

# Evaluation of the scientific production of countries by a resource-scaled two-dimensional approach

Seyyed Mehdi Hosseini Jenab\*, Ammar Nejati<sup>1</sup>

*\*Centre for Space Research, Potchefstroom Campus, North-West University, Potchefstroom, South Africa, <sup>1</sup>Institute of Physics and Bethel Center for Theoretical Physics, University of Bonn, D-53115 Bonn, Germany*

## ABSTRACT

The relative positions of country groups in the world of science are studied via application of a two-dimensional mapping method that is based on quantity and quality indicators of the scientific production as peer-reviewed articles. To obtain such indicators, different influential effects such as the background global trends, temporal fluctuations, disciplinary characteristics, and mainly, the effect of countries resources have been taken into account. Fifty countries with the highest scientific production are studied in 12 years (1996–2007). A common clustering algorithm is used to detect groups of co-evolving countries in the two-dimensional map, and thereby countries are classified into four major groups based on their relative positions in the map. The final results are contrasted with a non-resource-scaled map by considering some exemplary countries that have drastically different positions in the two maps.

**Keywords:** Countries' scientific production, quality-quantity measure, two-dimensional map, time evolution, clustering

## INTRODUCTION

Quantitative analysis of the productions of research units (countries, academic institutions, or research groups) has turned out to be one of the most significant issues since the early days of scientometrics. An important part of that analysis is devising methods and indicators to evaluate the multitude of aspects of such a complex production process, especially, its “quantity” and “quality.”<sup>[1,2]</sup> This quantitative assessment is becoming inevitable today, since such methods and findings are highly relevant for policy-makers of research units as well as academic agents directly involved in the process, to evaluate their achievements, and modify their approaches for further progress.

One of the essential problems in such analyses is that to provide a detailed picture from scores of different indicators, usually a host of graphs and tables are offered, and it is difficult to grasp a comprehensive picture or make an ultimate judgment from them.<sup>[3-6]</sup> As a remedy for this problem, composite indicators<sup>[7]</sup> and two-dimensional maps have been introduced.<sup>[8-10]</sup>

Along the same path, in this article, we have propounded a method to construct a two-dimensional map from two composite indicators – the first based on publications and the second based on citations thereof – aiming at a simultaneous representation and analysis of “quantity” and “quality” of the scientific productions of countries. However, we make certain adjustments in advance to take into account influential effects such as the background global or discipline-characteristic trends as well as temporal fluctuations, and mainly, the effect of different amount of resources accessible to countries. Regarding the latter, we have “scaled out” the effect of human and financial resources, and showed that accounting for them provides a different global pattern compared to the conventional ones found in the national or international reports on scientific production of countries.<sup>[3-6]</sup>

\*Address for correspondence:  
E-mail: mehdi.jenab@nwu.ac.za

### Access this article online

<b>Quick Response Code:</b>	<b>Website:</b> www.jscires.org
	<b>DOI:</b> 10.4103/2320-0057.153550

In this study, the topmost 50 countries in scientific production are considered in a period of 12 years (1996–2007). Since this period will certainly admit temporal fluctuations in positions, we implement a common statistical clustering method (*k-means* algorithm) to analyze the “resource-scaled” two-dimensional map. The resulting clusters (labeled by letters A, B, C, and D) indicate the relative positions of *country groups* compared to each other. We deliberately avoid the common convention of proposing some numeric “ranks” or “position markers” for countries, mainly because there seems to be an insignificant difference between countries belonging to a neighborhood in the range of values of an indicator; moreover, those ranks have proven to be volatile due to nonnegligible temporal fluctuations.

## METHOD

The proposed method involves a modified version of the previous scientometric indices to measure the quality and quantity of the explicit scientific production of a research unit in the form of journal papers. Usually, indices based on number of papers and their citations are used as measures of quantity and (academic) quality of the corresponding scientific production, respectively. From a set of quality–quantity indicators, we can obtain a graphical representation by plotting the quantity indicator on the horizontal axis, and the quality indicator on the vertical axis to produce a two-dimensional “quantity–quality map”. Here, we account for the effect of financial and human resources, background global trends on publications and citations, and discipline-specific characteristics of different academic fields by a simple procedure [Appendix].

