Research and Development: Bibliometric Analysis of Knowledge Flows of Brazilian Research 2005-2009

Branco Ponomariov and Hannes Toivanen

***Abstract*— By examining the knowledge in- and outflows in Brazilian research 2005-2009, we undertake comparative bibliometric analysis of the dynamics of knowledge creation in development context. Specifically, we analyse knowledge creation dynamics to find out how knowledge flows between developing countries and the "North" shape research and its exploitation.**

***Index Terms*— Research evaluation, Brazil, bibliometrics, knowledge transfer**

I. Introduction

Emphasis on the role of science, technology and innovation (STI) has increased on the global development agenda over the last decade[1],[2],[3]. Capturing the essence of this movement, the United Nations Millennium Project highlighted the critical importance of building national capacities in science, technology and innovation for achieving sustainable progress in delivering the Millennium Development Goals in most challenged countries and regions [4].

A central objective for harnessing STI for development is to build sustainable capacities in research in developing countries and emerging economies, so that research communities may address locally identified development challenges and can enhance local business and industry sectors. Although developing regions are highlighting impressive growth in research, measured as scientific publications, there is scant research to estimate how the increased research output is reflected in local capacities or focus of research.

The issue of dependence is recurrent in development, and of significance in the context of innovation and development too. In fact, analysts of African research have argued that it is increasingly dependent on European and North American research. African countries contribute relatively little to each other knowledge base, if measured as co-authored papers, and researchers have pointed out several other symptoms of uneven collaboration patterns between Western and African researchers [5],[6],[7],[8].

The sources of this growth are important to understand from both developmental and S&T policy perspective. Specifically, it is important to understand if this growth in publications extends and builds upon and expands domestic scientific capacity, or is it a derivation from externally produced knowledge. A related, and equally important question is whether domestic scientific capacity is sufficiently robust and whether it has influence extending beyond the local context. To shed light on these questions, this paper examines the origin of the knowledge base of Brazilian papers, approximated as the national origin of the papers cited in the Brazilian research literature, and it also explores the utilization of Brazilian knowledge by focusing on the national origin of citing papers.

The case of Brazil, on the other, has been cited as the counter example, as it has built science based industrial clusters in bioenergy, aerospace, forest industries, and petrochemicals, to name few. However, regional aspects of Brazilian STI are poorly analyzed, and capacities are heavily concentrated in few selected regions, such as São Paulo, Rio de Janeiro, Belo Horizonte, and few others. States and regions dealing with serious development challenges, as rural regions in general, or Western and North-Western Brazil rarely figure in assessments of Brazilian STI capacities.

This paper explores the role domestic research capabilities in building national innovation systems. We investigate knowledge in-flows and out-flows, as measured in bibliometric citations, and through analysis of regional aspects and characteristics of high-citing papers in Brazilian research output between 2005 and 2009. The objective of the papers it is to identify different

Manuscript received September 29, 2011. This paper has been prepared with financial support from Tekes – the Finnish Funding Agency for Technology and Innovation for the project "Smart Globalization: Evolving relationships between emerging economies and high-tech regions."

B. Ponomariov is with the University of Texas, San Antonio 501 W. Durango Blvd. San Antonio, TX 78207, USA Tel. +1 210 458 2619, Fax +1 210 458 2536 Email: branco.ponomariov@utsa.edu

Hannes Toivanen is with the VTT Technical Research Centre of Finland P.O.Box 1000, 02044 VTT, Finland Tel. +358 40 186 3882, Fax +358 20 722 7604 Email: hannes.toivanen@vtt.fi

contributing roles of domestic and foreign entities to Brazilian research system, and on the other hand to contribute to the development of methodologies and framework to analyze the building of innovation systems in emerging and developing countries.

II. Data and Methods

The analysis presented in this paper utilizes three data sets. The core data set (SOURCE dataset) is composed of a complete set of papers with at least one author with Brazilian address indexed by the Institute of Scientific Information (ISI) between 2005 and 2009 (inclusive). The total set of Brazilian papers for this period consists of 152,031 papers. The data was retrieved in XML format at the individual article level directly from the Thomson Reuters. Each article record contained the full text of all fields indexed by ISI (including citing and cited papers. Raw XML records provide the most comprehensive individual article bibliometric data available, and allow for analysis and aggregation at multiple levels. For this paper, we limited the SOURCE data set to articles and conference proceedings, consisting of 127,826 papers.

In addition, we obtained two other dataset covering all ISI-indexed articles that had been cited by paper in SOURCE dataset(CITED dataset), as well as dataset covering all papers that had made citation to SOURCE dataset paper (CITING dataset). Use and processing of these data sets with has been described in detail below.

We have not assessed how well the citing and cited data cover the references made in SOURCE papers, but they do cover all recent papers published in ISI-indexed publications, and thus do provide complete measure of citations within the ISI-universe of articles. Cut-of-date for the data delivery was May 2010. While most aspects of the paper remain the same, the number of citations received does change and in this regard the cut-off-date is significant.

Data in all three data sets was similar, though older article level data tends to have gaps. SOURCE, CITING, and CITED data were linked through unique ISI identifier for references, thus allowing us to analyze different aspects of citation behavior and patterns, as well as to add characteristics to the SOURCE papers.

An important part of the data processing was that we have assigned the state-address for author information for Brazilian addressees in SOURCE data. To this date, we have succeeded in adding state to 25,000 Brazilian records from 2005-2009, enhancing the coverage of this field up to 98% in the data. We believe this information yields a more detailed perspective in Brazilian research system, as described below.

The data was initially processed by means of a text-mining application – VantagePointTM software that, based on the XML field tags and regular expressions, parses and extracts the text data into fields, and thus makes the textual information amenable to export and analysis in a spread sheet. Spread sheet data, in turn, allows for easy generation of co-authorship matrices, which then can be processed in appropriate software. The analyses of research collaboration networks was performed in a social networks analysis software – UCINET TM [9].

III. Brazilian research

Currently, Brazil is cited as a benchmark for how to harness science and technology for development. Long-run trend in increasing country's knowledge production, the number of Brazilian science articles has increased over 18-fold during the last 3 decades, is often credited in enabling country's transition to knowledge economy and its emergence as one of the global economic and innovation hubs [10],[11]. This growth and expansion has been accompanied by transformation of the Brazilian research and innovation system [12],[13]. One aspect of the growth has been tremendous expansion of funding, as well as enhanced productivity of the Brazilian research system [10]. Research literature and interest on Brazilian innovation system is also growing rapidly, and ever more articles examine the dynamics and patterns behind its growth [14],[15].

