Science in India. A bibliometric study of national and institutional research performance in 1991-2006

Wolfgang Glänzel[1,2] B.M. Gupta[3]

15 July 2008

Abstract

In the present paper, an analysis of Indian S&T has been presented using its publication output in international peer reviewed journals covered in Web of Science database. It analyses India's publication activity in terms of global share, share of international collaborative publications, and visibility & citation impact for the period 1991-2006. It explores how far the trends in India S&T output mirror those of the other upcoming countries and what the main differences among those countries are. It discusses the findings in the light of the above-mentioned ongoing discussion on decline or emergence of Indian science.

1 Introduction

The global landscape of science and technology is undergoing radical changes. The old world picture of the 20th century has already been upended. Technological and scientific development is closely following the economic changes. Emerging and re-emerging economies have very soon been brought into the focus of both the public and the experts' interest. In 2003, Jim O'Neill coined the term BRIC (as abbreviation for Brazil, Russia, India, China) and only two years later, in 2005, he extended this idea by introducing the conception of the Next-11 (N11) comprising a larger list of eleven upcoming economies. This alone indicates the tremendous speed with which the development takes place. Reports on economic research have soon been followed by science policy and technology related studies which attempted to capture also the growth of science and technology mirroring the economic dynamics of these countries. All evidence points to the fact that the centre of gravity of scientific and technological advancement is moving to regions outside Europe and North-America.

Beyond any doubt, China has achieved the most spectacular progress in economy, science and technology among these emerging countries. Already two years ago, China has announced to have overtaken France as the world's fifth largest economy. Now China is challenging Germany as World's export leader. This trend is paralleled by its growth of science. Since 2006 China ranks second in terms of the world's publication output in the sciences as reflected by the Web of Science database and China strives to become the world's leading science power by

[1]Steunpunt O&O Indicatoren, Katholieke Universiteit Leuven, Leuven, Belgium,
Wolfgang dot Glanzel at econ dot kuleuven dot be
[2]Institute for Research Policy Studies, Hungarian Academy of Sciences, Budapest, Hungary,
glanzw at iif dot hu
[3]National Institute of Science, Technology and Development Studies (NISTADS), New Delhi, India,
bmgupta1 at gmail dot com

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin
Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

2050. This most spectacular development cannot, however, disguise that, overshadowed by China's rise, other countries emerged as well; and these nations are gaining or re-attaining the position as main global players in economy, research and technology. Recently O'Neill (2005) forecasted that India might become the world's third largest economic power.

In an earlier study by Glänzel et al. (2008) light was shed upon the dynamics of five emerging countries (China, Korea, Taiwan, Brazil and Turkey) in the mirror of bibliometric indicators. This paper and the study of research performance of the 'BRIC' countries by Zitt et al., (2006) pointed at quite different situations among the emerging countries. Besides the somewhat ambiguous and debated situation of Russia and Brazil, also India's evolution is a matter in dispute (e.g., Arunachalam 2002; Gupta and Dhawan 2006, 2008; Kedamani et al. 2007; Basu 2007).

In the present paper, the evolution of publication activity, visibility and citation impact of India is studied for the period 1991-2006. In this study India's bibliometric profile is analysed in the context of global trends as well as in comparison with five other emerging economies of comparable weight, particularly, with Korea, Russia, Brazil, Taiwan and Turkey. We will attempt to answer the question of in how far the trends in India mirror those of the other upcoming countries and what the main differences among those countries are.

2 Data sources and processing

All bibliometric data are based on bibliographic data extracted from the Web of Science (WoS) of Thomson Scientific (Philadelphia, PA, USA). The period for publication activity and scientific collaboration comprises the years 1991-2006. In order to obtain stable and reliable results, the underlying publication period is split up into several sub-periods of 2-3 years each. For the citation analysis, we used a three-year window each for papers published in 1992-2004. Only document types indexed in the Science Citation Index Expanded (SCIE) named as Articles, Letters, Notes and Reviews were taken into account.

Publications are assigned to countries on the basis of their corporate addresses as indicated in the by-line of the publication. An integer counting scheme is applied, i.e., all countries appearing in the address field are considered and multiple occurrence of a country within the same publication is de-duplicated. This approach results in counting publications with (at least) one author with an affiliation in the corresponding country. This counting scheme is best suited for analysing both the countries' weight and their international co-publication links, but as a consequence of its application, publications cannot be summed up over countries to the world total (cf. REIST-2 1997).