As a quantity measure, we use ‘publication per resource’ (PPR), which is similar to the ‘publication per population’ (PPP) introduced by Kronman and Wadskog,<sup>[11]</sup> except that to have a resource-unbiased comparison between countries, we divide the publication of each country by a resource indicator ( $R$ ), which is a dimensionless combination of its gross domestic product (GDP) as a financial resource measure, and its population as a measure of human resources.

As an academic quality measure, we use citation per publication (CPP)<sup>[11]</sup> to obtain the average academic quality of a journal publication.

In order to dispense with the background global trend in publication and citation during the studied period, both

indicators (PPR and CPP) are divided by their respective annual average over all of the studied countries; hence PPR<sub>m</sub> and CPP<sub>m</sub>. These indicators have been obtained for the “general science” (science as a whole) and 27 major subfields of science (see section A6). Those subfields are categorized into six major fields and finally, the quality and quantity indicators for the general science and each major field are provided (PPR<sub>a</sub> and CPP<sub>a</sub>). The resulting data are comparable across different years and different scientific fields.

Finally, there will be 600 data points on the two-dimensional map for each field constructed from the corresponding PPR<sub>a</sub>-CPP<sub>a</sub> data of 50 countries during 12 years (1996–2007). However, this map is still too intricate to decipher, mainly due to the complexities of the temporal evolution of countries. Therefore, we have used a clustering analysis to find out groups of co-evolving countries.

The data needed for the analyses are obtained from the SCImago Journal and Country Rank portal,<sup>[12]</sup> which provides Scopus data arranged according to the country, subfield of science and year.<sup>[13]</sup> The population and GDP data are obtained from the World Development Indicators database.<sup>[14]</sup> To provide a more detailed picture, the 27 subfields of science are categorized into six major fields: (1) Natural Sciences (*Nat*); (2) Formal Sciences (*Frm*); (3) Engineering Sciences (*Eng*); (4) Health Sciences (*Hlt*); (5) Social Sciences (*Soc*); (6) Arts and Humanities (*Arh*). Our focus of study is the whole body of science; yet, to offer a more detailed picture which includes also a rough measure of the relative positions of countries in different fields, we apply the same method to the six fields of science to find the co-evolving country groups.

## RESULTS AND DISCUSSION

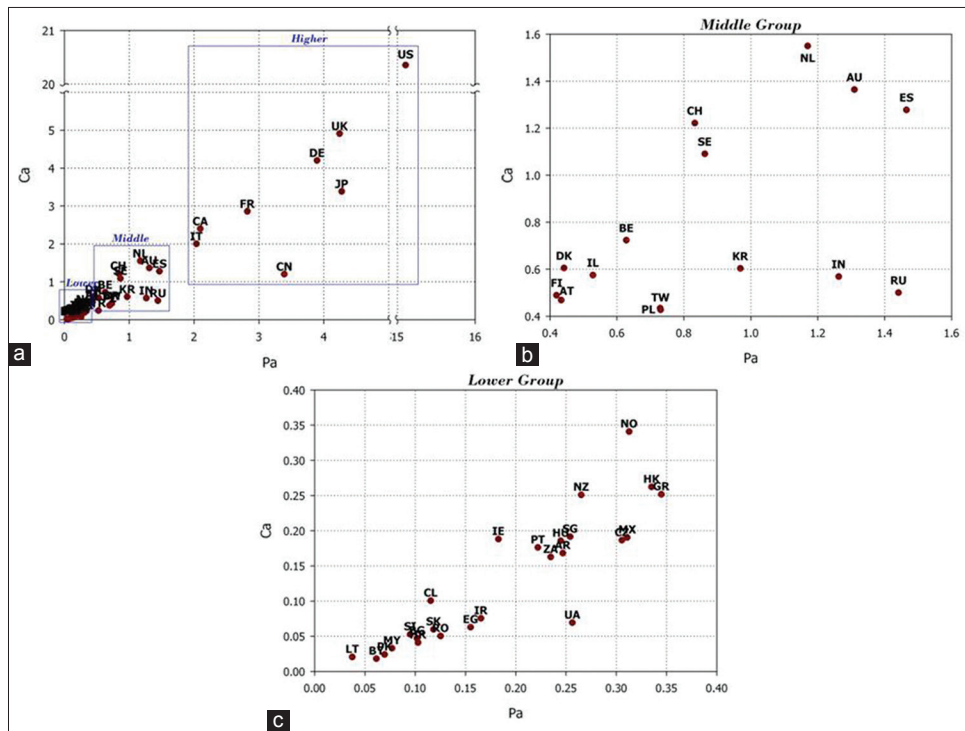
Before presenting the resource-scaled map and for the sake of comparison, we construct another map in which we have *only* corrected for the effect of the background global and field-characteristic trends on the quantity and quality indicators (i.e. the influence of resources has not been accounted for). Two indicators are used to construct the map, namely,  $P_u$  as the quantity and  $C_a$  as the quality indicator. The procedure to obtain these two indicators is analogous to PPR<sub>a</sub> and CPP<sub>a</sub>, except that we do not scale the number of publications  $P$  by the resource measure  $R$  and number of citations  $C$  by the number of publications (see section A7).

