

# Under-reporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice

Ismael Rafols\*, Tommaso Ciarli\*\* and Diego Chavarro\*\*\*

\*[i.rafols@ingenio.upv.es](mailto:i.rafols@ingenio.upv.es)

Ingenio (CSIC-UPV), Universitat Politècnica de València, València, 46022, Spain  
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL, UK  
Observatoire des Sciences et des Techniques (OST-HCERES), Paris, France

\*\*[t.ciarli@sussex.ac.uk](mailto:t.ciarli@sussex.ac.uk)

SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL (UK)

\*\*\*[d.chavarro@sussex.ac.uk](mailto:d.chavarro@sussex.ac.uk)

SPRU (Science and Technology Policy Research), University of Sussex, Brighton, BN1 9SL (UK)

## Abstract

Although the main bibliometric databases (Web of Science (WoS) and Scopus) claim to include journals on the basis of scientific and publication standards, there have long been concerns that its coverage is biased in favour of journals from industrialised countries and towards topics relevant to these countries. In this article, we investigate this claim for research on rice, comparing the database CAB Abstracts with the mainstream databases. We find clear evidence that for a field such as rice, statistics based on WoS and Scopus strongly under-represent the scientific production by developing countries, and over-represent production by industrialised countries. More importantly, we also find a substantial bias in coverage of different research topics. WoS and Scopus have a ~75% coverage of publications in molecular biology and issues related to consumption, but a much lower coverage (20-30% in WoS and 30-50% in Scopus) for research more directly related to rice production such as plant nutrition, diseases and characteristics. CAB Abstracts coverage is above 80% for all topics except consumption. The study suggests that statistics based on mainstream databases provide a significantly distorted view of the amount of research and diversity of agendas in most countries. Given that bibliometric statistics are often used for benchmarking and evaluation purposes, the database biases may translate into policy framings that undervalue domestic capabilities and research agendas more attuned to local needs in the global south.

## 1. Introduction

There is an increasing demand for science to help in addressing grand challenges or societal problems, such as tackling obesity, climate change or pandemics. In this context, it becomes important to understand the different types of research that can help to tackle these problems, and to deliberate towards which directions scientific research should be developed. A useful starting point is to investigate what is the existing science “supply,” and which research options might be better aligned to address grand challenges and societal “demands” (Sarewitz and Pielke, 2007). In order to map the science supply, a representation of the knowledge on research topics relevant for a problem is needed.

Bibliometrics can provide helpful tools for developing knowledge representations of existing science “supply”. However, these representations are dependent on the data and methods used. As a result, bibliometric tools or indicators often reproduce the selection choices in data collection and treatment. For example, it has been shown that conventional bibliometric analyses are biased against non-English languages (Van Leeuwen et al., 2001; Archambault et al., 2006), developing countries (Arvanitis and Chatelin, 1988), applied research (Van Eck et al., 2013), the social sciences and humanities (Martin et al., 2010) and interdisciplinary research (Rafols et al., 2012). The aim of this paper is to investigate the biases introduced by available databases in the representation of research topics, in particular in relation to developing countries and topics potentially relevant to disenfranchised populations.

In spite of notable differences between the two mainstream databases, comparisons between WoS and Scopus (which has broader coverage) have been shown to produce similar ordering in rankings regarding country production over different fields. This similarity in ranking has led some analysts to claim that ‘indicators of scientific production and citation at the country level are stable and largely independent of the database’ (Archambault et al., 2009, p. 1320). Thus, in international benchmarking of science by intergovernmental agencies, the main databases WoS (e.g UNESCO, 2010) and Scopus (e.g Royal Society, 2011) continue to be used. For example, in the 2014 Globelics conference in Addis Ababa, the African Observatory of Science, Technology and Innovation (AOSTI) presented a study on African publications based only on Scopus data in spite of Scopus of its very low coverage of African journals (AOSTI, 2014).

However, science policy analysts have repeatedly warned of that the partial coverage of the main commercial databases may lead to serious misrepresentation of science in developing countries (Arvanitis and Chatelin, 1988; Albornoz et al., 1999; Chavarro, 2013). Recommendations have been made on the need to improve scientometric indicators in order to “properly evaluate global science” even by studies using the main databases (e.g. Royal Society, 2011, p. 107, which used Scopus).

Agricultural research has been and remains an important endeavour in developing countries, as it is seen as a potential source of knowledge and innovation crucial for social and economic development. Therefore, scientists are under considerable pressure from client groups” so as to solve local problems rather than contribute to the development of “universal” knowledge (Velho, 1990, p. 503). The main reason for focussing on rice is that we would like to monitor a relatively basic crop (although the technology behind research on agricultural crops is far from basic), which serves a large number of people with different needs in different parts of the world. Rice is a crop (i) which feeds a huge number of people around the world, particularly in low and middle income countries; (ii) which was at the core of the *green revolution*, particularly in the 60's and 70's, when high yield varieties of rice were investigated and distributed across the world to reduce the problem of famines in low income countries; (iii) and which, being the symbol of the green revolution is also a controversial technology due to the negative effects such as impoverishment of diets, overuse of water, exhaustion of soils, pollution, etcetera. Given the applied orientation of a lot rice research, the local specificity of the topics and the relative lack of relevance of the topic for many developed countries, rice is an interesting case to test the extent to coverage by main bibliographic databases (Velho, 1986, 1990).

In this article, we investigate biases for rice research comparing coverage of WoS and Scopus with CAB Abstracts (CABI), a database specialised in agriculture and environment with a broader coverage of developing countries. To our knowledge, this is the one of the few studies using CAB Abstracts for a bibliometric comparison. MacDonald (1984) and Kawasaki, 2004) had carried out two case studies previously, while Albornoz et al. (1999) had carried out a full comparison of coverage by of different databases for Ibero-American countries. In agreement with these previous studies, we find that WoS and Scopus strongly under-represent the scientific production by researchers in developing countries, and over-represent that by researchers in industrialised countries.

The second and possibly more important contribution is to describe a substantive bias in coverage of different research topics. WoS and Scopus have a ~75% coverage of publications in molecular biology or related to consumption, but they have a much lower coverage (20-30% in WoS and 30-50% in Scopus) for research related to issues more directly relevant for production such as plant nutrition and pests. But CAB Abstracts coverage is above 80% for all topics except consumption.