As for subject classification, the hierarchical classification scheme developed by Glänzel and Schubert (2003) on the basis of ISI's journal assignment to Subject Categories is applied: Agriculture & Environment (AGRI), Biology (BIOL), Biosciences (BIOS), Biomedical research, Clinical & Experimental Medicine I (CLI1), Clinical & Experimental Medicine II (CLI2), Neuroscience & Behaviour (NEUR), Chemistry (CHEM), Physics (PHYS), Geosciences & Space Sciences (GEOS), Engineering (ENGN) and Mathematics (MATH). The level in between these major fields and the ISI Subject Categories comprises further 60 subfields. A science field has five subfields on average and a subfield aggregates about three ISI Subject Categories each.

For the citation analysis, a three-year citation window beginning with the publication year is applied for selected sub-periods of the above-mentioned publication period. Citations received by these publications have been determined on the basis of an item-by-item procedure, using special identification keys, made up of bibliographic data elements.

3 Methods and results

3.1 National research performance

Publication growth

In this section we focus on the trends and dynamics in the publication activity, the national publication profile, the evolution of citation

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin
Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

impact and the patterns and network international scientific collaboration of India. The set of advanced and standard indicators developed at ISSRU (Budapest) and SOOI/KU Leuven is used for the analysis. These measures comprise besides basic indicators and measures of activity and co-operativity, observed, field- and journal expected citation rates and relative citation rates, measures of high-impact activity and attractivity as well.

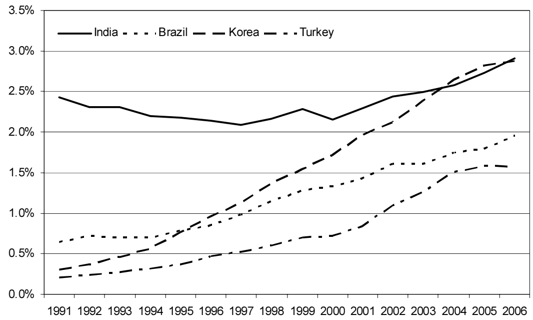

Figure 1 Evolution of India's publication output in terms of global share in the sciences in 1991-2006

The evolution of India's publication is long a matter in dispute output as has already mentioned in the outset. Indeed, the evolution of Indian publication activity with respect to the world total publication output does not provide an unambiguous picture. In particular, a period of decline (1991-1997) is followed by a short phase of stagnation around 1999 which, in turn, is followed by a period of strong growth (see Figure 1). It is not clear, in how far changes in the underlying database have contributed to these trends, but the pit in 1997-2000 cannot be explained with structural changes in the database alone. India's growth is not unambiguous; all one can conclude from the bibliographic data is that the increase of India's share in the world total of 19.4% during the last 15 years does simply not keep pace with the growth of others whose relative growth rate of their share amounts to even several 100% in the same period (cf. Figure 1).

The assessment of the evolution on the basis of ranks is certainly a difficult, even precarious endeavour. Nonetheless, we will have a look at the national ranking according to publication output in different years in the period 1991-2006 in order to gain insight in the changes of the global landscape of science and technology.

Table 1 Changing ranks of the world's leading contributors in publication output according to the SCIE

Country	Rank			
	1991	1996	2001	2006
USA	1	1	1	1
China PR	15	13	7	2
UK	2	2	3	3
Germany	4	4	4	4
Japan	3	3	2	5
France	6	5	5	6
Canada	7	6	8	7
Italy	8	7	6	8
Spain	12	10	10	9
Australia	10	9	11	10
India	**9**	**12**	**13**	**11**
South Korea	33	20	15	12
Netherlands	11	11	12	13
Russia	36	8	9	14
Brazil	22	22	17	15
Switzerland	14	15	16	16
Taiwan	25	19	18	17
Sweden	13	14	14	18
Turkey	38	31	25	19
Poland	18	18	19	20