By plotting these indicators on a two-dimensional map ( $Pa-Ca$  map), we obtain quite a revealing picture. Huge gaps are evident between groups of collocated countries so that one could recognize 3 distinct groups (“high”, “middle”, and “low”). Figure 1 shows the detailed picture by three subfigures of different scale. In this figure, we have shown the temporal average of  $Pa$  and  $Ca$  for each country to avoid the complexity due to temporal fluctuations. The detailed information about the relative positions of countries in this map is presented in the last column of Table 1.

In the second step, we account for both the effects of the resources and global and discipline-specific trends using  $PPRa$  and  $CPPa$  as our quantity and quality indicators. In addition, we apply the clustering method to this map in order to extract the inherent geometrical structure; that is, the country clusters. The clustering validity indices indicate that the optimal clustering is composed of four clusters that we have labeled by letters A, B, C and D (see below). It is remarkable to note that the optimal clustering for the major fields reveals almost the same pattern as that for the whole science; that is, four clusters with the same relative positioning. This is especially evident in natural ( $Nat$ ), health ( $Hlt$ ) and engineering ( $Eng$ ) fields.

The four clusters, A to D, are representatives of the global quantity–quality classes [Figure 2]. Members of cluster A are in an excellent level of both quantity and quality. Cluster B includes countries with rather the same quality as cluster A, but lower quantity. Cluster C is composed of countries with almost the same quantity level as that of cluster B, but lower quality level; thus, cluster C has a lower level of both quality and quantity in comparison with cluster A. Members of cluster D are in the lowest level of both quality and quantity compared to all the other three clusters. Some countries are ‘transitional’ which means that they move from one cluster to another in the course of their temporal evolution, and thence, are marked by two (or more) cluster labels.

Table 1 shows the results of clustering for the general science together with those for the six major fields. Countries are arranged according to their cluster labels in general science and the major fields. It should be reminded that the position of a country *within* each general-science cluster is arbitrary to some extent; this is due to the fact that changing the ordering of fields or adopting other definitions for fields can alter the positions within each general-science cluster.



**Figure 1:**  $Pa-Ca$  map (which is not scaled by the resources) for the 50 topmost countries. Each point represents the average value of  $Pa$  and  $Ca$  of a country throughout the studied period (1996–2007). Since the points are scattered in a very wide range for both  $Pa$  and  $Ca$ , the map is presented in three different ranges (a-c). (a) Three country groups are recognized – “higher”, “middle” and “lower” – according to the positions in the map. Notice the huge gap between USA and other countries in this map; (b) the “middle” group; (c) the “lower” group [Table 1]

**Table 1: Clustering results for PPRa – CPPa map (which is scaled by the resources) for the “general science” and its 6 major fields**