TABLE 1.
BRAZILIAN ARTICLES AND PROCEEDCINGS 2005-2009.

Year	Articles and proceedings	Articles	Annual growth articles and proceedings	Annual growth articles
2009	31674	28175	2,33 %	7,09 %
2008	30952	26309	19,81 %	23,23 %
2007	25835	21349	24,53 %	28,17 %
2006	20746	16657	11,42 %	14,48 %
2005	18619	14550		
TOTAL	**127826**	**107040**		

A central issue for this paper is to point out how the Brazilian research system is undergoing a transformation as a part its expansion process, and especially to look at aspects that can indicate to what degree Brazilian capabilities are contributing to growth and quality, and to what degree foreign capabilities are playing a role.

The number of Brazilian articles and proceedings between 2005 and 2009 increased about 70%, whereas the number of articles only increased almost 94%. Annual growth figures in our data vary greatly, and according to Thomson Reuters is due to their aggressive expansion of coverage in South American journals that was completed before 2009.[1] Yet, what stands, is that Brazilian research is growing at a strong rate and appears well above the global averages. (Table 1.)

Although there is growing number of research on the growth of Brazilian research system and output, relatively little research is available on state-level differences and dynamics. To examine in detail how the nature of Brazilian research system is evolving, we have looked at the growth of research at state levels. As is known,Sao Paulo towers Brazilian research with almost 40% share of all papers. After it, Rio de Janeiro, Minas Gerais and Rio Grande do Sul have shares around 10%, and top four states combined make up 70% of all research in Brazil 2005-2009. (Table 2.)

Growth rates for individual states vary greatly. Sao Paulo and Rio de Janeiro have slightly under 60% growth rate in 2005-2009, whereas Minas Gerais and Rio Grande do Sul have about doubled their research output, suggesting they are quickly catching up with Rio de Janeiro. Several other follow-up states have high growth rates, suggesting that the Brazilian system is getting more evenly balanced and less centered on Sao Paulo and Rio de Janeiro (Table 2.)

TABLE 2.
GROWTH AND DISTRIBUTION OF PAPERS BY BRAZILIAN STATES 2005-2009

	PAPERS	STATE	Growth 05-09	Share of BR papers	Cumulative Share
1	59435	SP	57,94 %	38,39 %	38,39 %
2	20944	RJ	56,36 %	13,53 %	51,92 %
3	14869	MG	107,95 %	9,61 %	61,53 %
4	13977	RS	95,89 %	9,03 %	70,56 %
5	8987	PR	107,29 %	5,81 %	76,36 %
6	5755	DF	68,24 %	3,72 %	80,08 %
7	5223	SC	88,54 %	3,37 %	83,45 %
8	4436	PE	106,24 %	2,87 %	86,32 %
9	3501	BA	87,55 %	2,26 %	88,58 %
10	3245	CE	93,81 %	2,10 %	90,68 %
11	2345	PB	146,60 %	1,51 %	92,19 %
12	1886	GO	118,36 %	1,22 %	93,41 %
13	1842	RN	96,09 %	1,19 %	94,60 %
14	1764	PA	94,92 %	1,14 %	95,74 %
15	1282	AM	124,85 %	0,83 %	96,57 %
16	1099	ES	179,84 %	0,71 %	97,28 %
17	1009	MS	256,99 %	0,65 %	97,93 %
18	646	MT	296,83 %	0,42 %	98,35 %
19	525	SE	228,26 %	0,34 %	98,69 %
20	509	AL	211,11 %	0,33 %	99,02 %
21	502	MA	207,69 %	0,32 %	99,34 %
22	353	PI	243,33 %	0,23 %	99,57 %
23	201	TO	428,57 %	0,13 %	99,70 %
24	178	RO	55,17 %	0,11 %	99,81 %
25	139	AC	366,67 %	0,09 %	99,90 %
26	99	RR	163,64 %	0,06 %	99,97 %
27	51	AP	183,33 %	0,03 %	100,00 %

N=154802. Note: Papers with multiple Brazilian addresses are counted multiple times. Dataset N=127823.

[1] Personal communication from Thomson Reuters.

IV. KNOWLEDGE IN-FLOWS - CITATIONS MADE

Citations are made in research papers typically to frame research in relevant theoretical and methodological framework, to discuss relevant research results published in other papers. References present also relatively stable set of data, as authors include in papers likely a comprehensive set of literature according to their best knowledge and understanding at the point of pubication. Thus, unlike citations received by individual paper, references do not change after publication [16].

When knowledge in- and outflows are measured exclusive through citations made and received, they address intellectual aspects of research production. By relying on certain group of papers to frame, substantiate, develop or critique, a citation is acknowledgment of intellectual debt. Yet, if we add geogprahic or organizational information, we can also address social ties [17].

Analysis of the geographic origin for research literature cited in research literature can be used as a proxy to identify knowledge flows and social structures. As a relatively new approach it cannot be relied upon alone, but it does provide additional perspective to evaluate and analyse research landscape. While we are not able to consider in our analysis the nature or intention of individual citations made, a comprehensive data set is likely to cast light on the structure and nature of national research effort, as well as to showcase the relative standing of other nations and regions.

We compare two sets of cited literature, for years 2005 and 2009, for Brazilian research literature to identify on what countries and regions Brazilian researchers are attuned to rely upon, and to identify possible changes. The analysis extracts for all Brazilian articles and conference proceedings cited literature that has been published within 10 years. Thus, our data includes for Brazilian papers from 2005 cited literature from between 1995 and 2005 (inclusive), and respective data for Brazilian papers published in 2009. This limitation is done to remove "noise" caused by possible ritualistic citations, as well as to focus analysis on knowledge flows between existing capabilities in different countries.

Citations made concentrate heavily on major global centers of scientific and technological research, as illustrated in Table 3. and Fig. 1. Europe is by far the most important source for cited research literature in Brazilian research, originating over 40 % of all citations made. North America captures about one fifth, whereas Brazil captures about 15%. Asia accounts for roughly 10 %, and after that regional shares decline significantly. The data also confirms how outward oriented Brazilian research is, as the share of South American literature remains relative small.