Table 1 presents the 20 countries contribution most to the world total publication output. The observed sharp rise of the share of national publication output in the world total of countries like of China, South Korea, Brazil, Taiwan and Turkey is certainly unambiguous, and has already been largely discussed in the relevant literature (e.g., Zhou and Leydesdorff 2006; Zitt et al. (2006); Glänzel et al. 2008; Glänzel 2008). The same unambiguousness applies to Russia, however, indicating a negative development. Relatively slight changes as found for India can be due to several external factors, too, as, for instance, changes of database coverage as pointed at above. The discussion of the evolution of India's position was mainly a matter of national concern (e.g., Arunachalam 2002; Gupta and Dhawan 2006, 2008; Kedamani et al. 2007; Basu 2007). Nonetheless, India holds a firm position around rank ten according to the world's most important producers of scientific literature (cf. Table 1).

Publication profile

National publication profiles can preferably be measured and visualised using the Relative Specialisation Index (RSI). This measure indicates whether a country has a relatively higher or lower share in world publications in a particular science field than its overall share in the world total (see REIST-2 1997), and is closely related to the Activity Index (AI) introduced by Frame (1977). Its definition and interpretation can be found in Glänzel (2001), therefore, a detailed description of these indicators is omitted here. RSI takes values in the interval [–1, +1]; RSI < 0 (RSI > 0) indicates a lower-than-average (higher-than-average) activity. RSI = 0 reflects a completely balanced situation. RSI is an indicator measuring the internal balance, therefore RSI > 0 for some fields implies RSI < 0 for others and RSI = 0 for all fields corresponds to the 'world standard'. National 'publication profiles' are determined on the basis of the twelve major science fields introduced in the Data sources and data processing section. Since subject classification on the basis of journal assignment does practically not result in disjoint subject areas, the classification scheme does not form a partition of the total. The twelve-component profile is therefore preferably visualised in 'clockwork diagrams', where each 'hour' represents one field. The graphical presentation of the 'world standard', i.e., RSI = 0 for all fields, is a regular dodecagon. Deviations from this standard result in to some extent characteristic deformations of the regular octagon. In earlier studies (e.g., REIST-2 1997), four basic paradigmatic patterns in publication profiles could be distinguished, particularly,

I. the 'western model' with clinical medicine and biomedical research as dominating fields,
II. the characteristic pattern of the former socialist countries with prevailing activity in chemistry and physics,
III. the 'bio-environmental model' with biology and earth and space sciences in the main focus,
IV. the 'Japanese model' with engineering and chemistry being predominant.

India's profile does not uniquely fit in any of the above categories (see Figure 2). It can rather be considered a mixture of Types II and III. The evolution is characterised by two general trends, particularly, by growing relative activity in the life sciences (except biology) and decreasing weight of agriculture & environment, geo- and earth sciences, mathematics and engineering. Thus the subject profile steadily evolves towards a more balanced situation but the pronounced predominance of natural sciences to the disadvantage of the life sciences still persists.

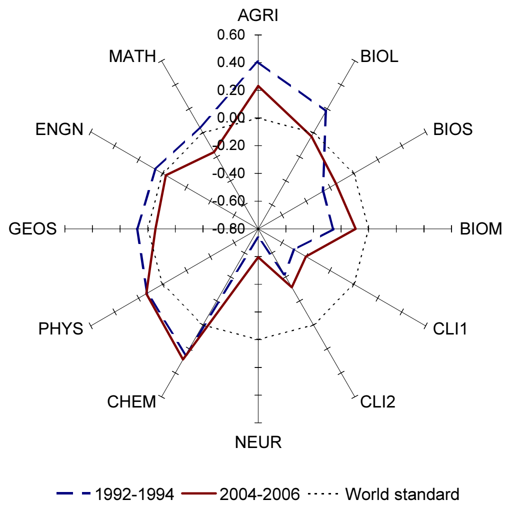

Figure 2 The change of India's publication profile over time

Citation impact

In this section we will have a look at the evolution of citation impact of Indian research in the life sciences, natural sciences and applied sciences. The following set of standard indicators is used for this analysis.

1. The Mean Observed Citation Rate (MOCR) is defined as the ratio of citation count to publication count.
2. Mean Expected Citation Rate (MECR). The expected citation rate of a single paper is defined as the average citation rate of all papers published in the same journal. MECR is defined as the average of these individual expectations over a given paper set.
3. Relative Citation Rate (RCR). RCR is defined as the ratio of the observed and journal-based expected citation impact. This indicator measures whether the publications of the unit under study attract more or less citations than ex-

pected on the basis of the journal impact measures of the journals in which they appeared.