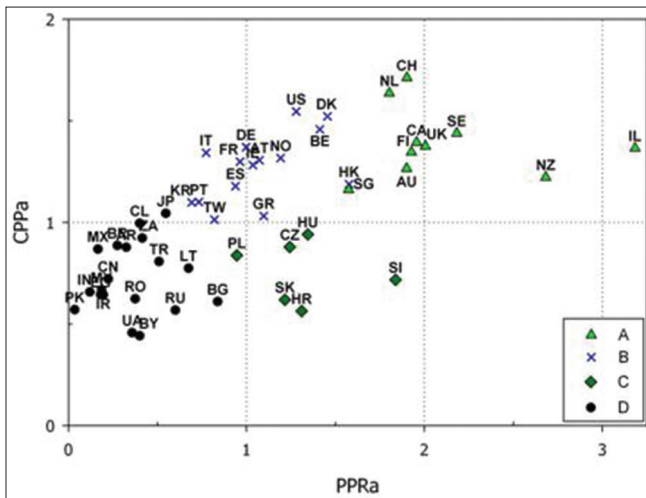
Name	Code	PPRa-CPPa map							Pa-Ca map
		General science	Science disciplines						
			Natural sciences	Engineering sciences	Health Sciences	Social sciences	Formal sciences	Arts and humanities	
Israel	IL	A	A	A	A	A	A	A	Mi
Sweden	SE	A	A	A	A	B	B	A	Mi
Finland	FI	A	A	A	A	B	C-B	A	Mi
Australia	AU	A	A	B	A	A	B	A	Mi
New Zealand	NZ	A	A	B	A	A	C	A	Lo
Canada	CA	A	A	B	A	B	C	A	Hi
Switzerland	CH	A	A	B	A	B-A	B	B	Mi
UK	UK	A	A-B	B	A	A	B	A	Hi
Singapore	SG	A	B	A	C-B	B	A	B	Lo
Netherlands	NL	A	B	B	A	B	B	A	Mi
Denmark	DK	B	A	B	A	B	B	A	Mi
Hong Kong	HK	B	B	A	B	A	C	A	Lo
Austria	AT	B	B	B	A	A	B	B	Mi
Belgium	BE	B	B	B	A	B	B	A	Mi
Germany	DE	B	B	B	B	A	B	B	Hi
France	FR	B	B	B	B	A	B	B	Hi
Italy	IT	B	B	B	B	A	B	B	Hi
Spain	ES	B	B	B	B	A	B	B-D	Mi
Portugal	PT	B	B	B	B	A	B	D	Lo
USA	US	B	B	B	B	B	B	A	Hi
Ireland	IE	B	B	B	B	B	B	A	Lo
Norway	NO	B	B	B	B	B	B	A-B	Lo
Greece	GR	B	B	B	C	A	C	A-B-D	Lo
Korea	KR	B	D	C	B	A	B-D	D	Mi
Taiwan	TW	B	D	C	B	A	C	B-D	Mi
Slovenia	SI	C	C	A	C	A	A	D	Lo
Poland	PL	C	C	C	B	D	C	D	Mi
Czech republic	CZ	C	C	C	C	A	C	B-D	Lo
Hungary	HU	C	C	C	C	A	C	B-D	Lo
Slovakia	SK	C	C	C	C	A	C	D	Lo
Croatia	HR	C	C	C	C	A	D	D	Lo
Bulgaria	BG	D	C	C	C	D	C-D	D	Lo
Russia	RU	D	C	C	D	D	D	D	Mi
Lithuania	LT	D	D	C	D	A-D	C	D	Lo
Belarus	BY	D	D	C	D	D	D	D	Lo
Romania	RO	D	D	C	D	D	D	D	Lo
Ukraine	UA	D	D	C	D	D	D	D	Lo
Japan	JP	D	D	C-B	B	D	D	B	Hi
Turkey	TR	D	D	D	C	D	D	D	Mi
South Africa	ZA	D	D	D	D	A	D	B	Lo
China	CN	D	D	D	D	D	D	B-D	Hi
Argentina	AR	D	D	D	D	D	D	D	Lo
Chile	CL	D	D	D	D	D	D	D	Lo
Egypt	EG	D	D	D	D	D	D	D	Lo
Iran	IR	D	D	D	D	D	D	D	Lo
Mexico	MX	D	D	D	D	D	D	D	Lo
Malaysia	MY	D	D	D	D	D	D	D	Lo

Contd...