Third, we explore the potential effects of the biases in policy. Given that bibliometric statistics are often used for benchmarking and evaluation purposes, we propose that the biases reported may translate into policy framings that undervalue domestic capabilities and research agendas more attuned to local needs in the global south. We discuss whether and how this might affect research prioritisation.

## 2. Methods and data

WoS and Scopus are well known mainstream databases provided by large information and publishing companies, Thomson-Reuters and Elsevier respectively. CAB Abstracts (<http://www.cabdirect.org/>) is a database focused on environment and agriculture. It is run by CABI, an inter-governmental, not-for-profit organization that was set up by a United Nations treaty, with 48 member countries (many of them belonging to the Commonwealth), with a mission of “providing information and applying scientific expertise to solve problems in agriculture and the environment”.<sup>1</sup> Therefore, both CAB Abstract (for agriculture and environment) and Global Health (for public health) are aimed at facilitating the retrieval of relevant information for practitioners, very much as the database MEDLINE for medical research (supported by the US National Institutes of Health), but with a focus on development.

Publications on rice were manually downloaded from the WoS (including SCI-Expanded, SSCI, A&HCI, CPCI-S i CPCI-SSH) searching “rice” or “oryza” in the field “topic”. Scopus records were also manually downloaded searching in title, abstract or keywords, i.e. TIT-ABS-KEY (“rice” OR “oryza”). Similarly, documents with “rice” or “oryza” were searched in title and abstract of the database CAB Abstracts. (CAB Abstract allows to retrieve documents classified as strongly related to “rice” in the field “Descriptor”, but for consistency with the search on WoS and Scopus, we did not use this option here). The records of the different databases were matched with multiple matching algorithms. The analysis was carried out using VantagePoint, the statistical package R and the visualisation programme VOSviewer.

---

<sup>1</sup><http://www.cabi.org/about-cabi/> (Retrieved March 1<sup>st</sup>, 2014).

Here we first present a set of descriptive statistics, providing information on the coverage in terms of number of publications by document type, language and year. Second, we focus our analysis in the comparison of the number of publications by countries and research topics of publications. We use the main author affiliation to retrieve information on the country. An important caveat is that CABI only reports the affiliation of the first author. In the case of WoS and Scopus, on the contrary, the affiliations of all authors are included. As a result, the shares of countries will tend to be higher in WoS and Scopus. In this preliminary version, this effect has not been corrected. The error is estimated (using small document samples) on a 10%-30% over-representation, depending on country. Third, we cluster rice research into six field on the basis of a co-occurrence of terms in abstracts and titles. Each publication is then fractionally assigned to a cluster and estimates are made of number of publications per cluster and country.

### 3. Description of document characteristics: type, language and year

Let us first describe the main differences between the documents retrieved from each database. Given that each database classifies documents into different type categories, we downloaded all the document types, with statistics described in Table 1. It is found that in all cases, journal articles have a dominant share, between 81% (WoS) and 94% (Scopus). Hence, the results that follow will be mainly explained by differences in the coverage of journals used to index articles. The second most important document type is conference proceedings/papers, which make between 3.5% (Scopus) and 7.3% (CABI) of the total publications. The category “Miscellaneous” in CABI (4.7%) deserves further investigation.

Table 1. Share of publications by document type in the three databases investigated.

WoS			Scopus			CAB Abstracts		
Doc type	%	Cum%	Doc type	%	Cum%	Doc type	%	Cum%
Article	81.2%	81.2%	JOUR	93.7%	93.7%	Journal article	84.8%	84.8%
Proceedings Paper	7.1%	88.3%	CONF	3.5%	97.2%	Conference paper	6.8%	91.6%
Review	3.4%	91.7%	SER	1.6%	98.8%	Miscellaneous	4.7%	96.3%
Meeting Abstract	2.7%	94.4%	INPR	0.9%	99.7%	Book chapter	2.0%	98.3%
Note	2.4%	96.8%	CHAP	0.3%	99.9%	Book	1.9%	100.2%
Book Review	1.6%	98.4%	BOOK	0.1%	100.0%	Annual report	0.9%	101.1%
Editorial Material	0.7%	99.1%				Bulletin	0.6%	101.7%
Letter	0.6%	99.6%				Conference proceedings	0.5%	102.2%
Correction	0.3%	99.9%				Bulletin article	0.4%	102.7%

**Note:** Some documents are classified into more than one category. For example, many CABI conference papers are also Journal articles (this is why cumulative counting is higher than 100%).

In terms of language, as shown in Table 2, CABI is much more comprehensive than WoS (Scopus data will be available soon), with almost 10% of the documents in Chinese, and 6.7% in

Japanese. WoS only covers a few journals in Japanese (2%) and Portuguese (1%). If we consider the actual number of publications rather than the ratio within the data base, the difference is even larger. For example, CABI has 7 times the number of publications on rice in Japanese and 5 times the number of publications on rice in Portuguese.

Table 2. Share of publication by original language.

Language	CABI		WoS	
	# docs	%	# docs	%
English	148577	71.84%	92554	94.93%
Chinese	20544	9.93%	490	0.50%
Japanese	13844	6.69%	2032	2.08%
Portuguese	5356	2.59%	1015	1.04%
French	3942	1.91%	560	0.57%
Spanish	3320	1.61%	307	0.31%
Korean	3018	1.46%	31	0.03%
Russian	2396	1.16%	162	0.17%
Italian	1546	0.75%	22	0.02%
German	1462	0.71%	214	0.22%
Persian	501	0.24%	0	0.00%
Dutch	440	0.21%	9	0.01%
Thai	421	0.20%	11	0.01%
Indonesian	285	0.14%	0	0.00%

**Note:** % documents is computed only over the documents with language reported in the database (98% in WoS, 91% in CABI).

Trends over time show that CABI has had historically a much broader coverage than WoS and Scopus, as shown in Figures 1 and 2. Before the 1980s, coverage by WoS and Scopus of publications on “rice” is very limited. CABI shows a great increase in rice publications from the postwar until the mid 1970s, particularly after the mid 1960s. This is in agreement with the diffusion of the “green revolution”. The postwar expansion is followed by a period of slow growth from 1975 until 2000, when a renewed growth is observed (perhaps in coincidence with the advent of genomic studies). Since the mid 1990s WoS and Scopus have been catching up with CABI and by 2012, WoS reaches 80% of CABI and Scopus 86% in total number of publications, though with substantial non-overlapping coverage (not shown).