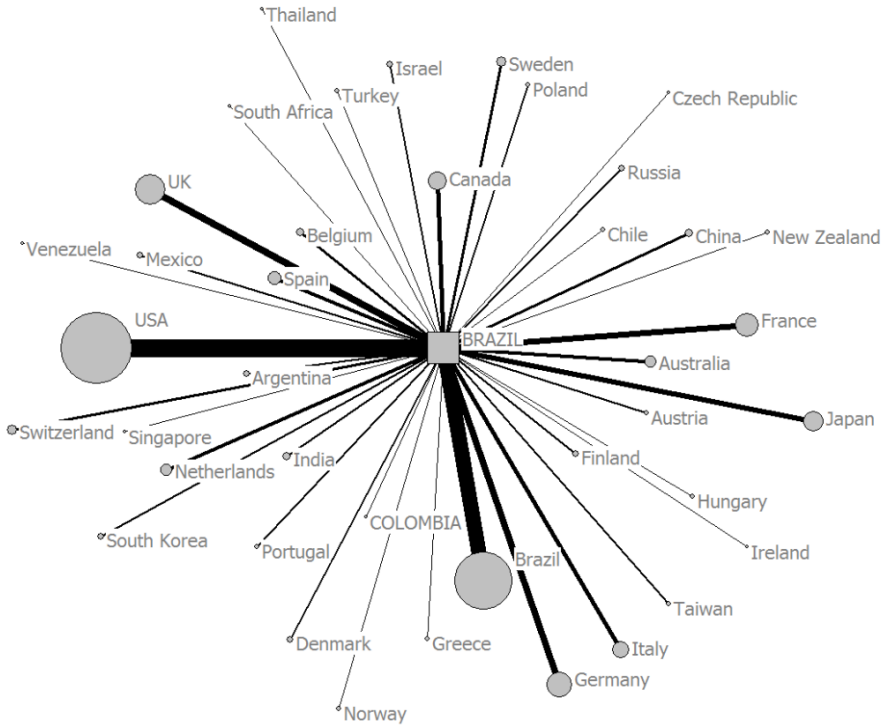

Fig. 1. Origin of cited research for Brazilian papers from 2005. Only countries with 500 or more papers cited.*

N for Brazilian articles and proceedings in 2005=18619; N for cited papers=238418; *Cited papers limited to publications from 1995-2005. **Note:** Scale for strength of ties is 1-20 and based on occurrence of citations. Scale for node size is 2-20 and based on all papers from the country, except for the square Brazil, whose size is based on publications in 2005.

Change in reference volumes indicates the relevance of a country's research effort in relation to that of Brazil. Although the ranking of individual regions does not change between 2005 and 2009, despite the 170% increase in Brazilian papers, Table 3**.** does demonstrate interesting changes.

TABLE 3.
CITATIONS MADE BY BRAZILIAN PAPERS I 2005 AND 2009 BY RESEARCH MAJOR REGION.

Region	Papers cited by BR papers 2005	Papers cited by BR papers 2009	Share of 2005 citations	Share of 2009 citations	%-change of citation share	Absolute Citation Change 05-09
Europe	101871	193917	42,73 %	40,94 %	-4,18 %	90,36 %
North America	50375	99987	21,13 %	21,11 %	-0,09 %	98,49 %
Brazil	35207	71659	14,77 %	15,13 %	2,45 %	103,54 %
Asia	25166	53405	10,56 %	11,28 %	6,82 %	112,21 %
Oceania	7191	15336	3,02 %	3,24 %	7,35 %	113,27 %
South America	6846	14440	2,87 %	3,05 %	6,17 %	110,93 %
Middle East	5064	11120	2,12 %	2,35 %	10,54 %	119,59 %
Central America and Carribean	3463	6931	1,45 %	1,46 %	0,75 %	100,14 %
Africa	3235	6847	1,36 %	1,45 %	6,54 %	111,65 %
TOTALS	**238418**	**473642**	**100,00 %**	**100,00 %**		

Most importantly, the share of citations made to European research literature drops over 4%, whereas North America maintains its share almost unchanged. Thus, the results suggest that the relevance of European research is declining relatively– at fastest rate of any major global region – for Brazilian research, whereas the U.S. is able to maintain its relevance. This is supports research that has demonstated that U.S. research literature and patents are receiving citations at higher rate than European, indicating that the quality – or relevance - of European research and technology trails that of American [18, 19]. At country level, European declines show big variation: the share of UK has declined only 1%, whereas Germany almost 8%, France almost 13 and Italy 6. Similary, in North America the U.S. has actually increased its share whereas Canada's has declined over 6% (Table 4.).

When we look at regions increasing their share, it stands out that Brazil appears also to rely increasingly on domestic research. It doubles the number of cited papers and increases almost 2.5% its share of citations made. Other regions gaining are all emerging and developing regions. At individual country level, China increases its share of citations almost 35% and India over 18%, showing largest gains in the data among top sources of citations. In absolute terms, China and India are also overtaking some major research countries, such as Switzerland and Sweden. (Table 3. and Table 4.)

In conclusion, we can boil down the results in Four key observations: 1) Brazilian research is relying increasingly on country's own knowledge base, suggesting that domestic capabilities are supportive of dynamics of research growth; 2) The relevance of European research for Brazilian research is declining, although there appears to be huge variation at national level; 3) The relevance of U.S. research remains stable for Brazilian research; 4) Emerging and developing countries, such as Brazil, China, India, Turkey and South Africa appear to have growing dependencies in knowledge production, here visible as strong growth as sources of cited literature for Brazilian research. Yet, these linkages are still of modest scale when compared to key relationships to the U.S. and Europe.