**Table 1: Contd...**

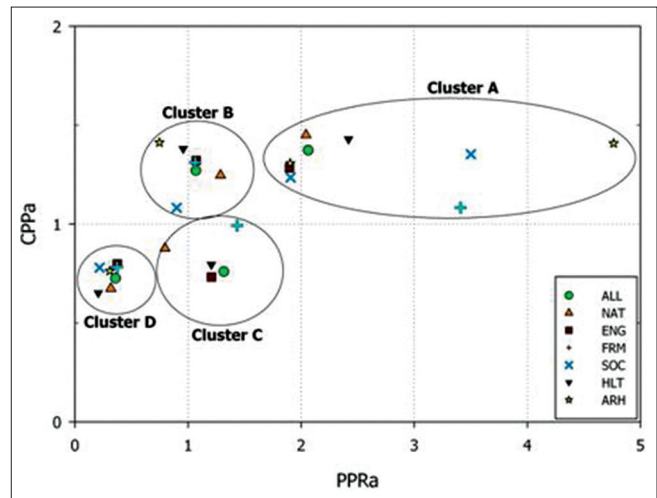
Name	Code	PPRa-CPPa map							Pa-Ca map
		General science	Science disciplines						
			Natural sciences	Engineering sciences	Health Sciences	Social sciences	Formal sciences	Arts and humanities	
Pakistan	PK	D	D	D	D	D	D	D	Lo
Brazil	BR	D	D	D	D	D	D	D	Mi
India	IN	D	D	D	D	D	D	D	Mi

The last column shows the relative position of countries in the Pa-Ca map (which is not scaled by the resources). Some countries are “transitional”, and thence, are marked by two (or more) cluster labels. PPR=Publication per resource, CPP=Citation per publication



**Figure 2:** PPRa-CPPa map (which is scaled by the resources) for the 50 topmost countries. Each point represents the average value of PPRa and CPPa of a specific country throughout the studied period (1996–2007). Unlike the Pa-Ca map (not scaled by the resources; Figure 1), there is no huge gap between the countries, and relative positions of some countries have changed drastically

Figure 2 shows the relative positions of countries in the resource-scaled map. Compared to the unscaled map [Figure 1], one notices that the huge gap between countries has disappeared; in addition, positions of some countries have been changed considerably. One of the countries that undergoes a drastic change in position upon resource-scaling is USA. In the unscaled map, it has a superior position relative to all others, especially its European competitors like UK and Germany. In the resource-scaled map, however, it appears as a group-B member with a considerable distance from topmost countries like Israel, Sweden or New Zealand. This is a remarkable instance of the consequences of accounting for the resource effects, in which countries with lower resources ‘outperform’ a country like USA, which has a considerable share of the global scientific production. Another significant instance is the case of China. It has a superior position (especially, as regards publication number) in the unscaled map, while after resource-scaling, it has been



**Figure 3:** Positions of centroids of four clusters (A to D) for “general science” (ALL) and its 6 major fields in the PPRa-CPPa map [Figure 2]

demoted to a very low position within group D, with a large distance from the average position in the world of science.

### CONCLUSION

In summary, it can be observed that accounting for the effect of resources considerably alters the relative positions of countries as the largest research units, and reveals some noticeable features, namely, co-evolving clusters of countries, which are observable in the general science as well as its major fields.

Some caveats are in order about interpreting the results. Firstly, GDP and population are merely rough indicators of the (academic) resources of a country when reliable detailed data about field-specific science spending and the academic skilled labor is available only for a few countries. Secondly, science is generally a cooperative task that is performed through extensive collaborations between research units like countries. Hence, the quality and quantity of the scientific output of a country is not only an outcome of intra-national efforts, but usually, a result of inter-national

fruitful collaborations. This might explain the top positions of countries like Switzerland and Israel which maintain a large number of efficient international collaborations.

The next step for this study would be considering different temporal periods to see how the clustering patterns vary by time. In addition, it is very interesting to investigate the details of the time-dependent network of international scientific collaborations to find out whether some co-evolving groups could be also recognized there.

### ACKNOWLEDGMENT

The authors would like to express their thanks to A. Bavali, R. Ghasemi, F. Jahangiri, and Z. Rezaee for fruitful discussions in the preliminary stages of the study that led to this article.