Figure 1: Number of publications on rice per year by database from 1902 until 1975.

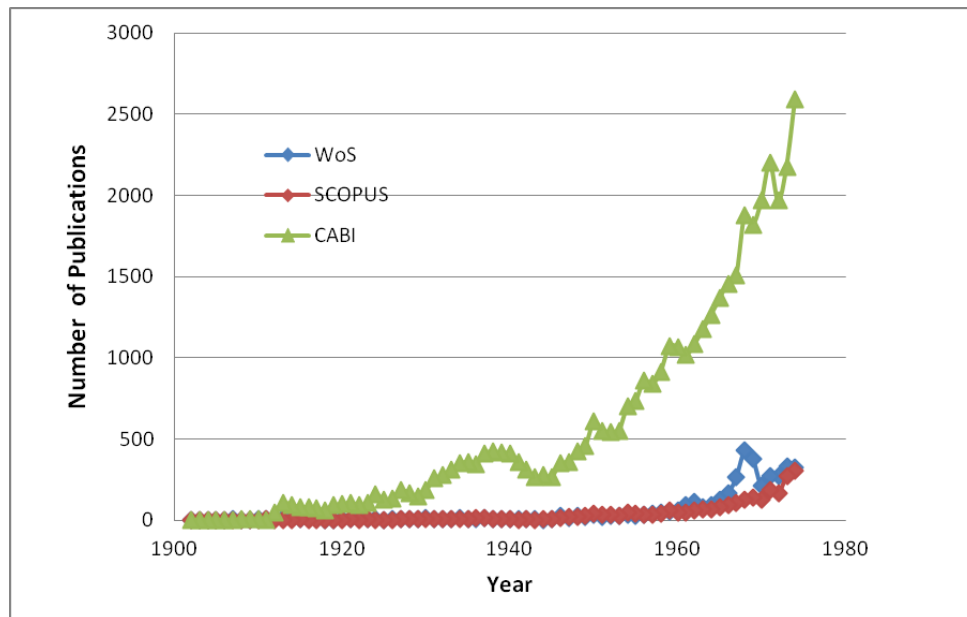
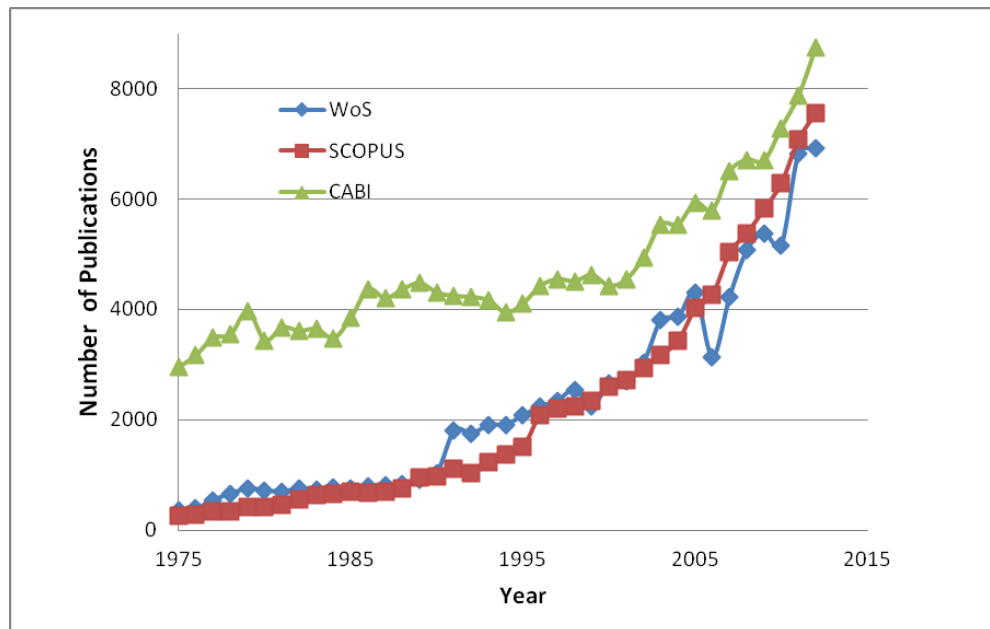


Figure 2: Number of publications on rice per year by database from 1975 until 2012.



#### 4. Comparison of coverage across countries

The countries publishing the most on rice are India, China, Japan and the US. China's publications on rice have sharply increased in the last twenty years (as expected from global publication trends, cf. Leydesdorff, 2012) whilst the shares by India, Japan and the US have decreased, as shown in Figure 3. Interestingly, all three databases agree on these trends.

However, there are major differences in the overall proportion of publication assigned to each country in each database. In the case of CABI, India was the most productive country until was caught by China in 2004. But whereas India's publications made 21% of the total in 2000-09 according to CABI, they represent less 9.6% and 8.4% in Scopus and WoS, respectively, as illustrated in Figure 4. Similarly, China's publications were 23% of CABI's publications, but only 16% and 13% according to Scopus and WoS. Oppositely, US publications were only 7% in CABI, but 15% and 16% in Scopus and WoS. Japan stands in the middle, with only a  $\pm 1\%$  difference depending on the database used. The differences in coverage between databases have narrowed in recent years, as shown in the right side of figure 4, but there is still a 2-fold difference in the percentage of publications assigned by CABI and WoS for the US and India.

Figure 3: Publications trends by country according to CABI data.