4. Normalised Mean Citation Rate (NMCR). The field-expected citation rate of a single paper is defined as the average citation rate of all papers published in the same subfield. Since subject assignment is often not unique a fractional counting scheme is applied. NMCR is defined as the ratio of the observed and field-based expected citation impact. This indicator gauges citation rates of the papers against the standards set by the specific subfields. This indicator is based on the 60 subfields according to the above-mentioned SOOI/ISSRU classification scheme.

A detailed description of definition, application and interpretation of these indicators can be found in earlier papers (e.g., Glänzel et al 2003). We just mention here that MOCR = 0 implies RCR = 0 and NMCR = 0, and corresponds to uncitedness; RCR (NMCR) < 1 represents a lower-than-the-average, RCR (NMCR) > 1 a higher-than-the-average situation according to the corresponding reference standards. Finally RCR (NMCR) = 1 means that the papers received the number of citations expected on the basis of the average citation rate of the publishing journals (subfields). A large deviation of RCR from NMCR means that the journals in which authors of the country under study are on average publishing, do usually not conform to the corresponding subject standards. This deviation may be positive or negative.

Although the pit found in relative publication activity in the second half of the 1990s is paralleled by a similar in citation impact (see Figure 3), normalised and relative citation impact generally increased during about the last decade after the ‘pit’ around 1994. However, it did, on an average, not reach the world standard. This latter characteristic is shared with the citation patterns of other emerging countries. Although both relative indicators (RCR and NMCR) show the same trends, RCR always considerably exceeds the corresponding NMCR values. This means, that Indian scientists publish, on an average, in journals with lower impact that the corresponding field standard. In a recent study, we have found similar patterns for other emerging economies, particularly for China, Brazil, Taiwan, Korea and Turkey (Glänzel et al. 2008). Both subject- and journal-normalised impact of these countries increased while the RCR always exceeded the corresponding NMCR value. However, Taiwan and Korea are in a more advantageous situation than the other countries since both relative citation rates are here relatively close to each other (see Glänzel et al. 2008).

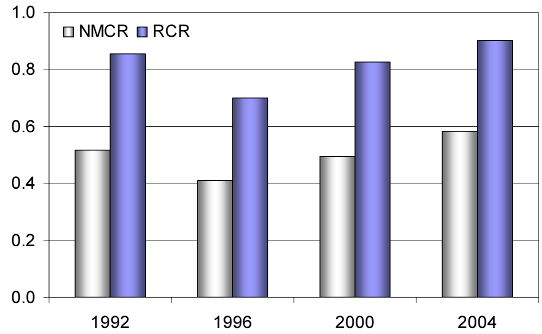

Figure 3 Evolution of citation impact of Indian publications

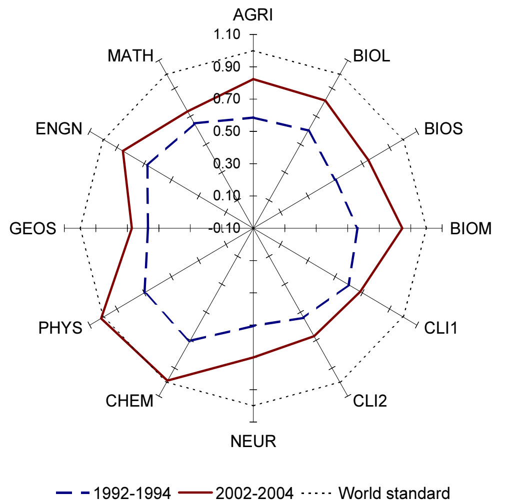

Figure 4 Change of India’s Relative Citation Rate over time

The breakdown of citation indicators by major fields substantiates that the increase of the relative citation rate has effect on all areas of the sciences. The effect is especially strong in the natural sciences where it already reached the world standard in 2002-2004 where observed citation impact has become in line with the journal-based expectation (see Figure 4).

