: towards 2030, <https://unesdoc.unesco.org/ark:/48223/pf0000235406>, downloaded from the UNESCO server on February 26th, 2019.
- Wildgaard, L., Schneider, J.W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics*, 101(1), 125-158.
- Yan, Z., Wu, Q., & Li, X. (2016) Do Hirsch-type indices behave the same in assessing single publications? An empirical study of 29 bibliometric indicators. *Scientometrics*, 109(3), 1815-1833.

Figure Captions

Fig. 1 Total number of citations as a function of the total number of publications for the Top-189 countries

Fig. 2 $\text{LN}(\text{GRO})$ vs $\text{GRO} - \text{Rank}$ for the Top-189 countries

Fig. 3 $\text{GROr}/\text{sum}(\text{GROr})$ as a function of SGr for Academic Institutions and “Other” Institutions

a: Academic Institutions. b: Other Institutions

Fig. 4 GRO as a function of $\text{sum}(\text{GROr})$ and RRO as a function of $\text{sum}(\text{RROr})$ for 189 countries

a: GRO as a function of $\text{sum}(\text{GROr})$; b: RRO as a function of $\text{sum}(\text{RROr})$

Fig. 5 Scores on the two principal components (sample of 55 countries)

Kenya (outlier) was omitted. Italic, italic and underlined, underlined, and regular roman letters correspond to Asian, North-African, Latin American, and Western (+ South Africa) countries respectively. Bold letters correspond to countries of former Soviet-Union and satellite countries.

Fig. 6 $\text{LN}(\text{GRO})$ as a function of $\text{LN}(\text{WTH})$ for 54 countries among the Top-56

Kenya and Saudi Arabia were omitted.

Fig. 7 $\text{LN}(\text{GRO})$ as a function of $\text{LN}(\text{WTH})$ for 48 USstates

Maryland and Massachusetts were omitted. Dots correspond to countries shown in Figure 6 for which $6 > \text{LN}(\text{WTH}) > 3.5$

Fig. 8 h-index as a function of GRO-index for institutions with a GRO-index higher than 2,000

Fig. 9 $\text{LN}(\text{Q1})$, $\text{LN}(\text{Top10})$, and $\text{LN}(\text{HCP})$ as a function of $\text{LN}(\text{p})$, $\text{LN}(\text{h})$ and $\text{LN}(\text{GRO})$ for Institutions displaying a GROindex higher than 2,000.

a: $\text{LN}(\text{X})$ as a function of $\text{LN}(\text{p})$; b: $\text{LN}(\text{X})$ as a function of $\text{LN}(\text{h})$; c: $\text{LN}(\text{X})$ as a function of $\text{LN}(\text{GRO})$. $\text{X} = \text{Q1}$ (closed circles), Top10 (grey circles) or HCP (open circles).