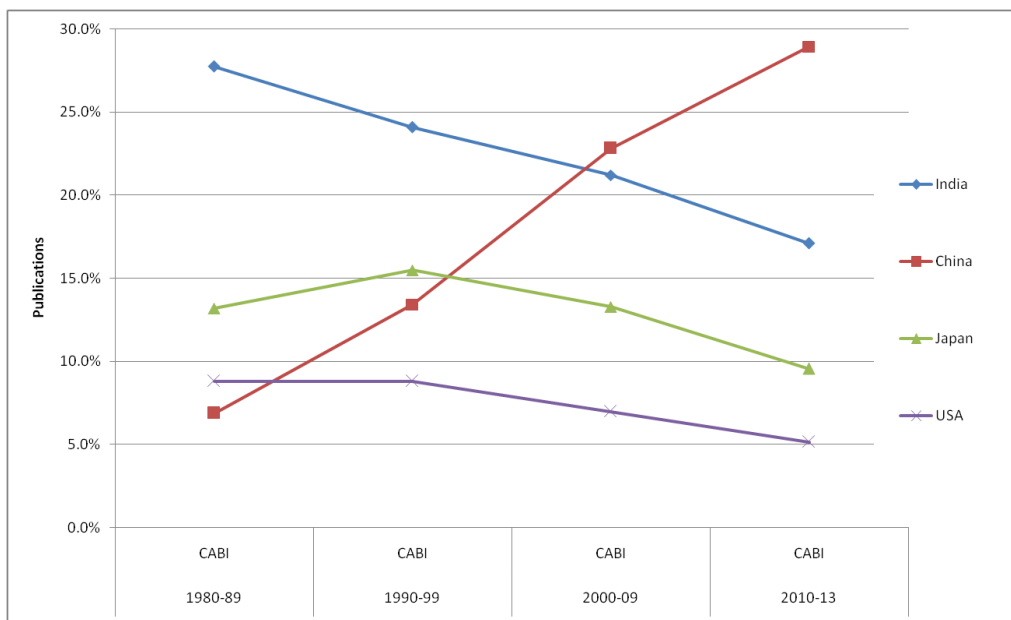


Figure 4: Percentage of publications on rice for large countries for different databases, in two periods.

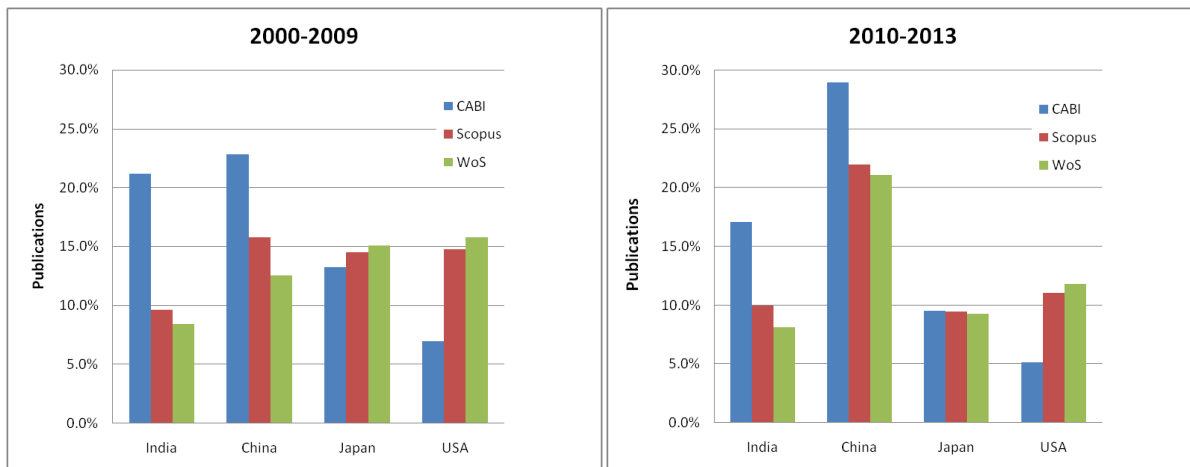


Figure 5: Percentage of publications on rice by countries for different databases, in two periods. Left hand side: countries with a relative higher CABI coverage. Centre: countries with similar coverage. Right hand: countries with higher coverage in WoS.

