Technical description of exploratory study on agriculture journals coverage by various bibliographic databases

Associated with the Leiden Madtrics Blog: ***Why coverage matters: Invisibility of agricultural research from the Global South may be an obstacle to development***

19 July 2023

Ismael Rafols[1], Muthu Madhan[2,3], Josep-Manuel Rodríguez Gairín[4], Marta Somoza Fernández[4] and Cristóbal Urbano[4]

[1] Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands. i.rafols@cwts.leidenuniv.nl
[2] DST-CPR, Indian Institute of Science, Bangalore, India. mu.madhan@gmail.com
[3] O P Jindal Global University, Sonipat, India.
[4] Department of Library and Information Sciences, Universitat de Barcelona, Barcelona, Spain. rodriguez.gairin@ub.edu, msomoza@ub.edu, urbano@ub.edu

Motivation

An agricultural development program usually consists of two vital elements: core agricultural research (i.e. on seeds, or on machinery) and extension (which may also include some research but whose focus is on implementation and adoption). These are complementary activities. Extension plays an important role in the success of agricultural development programs, especially in developing countries. Hence, agricultural researchers have an immediate responsibility to ensure that their research reaches the primary beneficiaries, namely farmers.

Agricultural researchers, like in others fields, use journals to share their ideas and findings. Some of these journals focus predominantly on advancements in core agricultural research, and some concentrate on extension research and cater to the needs of extension personnel in general. Extension journals cover topics relevant to specific locations (local, regional, national), offering insights that can influence policy decisions at various levels.

The commonly used citation databases often exclude journals from developing countries, especially those having focus on extension, due to their failure to meet the selection criteria, particularly related to citation-based parameters. Even the subject-specific databases (e.g. CAB Abstracts) which give special consideration to journals published from the developing countries, leave out many journals as these are unable to meet the required professional standards. This absence is a barrier for discovering information and conducting bibliometric analyses for monitoring or evaluation purposes. The objective of this brief study is to perform an exploratory analysis of how well agricultural journals are covered by the most widely used bibliographic databases.

Methods and data

Summary:

We downloaded the journals in the field of agriculture (excluding Forestry, Fisheries or Veterinary Science) as defined in DOAJ and Ulrichs on 14th March 2023, then searched these journals in the MIAR database (https://miar.ub.edu/), from which we obtained a sample of 1,662 with information

regarding the journal country and characteristics, which allowed us to compare bibliographic coverage across a battery of characteristics.

Detailed description:

To assess and evaluate the extent of coverage of agricultural journals indexed by the citation and subject databases, we created a sample set by selecting journals listed in the Directories (DOAJ and Ulrich) and then checked if they are indexed by the databases.

First, we examined the journals listed in DOAJ and considered a total of 588 journals, which were classified under five categories: Agriculture; Agriculture (General); Agriculture: Plant culture; Agriculture: Animal culture; and Agriculture: Cattle.

Second, we retrieved from Ulrichs 2,020 journals obtained from searching the field 'Subject' with the keywords "Agriculture" OR "Agricultural" OR "Horticulture"; Status="Active"; Serial type="Journal"; Format="Online". Of these only 1448 journals had ISSN.

The two lists were consolidated into one list of 2,232 journals of which 1,662 journals had ISSN. We used this list of 1,662 journals as the sample for our study. Both the initial list and the sample list are available in the complementary excel file in SocArxiv.

We used the platform MIAR (https://miar.ub.edu/), developed by the Faculty of Information and Audiovisual Media, University of Barcelona), which provides an information matrix for journal analysis, to obtain more detailed information on the 1,662 journals, particularly the journals' presence in multidisciplinary citation databases like Web of Science and Scopus, and in subject specific databases such as CAB Abstracts, Veterinary Sciences, Biosis, Agricultural & Environmental Science, and Food Science and Technology Abstracts.