Figure 1

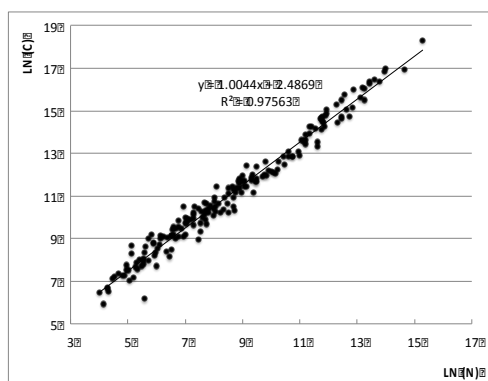


Figure 2

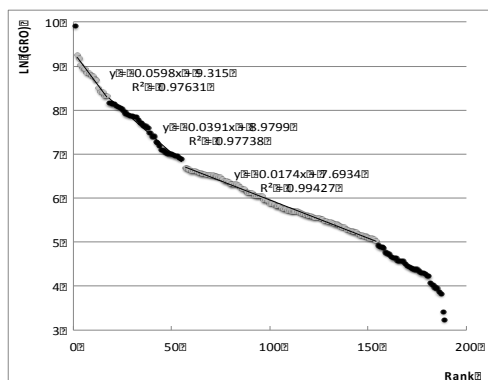


Figure 3

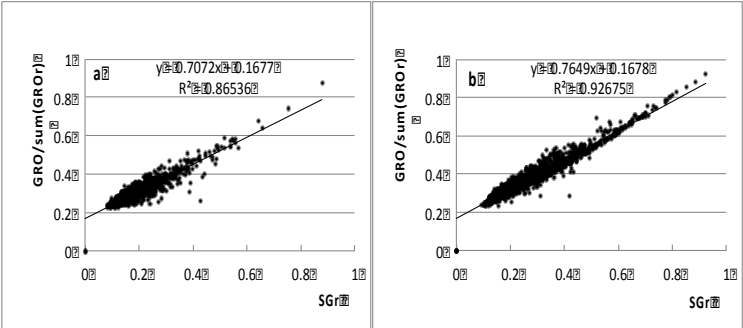


Figure 4

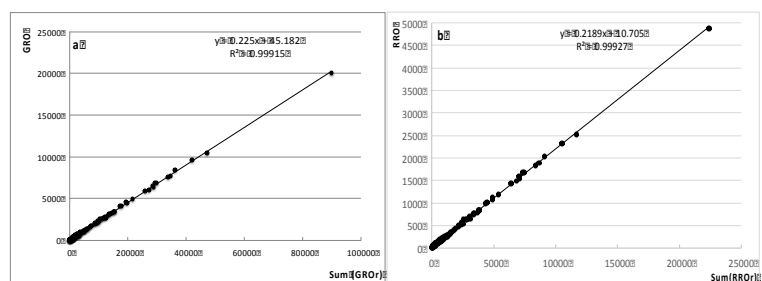


Figure 5

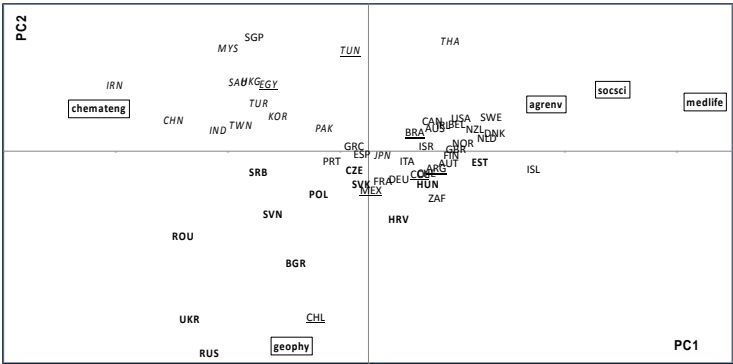


Figure 6

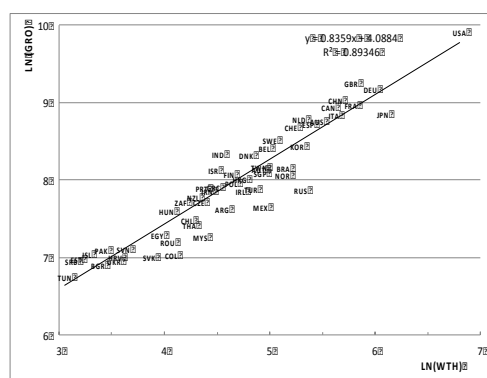


Figure 7

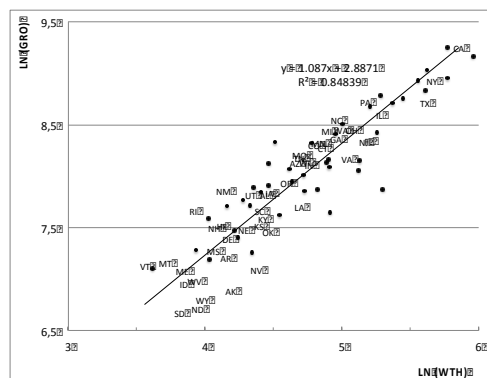


Figure 8

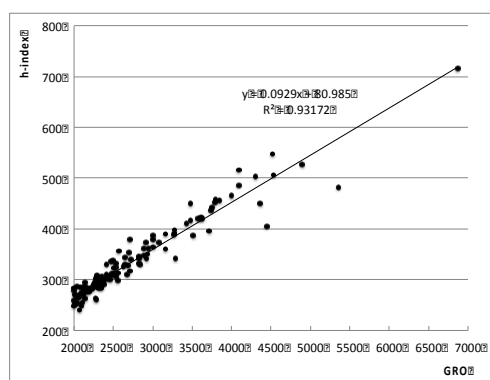


Figure 9

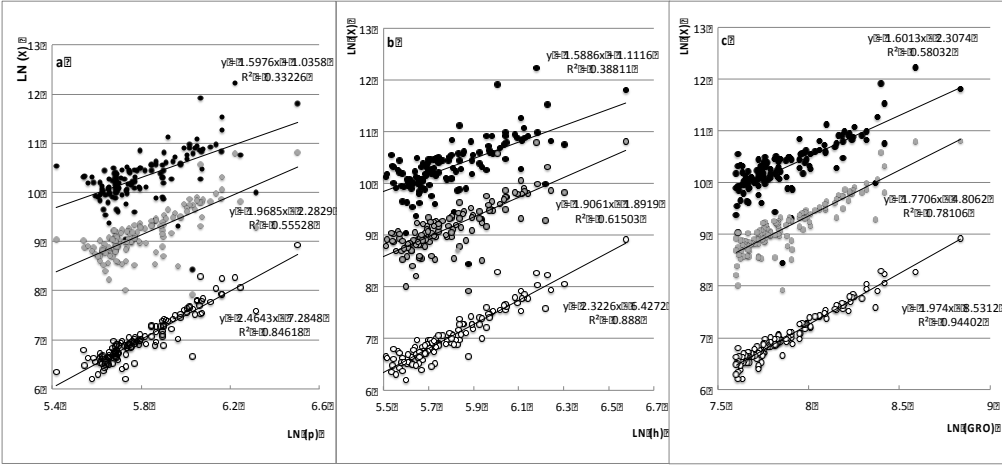


Table1 Institutions analysed in the paper

Type of Institution	Number
Academic (mainly universities)	2,965
Research Institutes	890
Health Institutions	331
Corporate Firms	186
Governmental organizations	136
Other institutions (Lab, Museum, Observatory...)	48

Table2 ISO three letter codes used in the paper

ARG: Argentina	EGY: Egypt	ISR: Israel	RUS: Russia
AUS: Australia	ESP: Spain	ITA: Italy	SAU: Saudi Arabia
AUT: Austria	EST: Estonia	JPN: Japan	SGP: Singapore
BEL: Belgium	FIN: Finland	KEN: Kenya	SRB: Serbia
BGR: Bulgaria	FRA: France	KOR: Korea	SVK: Slovakia
BRA: Brazil	GBR: United Kingdom	MEX: Mexico	SVN: Slovenia
CAN: Canada	GRC: Greece	MYS: Malaysia	SWE: Sweden
CHE: Switzerland	HKG: Hong Kong	NLD: Netherlands	THA: Thailand
CHL: Chile	HRV: Croatia	NOR: Norway	TUR: Turkey
CHN: China	HUN: Hungary	NZL: New Zealand	TUN: Tunisia
COL: Colombia	IND: India	PAK: Pakistan	TWN: Taiwan
CZE: Czech Republic	IRL: Ireland	POL: Poland	UKR: Ukraine
DEU: Germany	IRN: Iran	PRT: Portugal	USA: United States