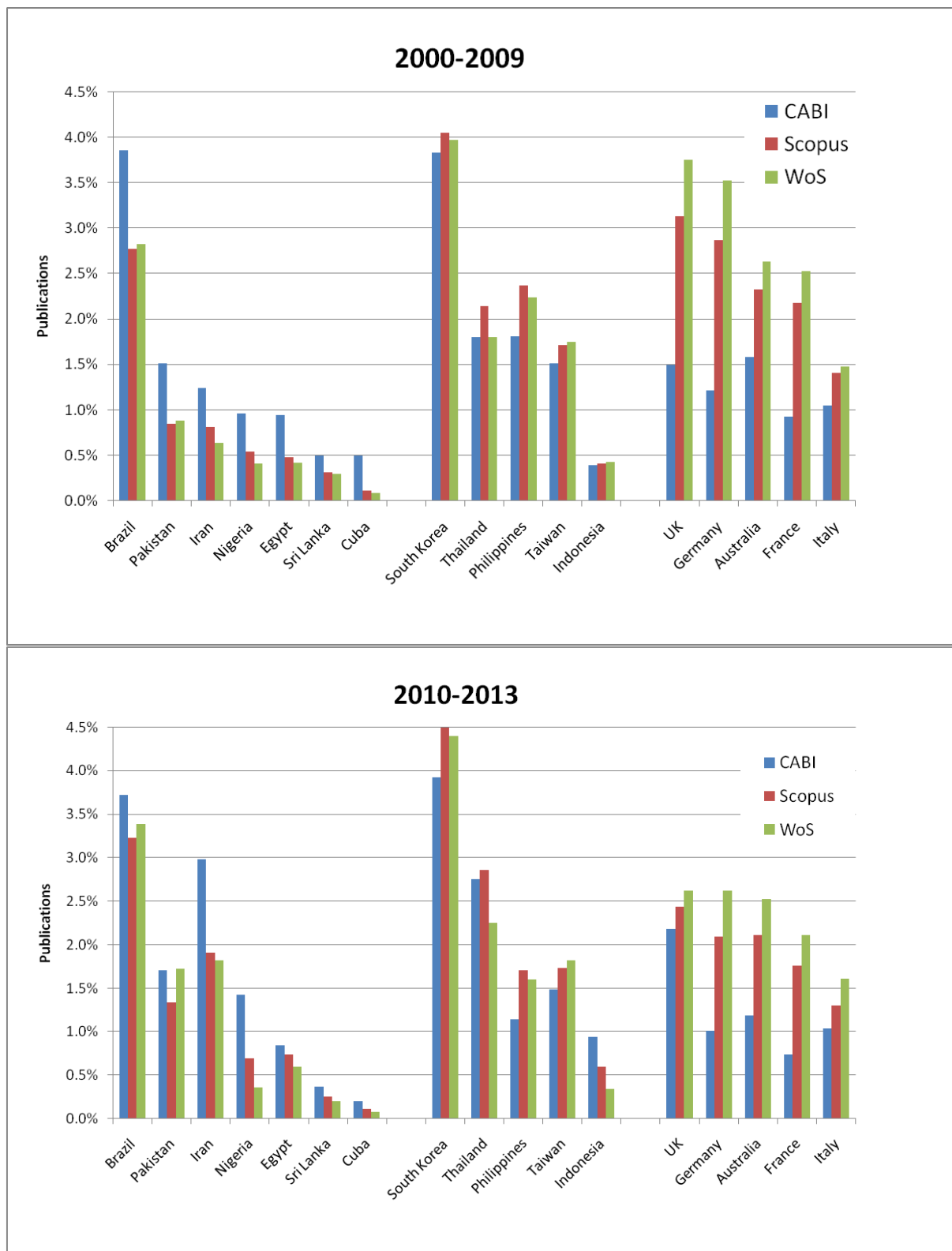




Figure 5 shows the percentage of publications for countries with smaller number of publications of rice for the three databases. As in the case of the large countries, we observe three patterns. On the left hand side, we show developing countries with much higher coverage in CABI. On the right hand side, we present industrialised countries with a much higher coverage in the WoS. In the middle, we have middle income countries from Asia that score similar shares in all three databases. In the latest period (2010-13) the differences between the countries are significantly reduced in some of the countries (e.g. in Brazil, perhaps due to incorporation of Brazilian journals into the commercial databases), but not in others (e.g. Cuba, Iran and Nigeria). In the case of Western countries, the differences in coverage mostly remain, while in the Asian middle income countries no clear trend is observed.

From the analysis of Figure 4 and 5 it follows that WoS and Scopus cover research published in North-America and Europe, whereas CABI is much more comprehensive in developing countries. As a result of CABI's larger coverage, Western countries relative contribution to scientific production on rice is much smaller than is usually acknowledged when using standard publication databases such as WoS and Scopus.

#### **4. Comparison of coverage across research topics**

The next question is the evenness in coverage across topics of the various databases. It is known that WoS for example has a wider coverage of the biomedical sciences, whereas Scopus has a relative better coverage of the social sciences and humanities than WoS. CABI, by construction, is expected to have a more comprehensive coverage of agronomy and environment. The problem for making a comparison across disciplines or topics is that a shared disciplinary classification across the three databases is needed –but of course, each database provides its own classification. In order to have a shared topic view across the three databases with a classification relevant to rice research, we pull together all the articles with abstract for the period 2002-2012 (a total of 78,225 articles) and carry out a clustering of terms using the freeware VOSviewer.

As shown in Figure 6, six research topics are found.:

- 1) a consumption cluster (red, bottom-left) that includes human nutrition (diet) and food processing (starch).
- 2) a cluster of productivity and plant nutrition (green, top-left) with research aiming to increase crop yield as well as socioeconomic issues.
- 3) a molecular biology cluster (blue, bottom-right), including genomics and transgenic research.
- 4) a genetics cluster (yellow, top-right) that appears to reflect hybridization approaches
- 5) a cluster describing plant characteristics (pink, top), such as panicle and grain length
- 6) a cluster related to diseases, pests and related efforts to protect the rice plant (turquoise, centre)

Given that this is a bottom up method for classifying with potential robustness problems, it was triangulated for with the results of an alternative clustering algorithm, based on direct semantic similarity between publications. The clusters differ slightly (for example the consumption cluster splits between food processing and human nutrition), but the overall findings are consistent.

Figure 6. Map of co-occurrence of terms in rice research in 2002-2012, with colours representing different research topics. Terms in near positions tend to appear together in abstracts and titles. The size of the nodes illustrates the number of times that terms occur.

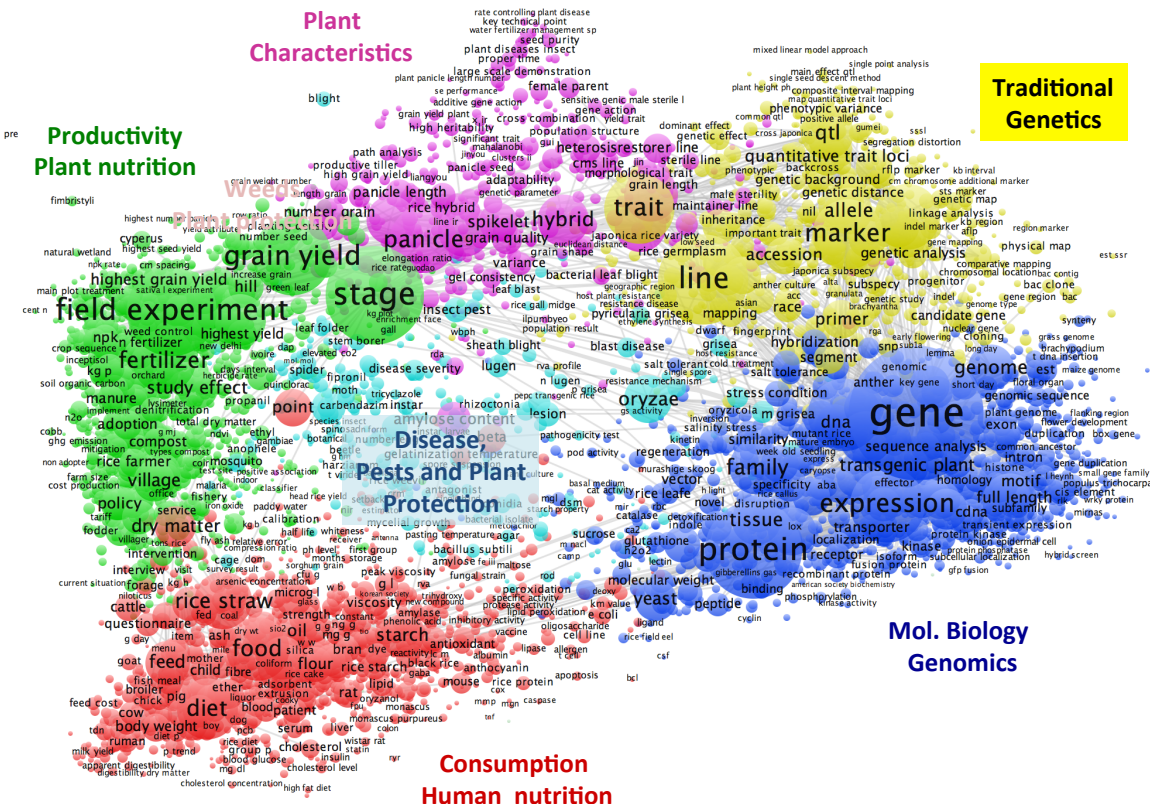


Figure 7. Relative coverage of each database for a given research topic found in Figure 6 (above). The reference set (full coverage) is based on the combination of the three databases.