International scientific collaboration

The duality of the co-authorship/co-operation relationship has long been discussed in the bibliometric literature (e.g., Katz and Martin 1997; Laudel 2002). At the level of individual authorship (Laudel) or at the institutional level (Katz and Martin), co-authorship does not depict research collaboration entirely, or might be distorted by scientists' multiple affiliations. Nonetheless, co-authorship proved a good proxy for 'higher-level' research collaboration between institutions, regions, and countries. Above all, international collaboration is usually well acknowledged in the published literature, and therefore a good indicator of co-operation at this level as well (Glänzel and Schubert 2004).

Several aspects of international collaboration are of special interest; besides the extent and share of international co-publications, the geopolitical profile of co-publications and the impact of collaborative research are preferred topics of bibliometric analysis. The dramatic intensification of international co-operation at all levels and the structural changes in the collaboration network has repeatedly been reported by several studies. An overview of this literature can be found, among others, in a recent study by Glänzel and Schubert (2004).

A first look at the publication data reveals a relatively low level of India's international co-operativity. In contrast to the general trend of intensifying collaboration, India's share of international co-publications in all papers has changed little. It ranged between 12% and 20% over the last 16 years and, again, there was a pit in the middle of the 1990s.

According to the regularity concerning the relation between foreign co-authorship ratio and the countries size found by Schubert and Braun (1990), one would expect a similar share of 'international papers' for India as found for countries of like size. Table 2 presents the corresponding percentages for all countries with 15,000-37,000 papers in 2006. India has the lowest co-operativity among these countries. The breakdown by partner countries, however, reveals slight structural changes in collaboration pattern.

Table 2 Share of internationally co-authored papers in all papers of selected countries in 2006

Country	Papers	Share
Australia	28778	42.0%
India	**28543**	**20.2%**
South Korea	28187	26.2%
Netherlands	24593	49.4%
Russia	21734	39.8%
Brazil	19135	30.0%
Switzerland	17973	60.0%
Taiwan	17842	18.7%
Sweden	17146	51.5%
Turkey	15392	16.1%

Table 3 Share in all Indian 'international papers' and strength of co-publications with India's most important partners (S=strong; M=medium)

Coutry/Region	1992-1994		2004-2006	
	Share	Str.	Share	Str.
EU15	*42.1%*	M	*40.6%*	S
USA	37.9%		34.5%	S
Germany	13.4%		14.9%	M
UK	12.8%		11.5%	M
Japan	6.8%		11.4%	M
France	7.1%		7.8%	M
South Korea	0.9%		6.2%	M
China PR	1.7%		5.4%	
Canada	8.9%		5.0%	
Italy	6.0%		4.4%	
Australia	2.8%		4.2%	
Taiwan	0.6%		3.6%	
Switzerland	2.7%		3.5%	
Russia	2.1%		3.2%	
Netherlands	2.9%		3.0%	
Spain	2.0%		2.8%	
Brazil	1.4%		2.4%	
Poland	0.7%		2.2%	
Sweden	2.1%		2.0%	

Table 3 shows the weight of India's ten most important partners in all papers with foreign partners in the two periods 1992-1994 and 2004-2006. The most impressing change concerns collaboration with the US and the EU . Co-operation with the US and the EU weakened but not all members of the EU are concerned. Collaboration with the Germany, France and Spain

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin
Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

slightly increased. Another remarkable trend concerns Japan, Taiwan, Brazil, China and Korea; the increase of collaboration with the latter four countries substantiates the growing importance of the emerging economies as partners for India. In spite of these changes, scientific collaboration with the EU and the US prevails.

The strength of bilateral links in the network of Indian international co-publications can readily be measured by Salton's cosine measure r (cf. Glänzel and Schubert, 2004). Two different thresholds ($r = 0.01$ and $r = 0.02$) are used to characterise different intensities of co-operation. The interval $0.01 \leq r < 0.02$ corresponds to medium strong links (M), whereas $0.02 \leq r$ corresponds to strong (S) links. The strength of links with the most important partners has in general increased. The strength of bilateral co-authorship links is given as supplementary indicator in Table 3.