### REFERENCES

1. May RM. The scientific wealth of nations. *Science* 1997;275:793-6.
2. King DA. The scientific impact of nations. *Nature* 2004;430:311-6.
3. Division of Science Resources Statistics (SRS), NSF. Science and Engineering Indicators (SEI); 2008. Available from: <http://www.nsf.gov/statistics/seind08>. [Last accessed on 2015 Feb 14].
4. Science and Technology Indicator Project Team, NISTEP, MEXT. Science and Technology Indicators: 2004, A Systematic Analysis of Science and Technology Activities in Japan; 2004. Available from: <http://www.nistep.go.jp/achiev/ftx/eng/rep073e/id×073e.html>. [Last accessed on 2015 Feb 14].
5. Organisation for Economic Co-operation and Development (OECD). OECD Science, Technology and Industry Outlook. Paris: OECD, 2008.
6. European Commission, Directorate-General for Research, Science, Technology and Competitiveness, Key Figures Report 2008/2009.
7. Vinkler P. Composite scientometric indicators for evaluating publications of research institutes. *Scientometrics* 2006;68:629-42.
8. Schubert A, Braun T. Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics* 1986;9:281-91.
9. Hirsch JE. An index to quantify an individual's scientific output. *Proc Natl Acad Sci* 2005;102:16569-72. <arXiv:physics/0508025v5>.
10. Nejati A, Hosseini Jenab SM. A two-dimensional approach to evaluate the scientific production of countries (Case Study: The Basic Sciences). *Scientometrics* 2010;84:357-64.
11. Rehn C, Kronman U, Wadskog D. Bibliometric Indicators Definitions and Usage at Karolinska Institute. ver. 1.0 (2007).
12. Further information. Available from: <http://www.scimagojr.com>. [Last accessed on 2015 Feb 14].
13. Further information. Available from: [http://www.help.scopus.com/Content/h\\_fldnames.htm](http://www.help.scopus.com/Content/h_fldnames.htm). [Last accessed on 2015 Feb 14].
14. Quoted in "Nation Master" website. Available from: <http://www.nationmaster.com/country-info/stats/People/Population>. [Last accessed on 2015 Feb 14].
15. McQueen J. Some methods for classification and analysis of multivariate observations. In: Cam L, Neyman J, editors. *Proceedings of the 5<sup>th</sup> Berkeley symposium on mathematical statistics and probability*. Statistics. Vol. 1. Berkeley, CA: University of California Press, 1967.
16. Gan G, Ma C, Wu J. *Data Clustering: Theory, Algorithms, and Applications*. Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics, 2007.

**How to cite this article:** Jenab SH, Nejati A. Evaluation of the scientific production of countries by a resource-scaled two-dimensional approach. *J Sci Res* 2014;3:95-103.

**Source of Support:** Nil, **Conflict of Interest:** None declared

## Appendix: The details of the method

### (A1) Notation

In this article, a variable  $X$  (e.g. publication indicator) depends on time duration  $y$  (in years), field or subfield  $f$ , and country  $c$ ; therefore  $X$  would be explicitly denoted as  $X(y, f, c)$ . Here,  $y$  ranges from 1996 to 2007,  $f$  from 1 to 6 for the major academic fields and 1 to 27 for subfields (based on the Scopus classification), and  $c$  denotes any of the 50 studied countries (indexed by numbers). For example in 2002, Germany has published 10866 papers in physics, so this number of publications ( $P$ ) can be denoted as:

$$P(2002, \text{physics}, \text{Germany}) = 10,866; \quad (1)$$

Those papers have received 245,096 citations ( $C$ ) up to year 2007:

$$C(2000 :: 2007, \text{physics}, \text{Germany}) = 245,096. \quad (2)$$

In addition,  $\langle X \rangle_i$  denotes an average of a variable  $X$  over a set  $\{i\}$ ; for instance,

$$\langle P \rangle_c = \frac{1}{50} \sum_c P(y, f, c) \quad (3)$$

is an average of publication  $P$  over the set of 50 countries  $\{c\}$ .

We will drop arguments  $y, f$ , or  $c$  when they are clear from the context.