VOSviewer

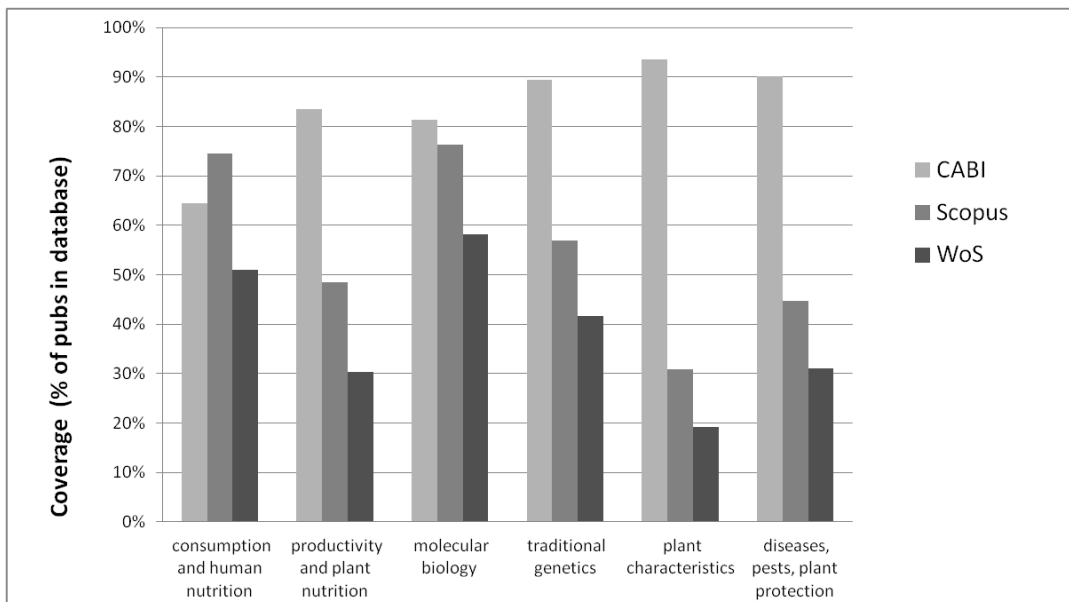


Figure 7 shows the coverage of the different databases per research topic. CABI coverage remains above 80%, except for the topic of consumption with a 65%. While the coverage of Scopus is above 70% for molecular biology and consumption, it is in the range of 30-50% for rice production, pests or plant characteristics that are more directly related to improving yields. Similarly WoS coverage is very low (20-30%) for production-related topics, but double for molecular biology and consumption (50-60%).

One may wonder the reason for such a disparate degree of coverage between topics by mainstream databases. An initial hypothesis is that the topics with low coverage are those where developing countries publish relatively more, and those with high coverage in mainstream are those where developed countries publish more. In table 3, we show the percentage of publications by country in each topic. At first look, it seems indeed that low income countries tend to focus relatively in the productivity cluster, whereas most developed countries focus in the molecular biology cluster, which account precisely for high and low coverage. Interestingly, though, there are some middle income countries such as Thailand and Malaysia with a high percentage of publications in the consumption area, which has a high coverage by WoS and Scopus. It seems thus, that publication pattern of countries do follow the social needs of their territory, as reported by Velho and Krige (1984) on the basis of interviews. Hence, in general there is a loose relation between lack of coverage and most of the topics relevant to developing countries, with the exception of consumption. But there are important corrections depending on the specific needs of the country.