TABLE 4.
CITATIONS MADE BY BRAZILIAN PAPERS 2005 AND 2009 BY RESEARCH COUNTRY – TOP-25 IN 2009

Rank 2009	Country	Papers cited by BR papers 2005	Papers cited by BR papers 2009	Share 2005 citations	Share 2009 citations	%-change of citation share 05-09	Absolute Citation Change 05-09
1	**TOTALS**	**238418**	**473642**				98,66 %
2	USA	41021	82523	17,21 %	17,42 %	1,26 %	101,17 %
3	Brazil	35207	71659	14,77 %	15,13 %	2,45 %	103,54 %
4	UK	17012	33464	7,14 %	7,07 %	-0,98 %	96,71 %
5	Germany	13993	25634	5,87 %	5,41 %	-7,79 %	83,19 %
6	France	13216	22923	5,54 %	4,84 %	-12,69 %	73,45 %
7	Japan	10362	18366	4,35 %	3,88 %	-10,78 %	77,24 %
8	Canada	9354	17464	3,92 %	3,69 %	-6,02 %	86,70 %
9	Italy	9194	17180	3,86 %	3,63 %	-5,94 %	86,86 %
10	Spain	7281	15341	3,05 %	3,24 %	6,06 %	110,70 %
11	Australia	5967	12673	2,50 %	2,68 %	6,91 %	112,38 %
12	Netherlands	5876	11591	2,46 %	2,45 %	-0,70 %	97,26 %
13	China	4075	10926	1,71 %	2,31 %	34,97 %	168,12 %
14	Switzerland	4849	9260	2,03 %	1,96 %	-3,87 %	90,97 %
15	Sweden	4538	8443	1,90 %	1,78 %	-6,35 %	86,05 %
16	India	3485	8202	1,46 %	1,73 %	18,47 %	135,35 %
17	Belgium	3664	7031	1,54 %	1,48 %	-3,41 %	91,89 %
18	Argentina	2928	6096	1,23 %	1,29 %	4,80 %	108,20 %
19	South Korea	2622	5601	1,10 %	1,18 %	7,53 %	113,62 %
20	Denmark	2566	5214	1,08 %	1,10 %	2,28 %	103,20 %
21	Israel	2716	4697	1,14 %	0,99 %	-12,95 %	72,94 %
22	Mexico	2141	4199	0,90 %	0,89 %	-1,28 %	96,12 %
23	Portugal	1903	4190	0,80 %	0,88 %	10,83 %	120,18 %
24	Finland	2106	3994	0,88 %	0,84 %	-4,54 %	89,65 %
25	Taiwan	1917	3917	0,80 %	0,83 %	2,85 %	104,33 %

V. KNOWLEDGE OUTFLOWS - CITATIONS RECEIVED

Here, we look at citations received by Brazilian papes from two perspectives: *First*, we look at the the geographical origin of papers citing Brazilian research in order to look more specifically at in what locations Brazilian research is demonstrating relevance. *Secondly*, we analyze the excellence in Brazilian research by looking in detail at the top-quartile of papers receiving citations.

When processing data for citations received, we have limited the data set making citations to articles and proceedings and excluded other type of papers, such as reviews, comments, editorials, etc. etc. This ensures that we focus essentially on knowledge flows, that is research outputs that make citations in the context of research output. Our citation data is not limited in years, but we include all papers making citations to Brazilian papers, yet the cut of date for our data was in early 2010, leaving little time for 2009 published papers to received citations.

A. Citations received by major regions

Utilization of research through citations can be taken as a proxy for relevance of the cited research. When looked at country or regional level, citations received by Brazil suggests what regions and to what degree they rely on Brazilian knowledge production. Table 5. details by source region and Brazil citations received by Brazilian papers.

TABLE 5.
CITATIONS RECEIVED BY BRAZILIAN PAPERS 2005 AND 2009 BY MAJOR REGIONS CITING

Country	Records 2005	Records 2009	Citations received 2005	Citations received 2009	Share of citations received 2005	Share of citations received 2009	% Change in share of received citations 05-09
Europe	133963	60469	91102	4834	31,96 %	27,45 %	-14,10 %
Brazil	111437	70141	74978	7338	26,30 %	41,67 %	**58,43 %**
North America	77331	35512	52949	2416	18,57 %	13,72 %	-26,14 %
Asia	56080	25699	38480	1535	13,50 %	8,72 %	-35,42 %
South America	13502	7064	8943	588	3,14 %	3,34 %	**6,44 %**
Middle East	10596	4782	6890	336	2,42 %	1,91 %	-21,06 %
Oceania	8125	3483	5019	222	1,76 %	1,26 %	-28,40 %
Africa	6028	2572	3195	144	1,12 %	0,82 %	-27,04 %
Central America and Carribean	5836	2794	3531	198	1,24 %	1,12 %	-9,23 %
TOTALS	**422898**	**212516**	**285087**	**17611**	**100,0 %**	**100,0 %**	

Fig. 2. presents what countries are drawing upon Brazilian research by citing Brazilian papers from 2009, demonstrating that domestic capacities are important for the current mode of growth in Brazil. It also demonstrates that Brazilian research is relatively quikcly distributed globally and that some important differences can be detected within less than years through analysis of citations received.

Change in regional or country level citation patterns can cast light on the dynamics behind the growth of Brazilian research. The share of Brazilian citations is highest for 2009, indicating self-citation and rapid domestic diffusion, but evens out when more time lapses to make citations: Indeed, Brazilian citations to 2009 papers capture 60% more of total citatations than they received for 2005 papers. South America has analogous trend, with higher citation rate for 2009 and lower for 2005, but this remains relatively insignificant as South America accounts in total around 3% of all citations received. European reserch appears to be relative quick to cite Brazilian research too, suggesting proximity in research interests or networks. Especially Spain, Switzerland, and Portugal cite quickly Brazilian research. (Tables 5. and 6.)

When citations received are broken down to individual countries, USA is easily the largest country with over 16% share in 2009. Second largest country in 2009 is China, with over 5% share, reinforcing impressions about its rise in global science and new South-South linkages. China, Germany, UK, and France make up a second tier group in making citations to Brazilian papers, followed by a third group of Italy, Spain, Japan and Canada. (Table 6).

Figure 2 provides a network map of countries citing Brazilian research from 2009, highlighting by tie strength the volume of citations. For future research, it would be useful to break down the citation flows by key research areas map their connections with research collaboration patterns.

Fig. 2. Origin of received citations for Brazilian papers from 2009. Only countries with over 40 citing papers.

N for Brazilian articles and proceedings in 2009=31497; N for citing articles=11423. **Note:** Scale for strength of ties is 1-20 and based on occurrence of citations. No scale for nodes.

Our analysis of geographical origin of citations received for Brazilian research in 2005 and 2009 show that the country continues to rely on its domestic capacities in research. While some of domestic citation behaviour may be driven by self-citation, there are also indication that suggest of strong domestic research agenda setting: The sheer volume of Brazilian research reduces the importance of self-citation. Secondly, there is high volume of international research groups making citations to Brazilian research, indicating its relevance to other countries.