### (A2) Effect of Resources

Countries have widely different financial and human resources which definitely have a strong influence on the quantity and quality of their scientific production, and, therefore, on their relative positions in the final quantity–quality map. Quite obviously, a country with large resources (like USA or China) is likely to publish more, and obtain more citations than a country with much smaller resources (like New Zealand). However, our main objective here is to analyze scientific performance of countries while taking into account the share of total resources they benefit. Therefore, to achieve such a resource-scaled picture of the relative positions of countries in science, we have to account for these resource effects in the quantity and quality indicators. In the following, we consider the cases of publications and citations separately:

### (A2-1) Publication

To account for the effect of resources on the published output, firstly, one has to combine the human and financial resource measures into a single composite resource indicator; a straightforward way is to render them dimensionless, and subsequently, sum them up into a (field-independent) resource indicator,  $R$ :

$$R(y, c) = \frac{Pop(y, c)}{\langle Pop(y, c) \rangle_c} + \frac{GDP(y, c)}{\langle GDP(y, c) \rangle_c} \quad (4)$$

where  $Pop$  stands for population and  $GDP$  is the gross domestic product. Here, population and  $GDP$  have been used, respectively, as measures for the human and financial resources of countries. Division by the average taken over all the studied countries in the same year is used in converting them to dimensionless ratios. By dividing the publication number (as a quantity indicator) by the resource indicator, a new quantity measure, ‘publication per resource’ ( $PPR$ ), for a field  $f$  and a country  $c$  in year  $y$  will be obtained:

$$PPR(y, f, c) = \frac{P(y, f, c)}{R(y, c)} \quad (5)$$

### (A2-2) Citation

A higher total number of citations do not guaranty a higher quality because those citations might also indicate a high number of low-quality productions. Hence, to obtain an honest measure of academic quality, one has to account for the total of number of publications. So, we have used the well-known indicator, citation per publication ( $CPP$ ).<sup>[11]</sup> In this study, the extent of the ‘citation window’ is from the year of publication to the final year of the studied period (2007).  $CPP$  of a country in a specific year is the total citations (obtained within the window) to the publications of the country in that year, divided by the total number of those publications:

$$CPP(y, f, c) = \frac{C(y :: 2007, f, c)}{P(y, f, c)} \quad (6)$$

### (A3) Effect of Background Global Trends on Publications and Citations

When investigating the evolution of positions in a period, the global trends in quality and quantity of scientific production will affect the positions of countries in the maps, and only by their elimination, the positions in different years will be comparable to each other.

**(A3-1) Temporal Growth in Publication**

An annual increase in number of publications and in publication per resource (PPR) is observed throughout the world. We rescale by the average to find the position of a country compared to the annual global average.

**(A3-2) Intrinsic Trend of Citation**

It has been empirically observed that one should wait some years after the publication of an article for the citations to grow considerably. Regarding this fact, a diminishing trend will be seen for the citations of papers and consequently, in citation per publication (CPP) near the end of the study period (2007). This intrinsic decreasing trend would conceal and misrepresent the actual temporal evolution of quality of the scientific production of a research unit.

To eliminate these two background trends (A3-1 and A3-2), a simple scaling method has been applied in which the value of every indicator in a specific year is divided by the average value of that indicator in the same year. The average is taken over all the 50 countries in that year:

$$PPR_m(y, f, c) = \frac{PPR(y, f, c)}{\langle PPR(y, f, c) \rangle_c} \tag{7}$$

$$CPP_m(y, f, c) = \frac{CPP(y, f, c)}{\langle CPP(y, f, c) \rangle_c} \tag{8}$$

This scaling-by-average routine helps to remove the effects of the background global trends and provides dimensionless ratios that can be used in comparisons across different years and disciplines.

**(A4) Discipline-Specific Characteristics of Different Branches**

When combining or comparing indicators corresponding to different fields (or subfields) of science, one should note that each field of science has its own distinctive trend of publication or citation. For instance, mathematical subfields have a lower publication or citation than medical ones. Therefore, to make a fair comparison of the quality of a part (or whole) of scientific productions of countries, one should doubtlessly take account of this fact. Hence, to arrive at indicators that can be compared across different fields, we calculate the scaled indicator  $X_m$  ( $X$  represents any indicator like  $PPR$  or  $CPP$ ) for every subfield in each year. Then, to obtain an indicator for a major field composed of some subfields, we average

over the  $X_m$ 's corresponding to the subfields that make up that field:

$$X_a(y, F, c) = \frac{X_m(y, f_1, c) + X_m(y, f_2, c) + \dots}{\text{Total number of subfields included in } F} \tag{9}$$

where  $F$  can be the 'general science' or any of the six major academic fields (see section A6). Therefore, e.g. when considering the whole body of science, all the 27 subfields are averaged over; more explicitly:

$$PPR_a(y, \text{general science}, c) = \frac{PPR_m(y, f_1, c) + PPR_m(y, f_2, c) + \dots}{\text{Total number of subfields included in "general science"}} \tag{10}$$

$$\equiv \langle PPR_m(y, f, c) \rangle_f \text{ in "general science"} \tag{11}$$

$$CPP_a(y, \text{general science}, c) = \langle CPP_m(y, f, c) \rangle_f \text{ in "general science"} \tag{12}$$