DNK: Denmark	ISL: Iceland	ROU: Romania	ZAF: South Africa

Table3 Raw Data and Calculation of the GRO-index and the various GROr for the world

All Fields	N_w	C_w	Q_w	q_w	GRO_w
	12669278	213945356	52062781	16.89	29651
Research Fields	N_{rw}	C_{rw}	Q_{rw}	q_{rw}	GRO_{rw}
Agricultural Sciences	350,182	4214,897	1214,900	12.04	3,824
Biology & Biochemistry	655,603	15040,990	3140,210	22.94	8,488
Chemistry	1489,725	28295,481	6492,495	18.99	11,105
Clinical Medicine	2350,035	40702,524	9780,202	17.32	13,015
Computer Science	304,964	3505,363	1033,929	11.49	3,447
Economics & Business	229,392	2820,097	804,306	12.29	3,145
Engineering	1006,427	11974,992	3471,593	11.90	6,427
Environment/Ecology	370,443	6883,427	1596,846	18.58	5,447
Geosciences	377,291	6472,810	1562,733	17.16	5,178
Immunology	224,428	5737,206	1134,720	25.56	5,386
Materials Science	658,410	10850,762	2672,873	16.48	6,637
Mathematics	370,480	2267,589	916,568	6.12	2,369
Microbiology	177,598	3687,959	809,305	20.77	4,099
Molecular Biol. & Genetics	394,274	13354,013	2294,589	33.87	8,816
Multidisciplinary	16,759	355,231	77,158	21.20	1,279
Neuroscience & Behavior	452,541	11128,242	2244,100	24.59	7,429
Pharmaco. & Toxicology	340,426	5950,164	1423,232	17.48	4,988
Physics	1024,499	15784,117	4021,295	15.41	7,871
Plant & Animal Science	644,911	8171,135	2295,573	12.67	5,393
Psychiatry/Psychology	339,808	5959,524	1423,058	17.54	4,996
Social Sciences, general	758,582	7542,918	2392,054	9.94	4,877
Space Science	132,500	3245,915	655,808	24.50	4,008