TABLE 6.
CITATIONS RECEIVED BY BRAZILIAN PAPERS FROM 2005 AND 2009 BY TOP—35 CITING COUNTRIES (FOR 2005 PAPERS)

Rank	Country	Records 2005	Records 2009	Citations received 2005	Citations received 2009	Share of received citations 2005	Share of received citations 2009	% Change in share of received citations 05-09
1	Brazil	111437	70141	74978	7338	26,30 %	41,67 %	**58,43 %**
2	USA	66677	30820	45711	2041	16,03 %	11,59 %	-27,72 %
3	China	21950	9835	15203	605	5,33 %	3,44 %	-35,58 %
4	Germany	19642	9106	13408	629	4,70 %	3,57 %	-24,06 %
5	UK	18349	8422	12569	621	4,41 %	3,53 %	-20,02 %
6	France	16879	7768	11559	597	4,05 %	3,39 %	-16,39 %
7	Italy	13129	5724	8935	407	3,13 %	2,31 %	-26,26 %
8	Spain	12839	5623	8879	561	3,11 %	3,19 %	**2,28 %**
9	Japan	11539	5429	7945	259	2,79 %	1,47 %	-47,23 %
10	Canada	10654	4692	7238	375	2,54 %	2,13 %	-16,13 %
11	India	7691	3811	5288	242	1,85 %	1,37 %	-25,92 %
12	Australia	7047	3016	4674	201	1,64 %	1,14 %	-30,39 %
13	Netherlands	5977	2557	4026	219	1,41 %	1,24 %	-11,94 %
14	South Korea	5784	2655	4025	157	1,41 %	0,89 %	-36,86 %
15	Switzerland	5277	2249	3666	238	1,29 %	1,35 %	**5,09 %**
16	Argentina	5530	3001	3663	206	1,28 %	1,17 %	-8,96 %
17	Belgium	4241	1838	2929	156	1,03 %	0,89 %	-13,78 %
18	Sweden	4266	1985	2906	149	1,02 %	0,85 %	-17,00 %
19	Russia	4134	1969	2827	128	0,99 %	0,73 %	-26,70 %
20	Turkey	4234	1958	2824	84	0,99 %	0,48 %	-51,85 %
21	Poland	3830	1739	2614	133	0,92 %	0,76 %	-17,64 %
22	Portugal	3911	1864	2608	189	0,91 %	1,07 %	**17,31 %**
23	Mexico	3994	1814	2544	123	0,89 %	0,70 %	-21,73 %
24	Taiwan	3640	1610	2509	96	0,88 %	0,55 %	-38,06 %
25	Iran	2693	1191	1871	95	0,66 %	0,54 %	-17,81 %
26	Israel	2351	1094	1686	100	0,59 %	0,57 %	-3,99 %
27	Chile	2435	1164	1652	93	0,58 %	0,53 %	-8,87 %
28	Czech Republic	2477	1262	1649	83	0,58 %	0,47 %	-18,52 %
29	Austria	2359	989	1636	152	0,57 %	0,86 %	**50,40 %**
30	Greece	2222	935	1529	62	0,54 %	0,35 %	-34,36 %
31	Denmark	2097	930	1407	68	0,49 %	0,39 %	-21,76 %
32	Finland	2046	835	1361	62	0,48 %	0,35 %	-26,26 %
33	Colombia	2073	1034	1346	114	0,47 %	0,65 %	**37,10 %**
34	South Africa	1597	688	1061	36	0,37 %	0,20 %	-45,07 %
35	Hungary	1439	672	978	52	0,34 %	0,30 %	-13,93 %
	TOTALS	**422898**	**212516**	**285087**	**17611**	**94,60 %**	**94,66 %**	

B. Distribution of citations received – excellence in Brazilian research

Citations received are known to be highly skewed so that high-cited papers represent very small number of total papers but capture bulk of citations. If we look at the highest quartile of all papers in 2005 and 2009, we observe that from Brazilian papers published in 2005 they captured 75% of all citations, and in 2009 about 77%.The top-5 % of papers published in 2005, receiving

24 or more citations and consisting of 981 papers, captured 36,72% of all citations, and the top-1% of papers in 2005, consisting of 196 papers with 52 or more citations by 2009, capture 18,46 % of all citations. (**Fig 3. and Fig 4**.).

For papers from 2009, the Top-5 % of papers have received 3 or more citations and consists of 2327 papers and capture 46% of all citations. Top-1 percent has received 7 or more citations and consists of 441 papers, capturing about 21% of all citations received. (**Fig 4.**).

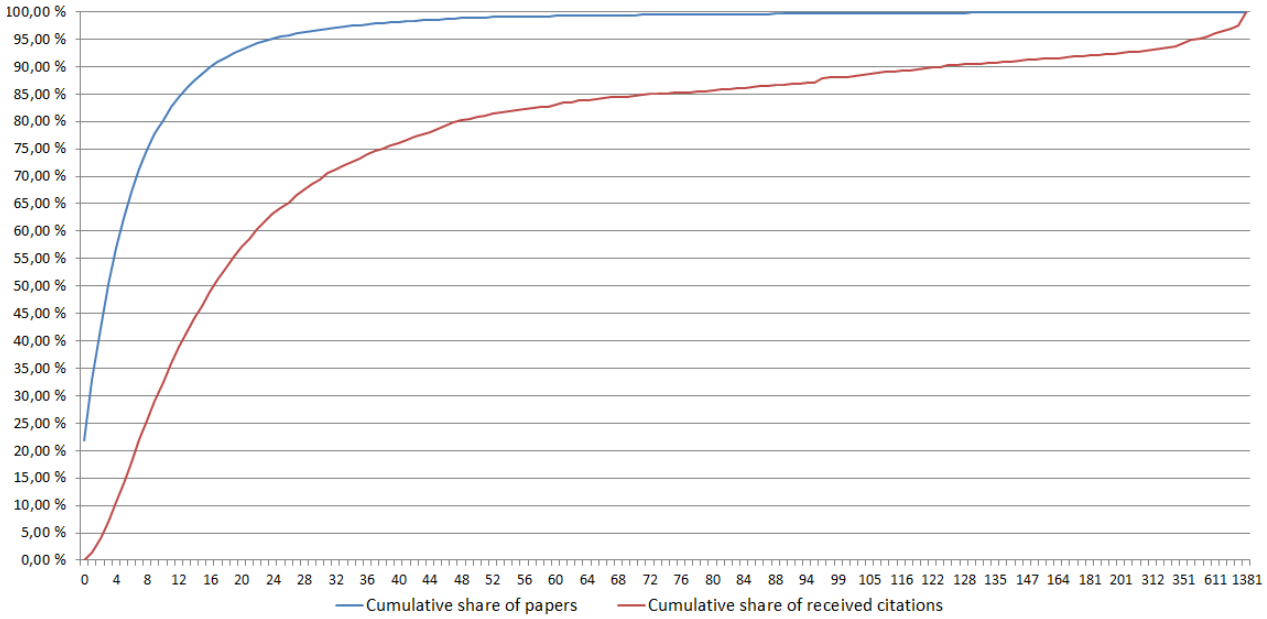

Fig. 3. Cumulative share of papers and received citations by number of citations for Brazilian papers published 2005.