**(A5) Clustering Method**

The aforementioned procedure (A2–A4) provides a time-series of  $PPR_a$ – $CPP_a$  data for countries throughout the studied period. Here, we will have 600 points on the 2-dimensional map constructed from  $PPR_a$ – $CPP_a$  data of 50 countries during 12 years (1996–2007) for each subfield or field of science. However, this map is still too complicated, mainly due to the temporal evolution of countries.

To obtain an intelligible picture out of this map, we use a clustering analysis to find out groups of co-evolving countries. Here, we apply the well-known *k-means* algorithm that is an unsupervised partitional clustering algorithm widely used in a variety of applications, and is relatively easy to understand and implement.<sup>[15]</sup> In addition, we use two 'internal' validity indices (*Dunn's* and *SD* criteria) to find out the optimum number of clusters. These criteria measure the compactness and separation of clusters; in addition, they are 'internal' since they evaluate the clustering results using only quantities and features acquired merely from the data set itself.<sup>[16]</sup> The optimum number of clusters would be the one corresponding to the lowest indices.

The resulting clusters are labeled by letters A to D. Notice that the cluster labels in all fields denote approximately the same position in the map as those for the general science; for



example, in social sciences (*Soc*) there is no cluster C, because in that case no cluster is located near cluster C of general science, but instead, two clusters of *Soc* are located near cluster A of general science [Figure 3] and hence, are labelled as A.

### (A6) Categorization of Subfields

We have categorized the initial 27 *Scopus* subfields of science into 6 major subfields: (1) Natural Sciences (*Nat*): Physics and Astronomy, Neuroscience, Immunology and Microbiology, Earth and Planetary Sciences, Chemistry, Biochemistry, Genetics and Molecular Biology, Agricultural and Biological Sciences; (2) Formal Sciences (*Frm*): Mathematics, Computer Science; (3) Engineering Sciences (*Eng*): Materials Science, Environmental Science, Engineering, Energy, Chemical Engineering; (4) Health Sciences (*Hlt*): Veterinary, Pharmacology, Toxicology and Pharmaceutics, Nursing, Medicine, Health Professions, Dentistry; (5) Social Sciences (*Soc*): Decision Sciences, Social Sciences, Psychology, Economics, Econometrics and Finance, Business, Management and Accounting; (6) Arts

and Humanities (*Arh*): Arts and Humanities. Indeed, this categorization is arbitrary.

### (A7) Unscaled Map

Two indicators are used to construct the unscaled map, namely, *Pa* as the quantity and *Ca* as the quality indicator. The procedure to obtain these two indicators is analogous to *PPRa* and *CPPa*, except that we do not scale the number of publications *P* by the resources, and the number of citations *C* by total number of papers. More explicitly:

$$Pm(y, f, c) = \frac{P(y, f, c)}{\langle P(y, f, c) \rangle_c} \quad (13)$$

$$Cm(y, f, c) = \frac{C(y, f, c)}{\langle C(y, f, c) \rangle_c} \quad (14)$$

$$Pa(y, f, c) = \langle Pm(y, f, c) \rangle_f \text{ in } F \quad (15)$$

$$Ca(y, f, c) = \langle Cm(y, f, c) \rangle_f \text{ in } F \quad (16)$$

### “Quick Response Code” link for full text articles

The journal issue has a unique new feature for reaching to the journal’s website without typing a single letter. Each article on its first page has a “Quick Response Code”. Using any mobile or other hand-held device with camera and GPRS/other internet source, one can reach to the full text of that particular article on the journal’s website. Start a QR-code reading software (see list of free applications from <http://tinyurl.com/yzlh2tc>) and point the camera to the QR-code printed in the journal. It will automatically take you to the HTML full text of that article. One can also use a desktop or laptop with web camera for similar functionality. See <http://tinyurl.com/2bw7fn3> or <http://tinyurl.com/3ysr3me> for the free applications.