Table 4 Rank, Name, GRO-index values of the Top-34 Institutions

Rank	Name	GRO
1	Harvard University	6,875
2	Centre National de la Recherche Scientifique (CNRS)	5,353
3	National Institutes of Health (NIH) - USA	4,902
4	United States Department of Energy (DOE)	4,525
5	VA Boston Healthcare System	4,522
6	Chinese Academy of Sciences	4,444
7	Max Planck Society	4,360
8	Howard Hughes Medical Institute	4,306
9	Stanford University	4,096
10	Massachusetts Institute of Technology (MIT)	4,088
11	Johns Hopkins University	3,994
12	University of Toronto	3,841
13	University of California Berkeley	3,794
14	University of Washington Seattle	3,780
15	University of California Los Angeles	3,753
16	University of Oxford	3,735
17	Institut National de la Sante et de la Recherche Medicale (Inserm)	3,712
18	University of Michigan	3,634
19	University of California San Francisco	3,606
20	University of Pennsylvania	3,606
21	University of Cambridge	3,574
22	University College London	3,514
23	Massachusetts General Hospital	3,482
24	Columbia University	3,473
25	University of California San Diego	3,434
26	Consejo Superior de Investigaciones Cientificas (CSIC)	3,289
27	Duke University	3,278
28	Yale University	3,276
29	Imperial College London	3,260
30	University of Chicago	3,162
31	Pierre & Marie Curie University - Paris VI	3,156
32	Cornell University	3,078
33	University of Pittsburgh	3,005
34	Washington University (WUSTL)	3,003

Table5 Rank, GRO-index of theTop-56 countries and their WTH values

Rank	Country	GRO	WTH	Rank	Country	GRO	WTH
1	USA	20,049	900.1	29	RUS	2,634	199.5
2	GBR	10,417	321.9	30	IRL	2,618	112.7
3	DEU	9,581	386.8	31	TUR	2,584	124.4
4	CHN	8,377	277.4	32	NZL	2,562	71.9
5	FRA	7,773	321.3	33	IRN	2,373	82.2
6	CAN	7,576	259.5	34	ZAF	2,239	63.9
7	JPN	6,903	434.8	35	CZE	2,234	75.8
8	ITA	6,874	273.7	36	MEX	2,099	136.2
9	NLD	6,555	196.7	37	ARG	2,038	94.2
10	AUS	6,370	232.1	38	HUN	1,979	55.8
11	ESP	6,081	214.3	39	CHL	1,767	67.7
12	CHE	5,878	182.0	40	THA	1,650	69.4
13	SWE	4,937	148.7	41	SAU	1,632	170.2
14	KOR	4,587	192.8	42	EGY	1,452	51.2
15	BEL	4,484	141.4	43	MYS	1,414	76.8
16	DNK	4,144	119.1	44	SVN	1,322	37.1
17	IND	4,115	91.0	45	PAK	1,215	30.1
18	AUT	3,500	132.8	46	ROU	1,197	56.4
19	BRA	3,472	169.2	47	ISL	1,138	26.2
20	TWN	3,419	134.5	48	COL	1,121	58.3
21	ISR	3,392	87.0	49	EST	1,099	23.7
22	SGP	3,250	135.7	50	HRV	1,092	34.4
23	FIN	3,201	101.0	51	SVK	1,072	47.4
24	NOR	3,144	167.9	52	SRB	1,043	22.6
25	HKG	3,011	112.0	53	UKR	1,042	33.9
26	POL	2,843	103.2	54	KEN	1,003	11.1
27	GRC	2,719	87.1	55	BGR	986	29.1
28	PRT	2,675	77.9	56	TUN	845	21.2

Table6 Cluster analysis on the 22 research fields' GRO values

Cluster	Research Fields
agrenv	Agricultural Sciences/ Environment & Ecology/ Plant & Animal Sciences
medlife	Biology & Biochemistry/ Molecular Biology & Genetics/ Clinical Medicine/ Microbiology
	Multidisciplinary/ Immunology/Neurosciences & Behavior/ Pharmacology & Toxicology
chemateng	Chemistry/ Materials Science/ Engineering/ Computer Science/ Mathematics
socsci	Psychiatry & Psychology/ Social Sciences, general/ Economics & Buisness
geophy	Physics/ Space Science/ Geosciences