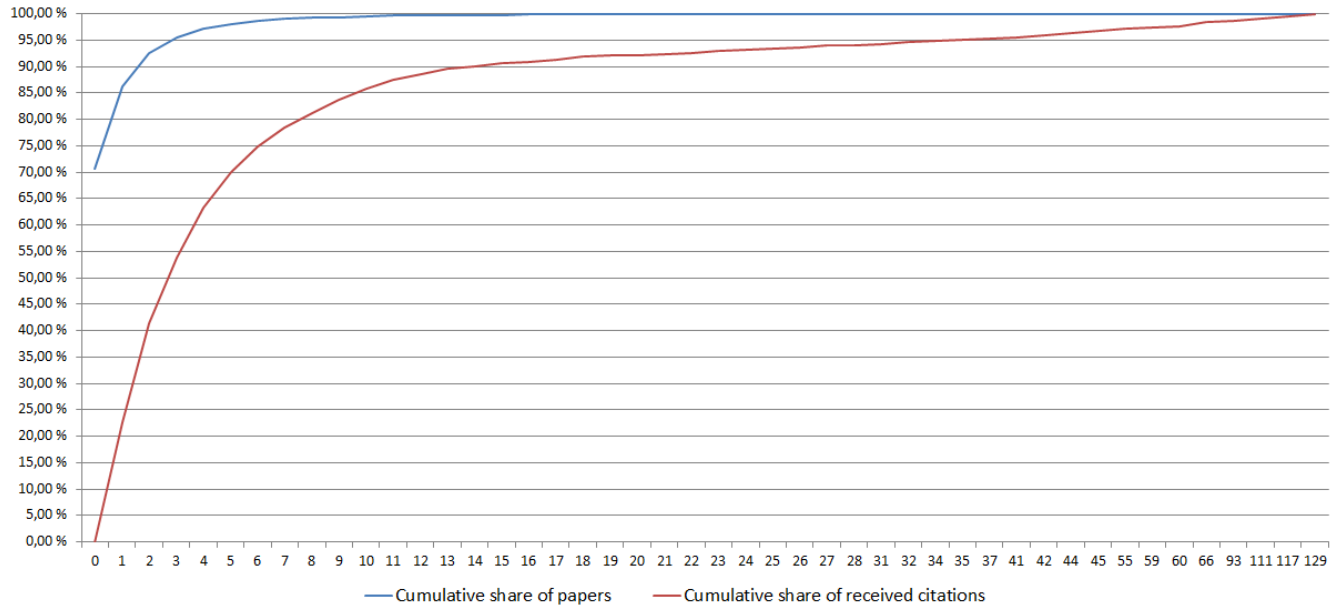

Fig. 4. Cumulative share of papers and received citations by number of citations for Brazilian papers published 2009.

Naturally, citations received are much more skewed with the 2005 data as there has been longer time for them to accumulate citations. Our data does not distuinguish between different fields, and thus includes a bulk of humanistic papers likely receive very little or no citations. However, distribution of cumulative citations suggest that there are distuingishable segments of excellence in Brazilian research.

When we look at the geographical distribution of high cited papers in Brazil by state-level information on research address, we are able to relate overall volume to quality of research, and also to state-level growth (Table 2)**.** High citing papers show some degree of concentration, but this is relatively modest. Of 27 states 22 are contributing to the top 5% pool of most cited papers, and 16 states belong to the top 1% pool.

TABLE 7.
DISTRIBUTION OF TOP-5 %, TOP-25% AND ALL BRAZILIAN 2005 RESEARCH PAPERS BY STATE.

	State	Top-5% records	Share		State	Top-25% records	Share		Share	All records	Share
1	SP	566	45,57 %	**1**	SP	2532	44,08 %	**1**	SP	8997	40,92 %
2	RJ	193	15,54 %	**2**	RJ	793	13,81 %	**2**	RJ	3231	14,69 %
3	RS	102	8,21 %	**3**	RS	531	9,24 %	**3**	MG	1911	8,69 %
4	MG	88	7,09 %	**4**	MG	445	7,75 %	**4**	RS	1847	8,40 %
5	DF	74	5,96 %	**5**	PR	271	4,72 %	**5**	PR	1166	5,30 %
6	PR	56	4,51 %	**6**	DF	262	4,56 %	**6**	DF	869	3,95 %
7	SC	41	3,30 %	**7**	SC	211	3,67 %	**7**	SC	707	3,22 %
8	BA	22	1,77 %	**8**	BA	125	2,18 %	**8**	PE	593	2,70 %
9	PE	15	1,21 %	**9**	CE	114	1,98 %	**9**	BA	474	2,16 %
10	CE	14	1,13 %	**10**	PE	105	1,83 %	**10**	CE	452	2,06 %
11	GO	14	1,13 %	**11**	PA	60	1,04 %	**11**	PB	294	1,34 %
12	PA	12	0,97 %	**12**	PB	54	0,94 %	**12**	GO	256	1,16 %
13	AM	10	0,81 %	**13**	GO	49	0,85 %	**13**	RN	256	1,16 %
14	PB	9	0,72 %	**14**	RN	45	0,78 %	**14**	PA	236	1,07 %
15	RN	9	0,72 %	**15**	AM	37	0,64 %	**15**	AM	169	0,77 %
16	ES	7	0,56 %	**16**	ES	28	0,49 %	**16**	ES	124	0,56 %
17	AL	3	0,24 %	**17**	MA	15	0,26 %	**17**	MS	93	0,42 %
18	MA	2	0,16 %	**18**	MS	14	0,24 %	**18**	MT	63	0,29 %
19	MS	2	0,16 %	**19**	AL	11	0,19 %	**19**	AL	54	0,25 %
20	AC	1	0,08 %	**20**	RO	11	0,19 %	**20**	MA	52	0,24 %
21	AP	1	0,08 %	**21**	MT	9	0,16 %	**21**	SE	46	0,21 %
22	MT	1	0,08 %	**22**	SE	8	0,14 %	**22**	PI	30	0,14 %
	Total	**1242**		**23**	PI	7	0,12 %	**23**	RO	29	0,13 %
				24	AC	3	0,05 %	**24**	TO	14	0,06 %
				25	TO	2	0,03 %	**25**	RR	11	0,05 %
				26	AP	1	0,02 %	**26**	AC	9	0,04 %
				27	RR	1	0,02 %	**27**	AP	6	0,03 %
					Total	**5744**			**Total**	**21989**	

Yet, when we look at the state-level concentrations, São Paulo stands out. It is the leading center of Brazilian research, producing about 41% of all national reserch in 2005. Yet, it's share of top-25% cited papers is 44,08% and 45,57 of top-1% papers. In addition, some states have higher share of highly cited papers than their overall share of papers. These include Rio Grande do Sul, Distrito Federal and Santa Catarina.