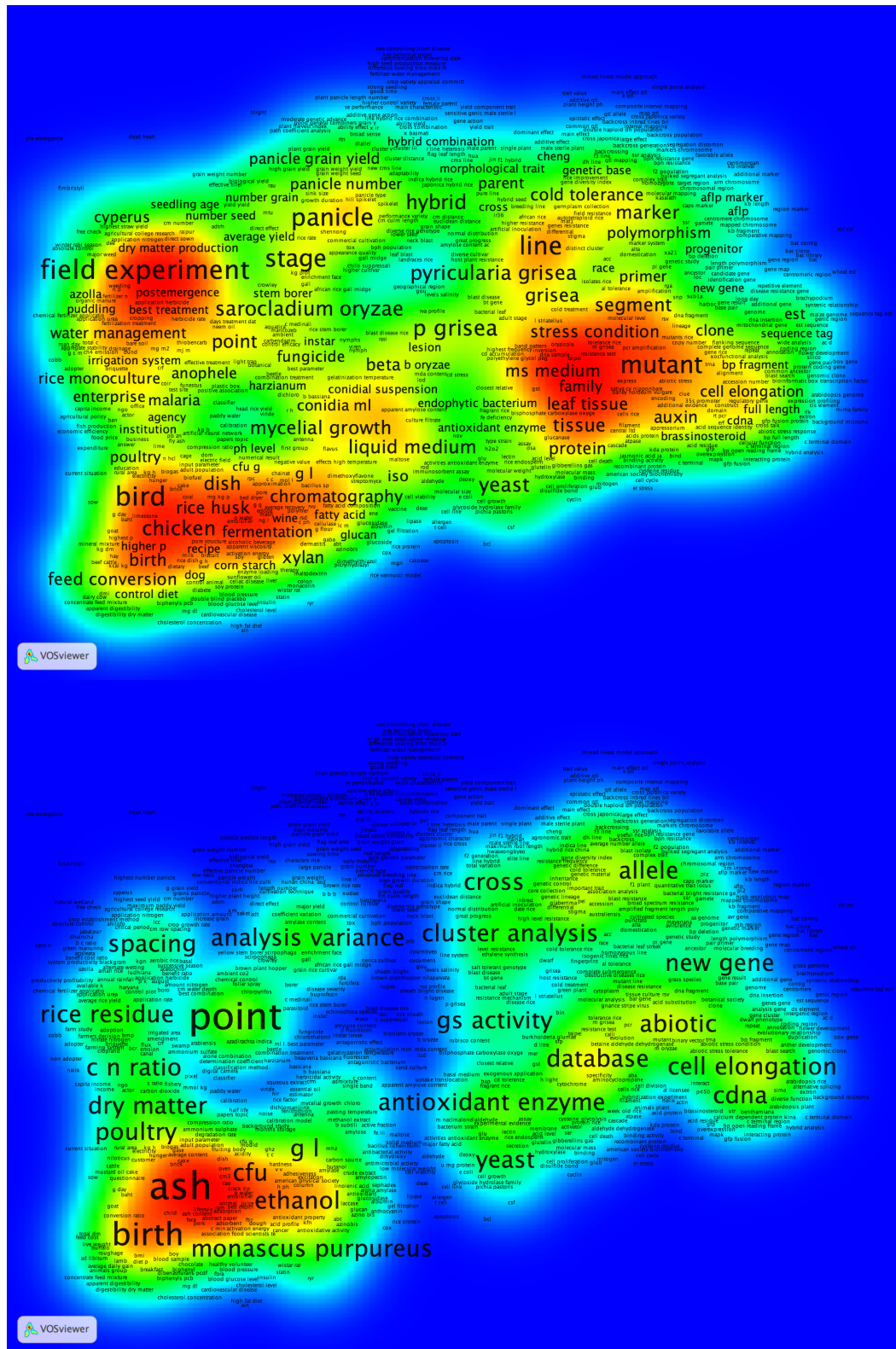
Whatever the reason, there is no doubt that the low coverage of many low and middle income countries in the mainstream databases result in a very distorted perspective of the research they carry out. This is illustrated in Figure 8, comparing the publication portfolio of rice for Iran according to CABI (top) and WoS (bottom). In the case of CABI, one observes expertise in rice production and consumption, with some work on genetics. In the case of WoS, only the consumption side is visible. Similarly Figure 9 compares the publication portfolio of rice for Cuba according to CABI (top) and WoS (bottom). This case is atypical but interesting because it shows Cuba publications in molecular biology are not covered in mainstream databases.

Table 3. Percentage of publications in each research topic by country. Ordered according to high percentage in productivity (second row), which is most directly associated with needs in global south.

Country	Consumpt. and human nutrition	Productiv. and plant nutrition	Molecul. biology	Genetics	Plant charact	Diseases, pests,
Bangladesh	17,98%	57,31%	7,08%	5,15%	7,23%	5,25%
Cote Ivoire	12,02%	54,68%	8,15%	9,15%	2,90%	13,10%
Ghana	37,76%	48,22%	4,30%	4,50%	3,66%	1,55%
Nigeria	35,48%	47,17%	4,85%	3,56%	4,57%	4,37%
India	20,16%	44,89%	10,40%	6,21%	8,89%	9,44%
Benin	7,12%	44,01%	9,00%	26,63%	8,61%	4,62%
Philippines	12,70%	41,22%	12,79%	21,90%	7,04%	4,35%
Pakistan	24,35%	40,48%	11,55%	7,78%	7,60%	8,24%
Uruguay	38,65%	37,22%	10,90%	5,26%	2,35%	5,63%
Australia	23,83%	33,46%	29,77%	7,66%	2,10%	3,17%
UK	27,04%	32,79%	24,16%	8,25%	1,81%	5,95%
Iran	35,27%	30,63%	10,27%	8,58%	8,85%	6,40%
Brazil	42,45%	30,62%	10,70%	6,08%	4,24%	5,90%
<b>All distrib.</b>	<b>28,31%</b>	<b>29,45%</b>	<b>21,14%</b>	<b>9,31%</b>	<b>6,55%</b>	<b>5,24%</b>
Colombia	39,76%	28,17%	14,68%	7,86%	4,07%	5,46%
Egypt	54,02%	26,12%	7,38%	2,87%	3,73%	5,88%
Germany	27,91%	24,85%	36,60%	6,75%	1,18%	2,71%
USA	30,37%	24,30%	30,13%	8,92%	2,15%	4,13%
Malaysia	54,45%	23,72%	9,94%	5,59%	2,89%	3,41%
China	19,46%	23,66%	26,72%	14,57%	11,46%	4,13%
Cuba	24,66%	23,63%	21,49%	8,89%	6,47%	14,85%
Argentina	39,40%	22,69%	20,31%	7,00%	2,46%	8,15%
Japan	27,32%	22,46%	32,42%	10,24%	3,85%	3,71%
France	26,23%	21,69%	36,52%	11,00%	1,26%	3,30%
Thailand	58,44%	21,41%	10,25%	4,39%	2,22%	3,30%
Spain	48,78%	20,23%	21,76%	3,93%	1,15%	4,15%
Taiwan	45,54%	19,37%	25,59%	4,89%	1,99%	2,62%
S. Korea	42,18%	12,51%	25,55%	8,98%	6,74%	4,03%



Figure 9. Publication density for rice research in Cuba for the database CABI (top) and WoS (bottom). CABI captures publications in various areas, whereas WoS portrait suggests a very patchy landscape with a focus on consumption.



## Discussion and conclusions

The results of this article suggest that previous assumptions on the stability of indicators across databases of scientific production are incorrect (Archambault et al., 2009). Instead, this case study on rice research shows that the indicator of number of publications per country is very dependent on the database when one analyses low and middle income countries, as suggested by much older studies (Arvanitis and Chatelin, 1988; Albornoz et al., 1999). A new important finding of this study is that there is also a serious bias on coverage of research topic. One may imagine that these differences are likely to apply as well to other fields of agricultural research, particularly those related to crops less relevant in the temperate climates of industrialised nations. These results are potentially important for international organisations such as FAO, IFRI or UNESCO (UNESCO, 2010) that aim to work on human development.

Nevertheless, the findings do not come as a surprise given the proliferation in the last two decades of journal indexing systems at the regional level, such as Scielo<sup>2</sup> or Redalyc<sup>3</sup> that aim to provide visibility to local journals, often in languages other than English (Chavarro, 2013), precisely to compensate for fact that the local science and its journals are not perceived as participating in “international” science (Velho, 1986).