Additionally, we conducted a separate estimate of the journals indexed in Dimensions using CWTS internal database. For this process, we relied on a single ISSN number provided with the source to compare the records. However, we did not utilise the second ISSN (L) in *Dimensions*, which resulted in some missing records (estimated at 12%). Therefore, coverage in Dimensions is an estimation.

Results

We compared the sample set of 1,662 journals with the journals indexed by a total of 8 databases. Among these, three are multidisciplinary citation databases, while the remaining five are subject-specific databases related to agriculture and allied areas. In the sample set, 1,332 journals were covered in at least one of the eight databases, and 300 were not covered by any (part of them might be journals created in recent years and not yet indexed).

The five databases with the higher coverage are Dimensions (1160 journals), CAB Abstracts (805), Scopus (562), Veterinary Science (550) and Web of Science (WoS) (435). Other relevant databases included BIOSIS (262 journals), Agricultural & Environmental Science (260) and Food Science and Technology Abstracts (251).

Dimensions has included a larger proportion of journals (69.8%) from the sample list compared to other databases we studied. This is because publications data in Dimensions is indexed at the article level, allowing it to encompass a more comprehensive range of sources. Dimensions entrusts users with the responsibility of evaluating the quality of papers and their sources, enabling them to filter and customise their searches according to their requirements. This means that some of the journals included have various degrees of poor or problematic editorial or review practices (the term

'predatory' is unhelpful here for analytical purposes). In contrast, the other databases have journal selection criteria and index papers only from a selected list of journals. Most of the publication data in Dimensions is sourced directly from open sources, which means that articles published in journals listed in DOAJ may have been comprehensively covered.

There are overlaps in the coverage of journals between databases, and each database contains unique journals not found in the others. When comparing Dimensions and CAB Abstracts, which offer better coverage compared to other databases, we find that out of the 1662 journals in the sample set, 673 journals are covered by both Dimensions and CAB Abstracts. Additionally, 487 journals are indexed only in Dimensions, while 132 journals are indexed only in CAB Abstracts.
And 370 journals are not covered by any database. Of these, 18 journals are indexed only in either Web of Science or Scopus. (Figure 1)

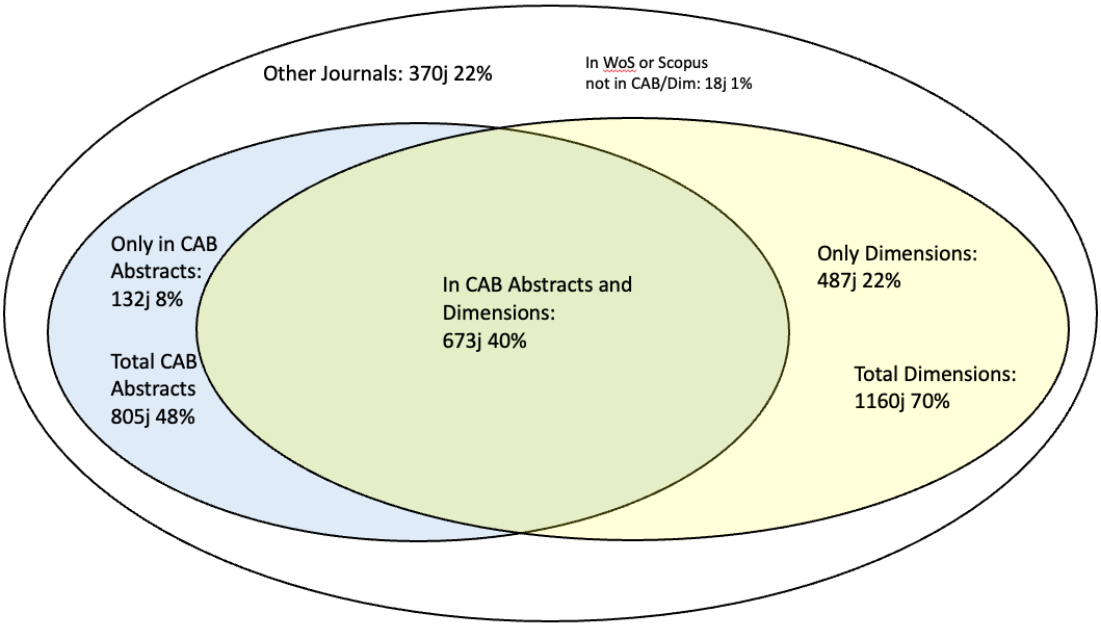

Figure 1. Coverage of agriculture-related journals by Dimensions and CAB Abstracts.