Fast growing states of Minas Gerais and Parana, on the other hand, have lower share of high citing papers than they have from overall pool of papers. Minas Gerais produces about 8,69% of all Brazilian papers, but only 7,75% of top-25% papers, and Parana 5,30% and 4,72%, respectively. Their shares drop even further when we look at the top-1% of papers, but only modestly. To fully relate how growth relates to quality of research, we are working to expand our data and analysis to include all years.

C. Origin of research and research excellence

A central issue in building research capacity is the relationship between domestic knowledge procuction and collaborative research. Each WOS-ISI record includes the Re-print address, which can used as a proxy for the leader of the research project. While it is difficult to estimate the degree of leadership or ownership of authorship, the reprint address provides a proxy to identify among the contributing authors the one with most weight, acknowledging that his or her role may be only slightly more significant than that of others. Yet, it remains, that the authors have agreed to assign the reprint address to that person.

Table 9. breaks down the reprint address for all Brazilian 2009 papers by major regions and Brazil, and shows how the share of Brazilian papers declines in relation to citations received. Brazilian reprint address is 88% of all papers in 2009, slightly above the share in 2005 that was 85%. Brazil's share of reprting addresses in top-quartile is 78% (74 in 2005), and 66% in top-5% papers (57%) and 47% (35%).

Share of European and North American reprint addresses in 2009 papers follows the pattern from 2005, but remain at somewhat lower level. Europe accounts for second largest share of reprint addresses in all papers and top-quartile, but is overtaken by North American in Top-5% and Top-1% papers.

TABLE 8.
SHARE OF REPRINT ADDRESSES IN TOP 1%, 5%, 25%, AND ALL BRAZILIAN ARTICLES AND PROCEEDINGS FROM 2005.

	Top-1% 2005					Top-5% -2005		
Rank	**RP Country**	**Records**	**Share**		**Rank**	**RP Country**	**Records**	**Share**
1	North America	69	35,38 %		1	Brazil	557	56,95 %
2	Brazil	63	32,31 %		2	North America	205	20,96 %
3	Europe	53	27,18 %		3	Europe	164	16,77 %
4	Asia	7	3,59 %		4	Asia	19	1,94 %
5	Middle East	1	0,51 %		5	South America	16	1,64 %
6	Oceania	1	0,51 %		6	Oceania	8	0,82 %
7	South America	1	0,51 %		7	Middle East	5	0,51 %
	TOTAL	195	100,00 %		8	Central America and Carribean	3	0,31 %
					9	Africa	1	0,10 %
						TOTAL	978	100,00 %

TABLE 8.
CONTINUED

	Top-Quartile 2005						**2005 All records**	
Rank	**RP Country**	**Records**	**Share**		**Rank**	**Records**	**# Records**	**Share**
1	Brazil	3459	74,05 %		1	Brazil	15725	85,24 %
2	Europe	531	11,37 %		2	Europe	1263	6,85 %
3	North America	502	10,75 %		3	North America	976	5,29 %
4	South America	77	1,65 %		4	South America	211	1,14 %
5	Asia	55	1,18 %		5	Asia	122	0,66 %
6	Central America and Carribean	19	0,41 %		6	Central America and Carribean	66	0,36 %
7	Oceania	16	0,34 %		7	Oceania	44	0,24 %
8	Middle East	8	0,17 %		8	Africa	21	0,11 %
9	Africa	4	0,09 %		9	Middle East	20	0,11 %
	TOTAL	4671	100,00 %			**TOTAL**	18448	100,00 %

TABLE 9.
SHARE OF REPRINT ADDRESSES IN TOP 1%, 5%, 25%, AND ALL BRAZILIAN ARTICLES AND PROCEEDINGS FROM 2009.

	Top-1% 2009					Top-5% 2009		
Rank	**Country**	**Papers**	**Share**		**Rank**	**Country**	**Papers**	**Share**
1	Brazil	211	47,85 %		1	Brazil	1532	65,92 %
2	North America	114	25,85 %		2	Europe	375	16,14 %
3	Europe	95	21,54 %		3	North America	318	13,68 %
4	Asia	7	1,59 %		4	South America	45	1,94 %
5	South America	5	1,13 %		5	Asia	21	0,90 %
6	Oceania	4	0,91 %		6	Oceania	18	0,77 %
7	Middle East	3	0,68 %		7	Central America and Carribean	6	0,26 %
8	Central America and Carribean	2	0,45 %		8	Africa	5	0,22 %
	TOTAL	**441**	100,00 %		9	Middle East	4	0,17 %
						TOTAL	**2324**	100,00 %

TABLE 9.
CONTINUED

		Top-Quartile 2009			All 2009			
Rank	**Country**	**Papers**	**Share**		**Rank**	**Country**	**Papers**	**Share**
1	Brazil	7265	78,46 %		1	Brazil	27599	88,01 %
2	Europe	942	10,17 %		2	Europe	1691	5,39 %
3	North America	732	7,91 %		3	North America	1301	4,15 %
4	South America	153	1,65 %		4	South America	373	1,19 %
5	Asia	73	0,79 %		5	Asia	166	0,53 %
6	Oceania	37	0,40 %		6	Central America and Carribean	91	0,29 %
7	Central America and Carribean	26	0,28 %		7	Oceania	76	0,24 %
8	Africa	21	0,23 %		8	Africa	35	0,11 %
9	Middle East	10	0,11 %		9	Middle East	27	0,09 %
	TOTAL	**9259**	100,00 %			**TOTAL**	**31359**	100,00 %

Thus, papers with foreign reprint addresses contribute significantly to Brazilian reseach excellence. This reinforces the impression of the relative relevance and quality of European and North American research. Both regions remain main contributors to Brazilian research and its relevance, a difference between quality and excellence persists: Europe is the largest foreign source for reprint address for all Brazilian papers in 2005, top-quartile and top-5%, but is replaced by North America in Top-1%. Here, Brazil has highest share of reprint addressees, but this is relatively likely to change as there has been more time to award citations (compare **Table 6.).**

What the relative shares of reprint addressees do demonstrate is that North American and European research collaboration matters greatly for quality – when measured as high citation rates – for Brazilian research. Furthermore, it also shows that North American, especially American one, matters more for high cited Brazilian reserch than European one, adding one more perspective to the discussion of relevance and volume of these two major research areas.