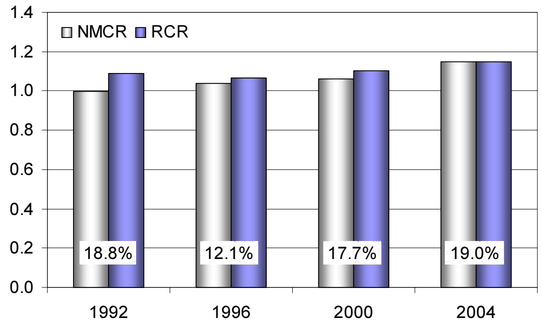

Figure 5 Evolution of citation impact of Indian international co-publications (percentage of 'international papers' is indicated on the bars)

Finally, we have a look at the impact of collaborative research. The comparison of the citation impact attracted by international co-publications with the 'national standard' confirms the expectations according to which international collaboration increases visibility and observed impact. Figure 5 presents the two relative citation rates for international co-publications. The results can directly be compared with those presented in Figure 3. On average, India clearly benefits from foreign co-authorship. The relative citation impact has slightly increased and exceeds the "standard value" of 1.0 during in the period 1992-2004. Except for the most recent available publication year, the impact of the journals where the international co-publications appeared still remains somewhat below the corresponding field-expectations.

4 Conclusions

After a down-leap in the middle of the 1990s India's science recovered and turned into steady growth in the new millennium. The growth does, however, not reach that of the other emerging economies like China, Korea, Taiwan or Brazil.

The relative growth with respect to the world total, which extends to all major fields of sciences, is accompanied by an increase of visibility and citation impact; however, the gap between journal and subfield standard still remains. Co-authorship with the EU and the US has lost weight in favour of collaboration with the Far East.

References

Arunachalam, S. 2002, Is science in India on the decline? *Current Science*, 83 (2), 107-108.

Basu, A. (2007), Scientific productivity and citation of scientific papers: Where do we stand? *Current Science*, 93 (6), 750-751.

Frame, J. D. (1977), Mainstream research in Latin America and the Caribbean, *Interciencia*, 2, 143-148.

Glänzel, W., Schubert, A. (2004), *Analyzing scientific networks through co-authorship*, In: H.F.M. Moed, W. Glänzel, U. Schmoch (Eds), Handbook of Quantitative science and Technology Research. The use of Publication and patent statistics in studies on S&T Systems. Kluwer Academic Publishers, Dordrecht, The Netherlands, 257-276.

Glänzel, W., Debackere, K., Meyer, M. (2008), 'Triad' or 'Tetrad'? On global changes in a dynamic world, *Scientometrics*, 74 (1), 71-87.

Glänzel, W. (2008), Turkey on the way to the European Union? On a scientific power rising next door, *ISSI Newsletter*, 4 (1), 10-17.

Glänzel, W., Thijs, B., Schubert, A., Debackere, K. (2009), Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance, *Scientometrics*, forthcoming.

Gupta, B. M., Dhawan, S. M. (2006). *Measures of progress of science in India: An analysis of publications output in science and technology*. New Delhi; Office of the Principal Scientific Advisor to the Government of India, November 2006 (PSA/2006/4), Accessible via: http://psa.gov.in

Gupta B. M., Dhawan, S. M. (2008). *Status of India in science and technology as reflected in its publications output in Scopus international database, 1996-2006*. Submitted for publications in Scientometrics (February 2008)

Katz, J. S., Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.

Kedamani, B. S., Sagar, A., Vijay Kumar, Gupta B. M. (2007). Mapping of Indian publications in S&T: A scientometrics analysis of publications in Science Citation Index. *DESIDOC Bulletin of Information Technology*, 27(1), 17-34.

Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11, 3–15.

O'Neill, J. (2005), Attention Europe: The BRICs are coming! *Internationale Politik*, 60 (5), 78-79.

REIST-2 (1997), *The European Report on Science and Technology Indicators 1997*. EUR 17639. European Commission, Brussels.

Schubert, A., Braun, T. (1990), International collaboration in the sciences, 1981-1985, *Scientometrics*, 19 (1-2), 3-10.

Zitt, M., Thèves, J., Laurens, P., Ramanana-Rahary, S., Bassecaulard, E., Filliatreau G. (2006), *Assessing international visibility of nations in science: citations performances of large emerging actors (BRIC)*. Poster presented at the 9th Science & Technology Indicators Conference, 7-9 September Leuven (Belgium).

Zhou, P., Leydesdorff, L. (2006). The emergence of China as a leading nation in science, *Research Policy*, 35(1), 83-104.