When examining the journals indexed in Web of Science, Scopus, and the CAB Abstracts, three well-established databases, we found that 356 journals are covered by all three databases. Moreover, 125 journals are covered by both Scopus and CAB Abstracts, 29 journals are covered by both Web of Science and CAB Abstracts, and 35 journals are covered by both Web of Science and Scopus. Furthermore, there are 295 journals listed exclusively in CAB Abstracts, 46 journals exclusive to Scopus, and 15 journals listed solely in Web of Science. (Figure 2)

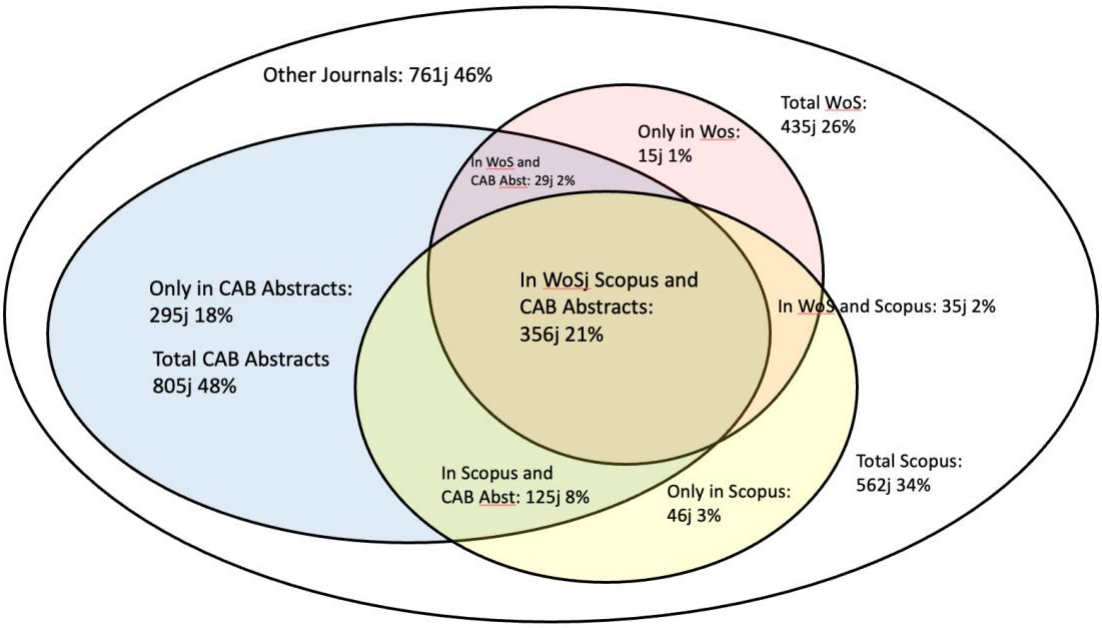

Figure 2. Coverage of agriculture-related journals by Web of Science, Scopus and CAB Abstracts.

After arranging the journals in the sample set according to their publication countries and grouping them into seven regions (Table 1), we noted the following distribution: 579 journals are published in the Europe & Central Asia region, 239 in North America, 228 in East Asia & Pacific, 211 in Latin America and the Caribbean, 128 in the Middle East and North Africa, and 113 in Sub-Saharan Africa.