VI. CONCLUSIONS

This paper has focussed on analyzing the knowledge in- and outflows of Brazilian research, as measured in citations made and received. Based on comprehensive Institute of Scientific Information data set on all Brazilian research publications between 2005 and 2009 (SOURCE data set), as well as two additional data sets consisting of all literature cited by SOURCE data set and of all literature making citations to SOURCE data. For this paper, the SOURCE data was limited to articles and proceedings, consisting of 127,826 papers. In addition, we added to the data state addresses to authors for about 25,000 records and resulting 98% coverage in the data.

Brazilian research is growing rapidly, and country's research system is being transformed with new funding, new and strengthened institutions and the overall build-up of the capabilities of Brazilian innovation system. One key aspect of this transformation is that the states Minas Gerais and Rio Grande do Sul, and some other catch-up states, are increasing their research output at almost double rate when compared with Sao Paulo and Rio de Janeiro.

Our analysis of citations received and citations made is suggestive of Brazil's relative standing to other countries and regions, as well as casts light on broader global shifts in knowledge production. Our data points out that Brazilian research makes most citations to European, Brazilian, North American and Asian research literature. These positions are mirrorred in the standing of regions making citations to Brazilian research from 2005.

Change over years in citations made by Brazilian papers reveal also how the relevance of research in major global research areas is evolving from the Brazilian perspective. A key outcome is that Brazilian research builds extensively on domestic capabilities, as Brazilian papers cite at very large numbers Brazilian research, suggesting that domestic capabilities are playing an important role and suggesting that Brazilian research community is able to shape up its own agenda.

It also shows that American research remains highly relevant for Brazilian research, whereas there is a drop in the use of European research between 2005 and 2009. This echoes other results suggesting that European research is surpassing USA in volume but trailing in quality. Furthermore, we observe that research from other emerging economies and developing countries is increasingly cited by Brasilian publications. Although the overall volume of South-South links is relatively weak, they are growing, and the strengthening role of China and India is especially noteworthy.

A central question for this paper has been to examine to what degree domestic capabilities drive Brazilian research and its excellence, and to that end we looked at the share of reprint addresses among papers. Reprint address is a vague suggestion of leadership role in writing the article, and in the data set it is assigned in well over 80% of papers to Brazilian author. Such a high

degree suggests a strong ownership of the research, but, on the other, this share drops quickly as we look at paper cohorts that received most citations, being mere one third within the top-1% and roughly half of the top-5% of all papers. Yet, in summary, Brazilian research appears to be its major pillar, contributing in significant and increasing manner to its nature as one of the emerging global knowledge hubs.

REFERENCES

[1] OECD and The International Development Research Centre, "Innovation and the development agenda," E. Kraemer-Mbula and W. Wamae, Eds., Paris, OECD Publishing, 2010.

[2] OECD, "Integrating science and technology into development policies: an international perspective," Paris, OECD Publishing, 2007.

[3] J. Chataway, J. Smith, and D. Wield, "Partnerships and building capabilities for science, technology and innovation and development in Africa." Research Centre on Innovation, Knowledge and Development, Open University, Milton Keynes, UK, Working Paper No. 2., 2005.

[4] UN Millennium Project, "Investing in development: A practical plan to achieve the millennium development goals," London, Earthscan, 2005.

[5] H. Toivanen and B. Ponomariov, "African regional innovation systems: bibliometric analysis of research collaboration patterns 2005-2009," *Scientometrics*, vol. 88, no. 2, pp. 471-493, 2011.

[6] N. Boshoff, "South-South research collaboration of countries in the Southern African Development Community (SADC)," *Scientometrics*, vol. 84, no. 2, pp. 481-503, 2010.

[7] M. Jeenah, A. Pouris "South African research in the context of Africa and globally," *South African Journal of Science* vol. 104, pp. 351-354, 2008.

[8] O. B. Onyancha and J. R. Maluleka, "Knowledge production through collaborative research in Sub-Saharan Africa: how much do countries contribute to each other's knowledge output and citation impact?" *Scientometrics* vol. 87, no. 2, p. 315-336, 2011.

[9] S. P. Borgatti, M. G. Everett, and L. C. Freeman, "Ucinet for Windows: Software for Social Network Analysis," Harvard, MA, Analytic Technologies, 2002.

[10] A. F. Helene and P. D. Ribeiro, "Brazilian scientific production, financial support, established investigators and doctoral graduates," *Scientometrics* Online First 2011.

[11] UNESCO, "UNESCO science report 2010: the current status of science around the world," UNESCO Publishing, Paris, 2010.

[12] J. Leta and H. Chaimovich, "Recognition and international collaboration: the Brazilian case," *Scientometrics* vol 53., no. 3, pp. 325-335, 2002.

[13] W. Glänzel, J. Leta, and B. Thijs, "Science in Brazil. part 1: A macro-level comparative study," *Scientometrics* vol. 67, no. 1, pp. 67-86, 2006.

[14] D. J. Schoeneck, A. L. Porter, R. N. Kostoff, and E. M. Berger, "Assessment of Brazil's research literature," Defense Technical Information Center, Fort Belvoir, VA, DTIC Technical Report Number ADA515318, 2007.

[15] D. J. Schoeneck, A. L. Porter, R. N. Kostoff, and E. M. Berger, "Assessment of Brazil's research literature," *Technology Analysis & Strategic Management*, vol. 23, no. 6, pp. 601-621, 2011.

[16] H. F. Moed, *Citation Analysis in Research Evaluation*. Dordrecht: Springer, 2005.

[17] H. Small, "On the shoulders of Robert Merton: Towards a normative theory of citation," *Scientometrics* vol. 60., no. 1., pp. 71-79, 2004.

[18] P. Albarrán, J. A. Crespo, I. Ortũno, and R. Ruiz-Castillo, "A comparison of the scientific performance of the U.S. and the European Union at the turn of the 21st century," *Scientometrics* vol. 85, pp. 329-344, 2010.

[19] C. K. N. Chang and A. Breitzman, "Using patents prospectively to identify emerging, high-impact technological clusters," *Research evaluation* vol. 18, no. 5, pp. 357-364, 2009.