There are a number of limitations in the empirical strategy adopted here. At this stage, we are not correcting the data for the fact that CABI only reports the affiliation of the first author. Second, although CABI coverage of publication is possibly the largest on a subject such as rice, publications still represent a subset of the research actually carried out. In research on agriculture many research outputs are not accounted for in publications, such as developments in the field, but also a lot of the research done by private companies and public organisations. Third, a bottom-up classification in terms of research topics is inevitably open to debate –although, for robustness, we have tested that the same findings are obtained with a different clustering methods.

Since knowledge representation can play a significant role in framing evaluation, research strategies, and technological development policies, we believe that topic bias in the dominant bibliometric databases deserve further conceptual and empirical discussion. The question whether and how such biases in the representation of knowledge may have an effect in research strategies is an important one. This study does not answer this question, but it shows that it is a meaningful issue. In particular, regarding the possibility that bibliometric reports of developing countries using dominant databases lead to inappropriate assumptions and choices regarding the domestic science bases and capabilities.

At the heart of this debate lies the question of how organisations or countries should manage research priorities so as to better align science “supply” with societal needs or “demands” (Pielke and Sarewitz, 2007; Kreimer and Zabala, 2007). Since bibliometric studies are part of the governance of science and innovation, the way they represent knowledge is bound to have an effect on the understanding and prioritisation of research portfolios (Wallace and Rafols, 2015).

---

<sup>2</sup>[www.scielo.org](http://www.scielo.org)

<sup>3</sup><http://www.redalyc.org>

Already in his classic discussion on research prioritisation Weinberg explains that when peer review breaks down –which is as soon as comparisons are needed between different subdisciplines— judgement about the relative worth of research is influenced by dominant norms in science that tend to value more universal over local findings, pure over applied, laboratory based over field work, etcetera (Weinberg, 1963). These normative views are likely to trickle down. The topic biases reported here may be both the result and further reinforce these biases with the ultimate effect of making science and technology less appropriate for the local needs of farmers in the global south. For these reasons, we believe that this apparently technical study is relevant for discussion on how the governance of science relates to inclusive innovation.

### Acknowledgments

We are grateful to Jim Sumberg and Andy Stirling for discussions. We acknowledge support from the EU (Marie Curie Integration fellowship to IR), the UK ESRC (RES-360-25-0076) and the US NSF (Award #1064146). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the funders.

### References

- Albornoz, M., Fernández Polcuch, E., & Arber, G. 1999. *Principales Indicadores de Ciencia y Tecnología*. Buenos Aires: RICYT.
- AOSTI. 2014. *Assessment of Scientific Production in the African Union, 2005–2010*. Malabo: AOSTI (African Observatory of Science, Technology and Innovation).
- Archambault, É., Campbell, D., Gingras, Y., Larivière, V., 2009. Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology* 60, 1320–1326.
- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., Gingras, Y., 2006. Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics* 68, 329–342.
- Arvanitis, R., & Chatelin, Y. 1988. National Scientific Strategies in Tropical Soil Sciences. *Social Studies of Science*, 18, 113–146.
- Chavarro, D., 2013. ¿ Son los sistemas de indexación y resumen un indicador de la buena calidad editorial de las revistas académicas? Presented at the RICyT Conference, Bogotá. <http://congreso2013.ricyt.org/files/mesas/1dProduccioncientifica/Chavarro.pdf>. Accessed on 1st July 2014.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. 2015. [The Leiden Manifesto for Research Metrics](#). *Nature*, 520, 429–431.
- Kawasaki, J.L., 2004. Agriculture Journal Literature Indexed in Life Sciences Databases. Issues in Science and Technology Librarianship Summer 2004. Available at: <http://www.istl.org/04-summer/article4.html>. Accessed 1st July 2014.
- Kreimer, P., Zabala, J.P., 2007. Chagas Disease in Argentina: Reciprocal Construction of Social and Scientific Problems. *Science, Technology and Society* 12, 49–72.
- Leydesdorff, L., 2012. World shares of publications of the USA, EU-27, and China compared and predicted using the new Web of Science interface versus Scopus”. *El profesional de la información*, 21(1), 43-49.
- McDonald, P., 1994. Characteristics of soil science literature, in: *The Literature of Soil Science*. Cornell University Press, pp. 43–73.



- Martin, B. R., Tang, P., Morgan, M., & al. 2010. *Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project* (A report for DFG, ESRC, AHRC, NWO, ANR and ESF). Brighton, UK: SPRU.
- Rafols, I., Leydesdorff, L., O’Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinarity. The case of innovation studies and business and management. *Research Policy*, 41(7), 1262–1282.
- Royal Society, 2011. Knowledge, networks and nations: Global scientific collaboration in the 21st century. The Royal Society, London.
- Sarewitz, D., & Pielke, R. a. (2007). The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, 10(1), 5–16.
- UNESCO, 2010. UNESCO science report 2010. The current status of science around the world. UNESCO, Paris.
- Van Eck, N. J., Waltman, L., van Raan, A. F. J., Klautz, R. J. M., & Peul, W. C. (2013). Citation Analysis May Severely Underestimate the Impact of Clinical Research as Compared to Basic Research. *PLoS ONE*, 8(4), e62395. <http://doi.org/10.1371/journal.pone.0062395>
- Van Leeuwen, T., Moed, H., Tijssen, R. W., Visser, M., & Van Raan, A. J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51(1), 335–346. <http://doi.org/10.1023/A:1010549719484>
- Velho, L., & Krige, J. 1984. Publication and citation practices of Brazilian agricultural scientists. *Social Studies of Science*, 14(1), 45-62.
- Velho, L., 1986. The “meaning” of citation in the context of a scientifically peripheral country. *Scientometrics* 9, 71–89.
- Velho, L., 1990. Sources of Influence on Problem Choice in Brazilian University Agricultural Science. *Social Studies of Science* 20, 503–517.
- Wallace, M.L., Rafols, I., 2015. [Research portfolios in science policy: moving from financial returns to societal benefits](#). *Minerva* (Online). doi: 10.1007/s11024-015-9271-8
- Weinberg, Alvin M. 1963. The Criteria for Scientific Choice. *Minerva* 1: 159–171.