We found that Dimensions stands out as the only database offering a significantly broader and more comprehensive coverage of journals. It includes 80% of the journals published from Latin America & the Caribbean, 79% from East Asia & the Pacific, 75% from Europe & Central Asia, and 67% of the journals published from North America and South Asia. Although Dimensions covers less than 50% of the journals published from the Middle East & North Africa (47%) and Sub-Saharan Africa (41%), its coverage is still comparatively higher than the other databases studied here. CAB Abstracts ranks as the second-largest database in terms of the number and coverage, representing 48.4% of all journals in the sample, and it includes journals from various regions. More than 60% of the journals published from Latin America & the Caribbean and Europe and Central Asia are included in CAB Abstracts. Web of Science and Scopus have a better coverage of journals published from Europe and Central Asia compared to other regions. Their coverage of journals from North America is the second highest, while Latin America follows closely as the third.

An examination of how these databases cover journals published from different countries reveals that journals published from the Netherlands are better covered in all the databases. Dimensions has indexed over 90% of the journals published from six countries (Colombia, Indonesia, Netherlands, Ukraine, Brazil and Poland) and at least 50% of the journals published from 25 countries (Table 2). The coverage of Veterinary Science is relatively higher in low and middle income countries.

An analysis by countries (see Table 2) shows a more complex pattern, with significant differences in coverage within a world region (e.g. between the Netherlands, Poland and Romania).

The Venn diagram between CAB Abstracts (Figure 1), Scopus and WoS shows that while CAB Abstracts provides a much larger coverage than WoS and Scopus, among the three of them, they still miss almost half the journals (46%). The Venn diagram between CAB Abstracts and Dimensions (Figure 2) shows

that while Dimensions is much larger (70% vs. 48%), CAB Abstracts adds 8% of additional journals. And still there are 370 journals not yet covered.

Table 1. Number of journals covered by database in volume (#) and percentage (%) in total and by world region.

		WoS		**Scopus**		**Dimensions**		**CAB Abst.**		**Vet. Sci.**	
	Total	**#**	***%***	**#**	***%***	**#**	***%***	**#**	***%***	**#**	***%***
Total (baseline)	**1662**	**435**	***26%***	**562**	***34%***	**1160**	***70%***	**805**	***48%***	**550**	***33%***
Europe & Central Asia	579	227	*39%*	288	*50%*	434	*75%*	352	*61%*	238	*41%*
North America	239	78	*33%*	89	*37%*	161	*67%*	93	*39%*	75	*31%*
East Asia & Pacific	228	39	*17%*	55	*24%*	180	*79%*	67	*29%*	30	*13%*
Latin Am. & Caribbean	211	63	*30%*	68	*32%*	169	*80%*	136	*64%*	102	*48%*
South Asia	164	15	*9%*	36	*22%*	110	*67%*	70	*43%*	51	*31%*
Mid. East & N. Africa	128	9	*7%*	20	*16%*	60	*47%*	59	*46%*	35	*27%*
Sub-Saharan Africa	113	4	*4%*	6	*5%*	46	*41%*	28	*25%*	19	*17%*

Table 2. Number of journals covered by database in volume (#) and percentage (%) in total and by top countries.

		WoS		**Scopus**		**Dimensions**		**CAB Abstr.**		**Vet. Sci.**	
	Total	#	%	#	%	#	%	#	%	#	%
Total	**1662**	**435**	**26%**	**562**	**34%**	**1160**	**70%**	**805**	**48%**	**550**	**33%**
United States	227	75	33%	87	38%	155	68%	88	39%	71	31%
UK	133	87	65%	93	70%	110	83%	87	65%	71	53%
India	122	11	9%	26	21%	80	66%	55	45%	42	34%
Indonesia	112	6	5%	9	8%	107	96%	23	21%	5	4%
Brazil	101	35	35%	41	41%	94	93%	73	72%	55	54%
Nigeria	80	1	1%	1	1%	29	36%	16	20%	15	19%
Iran	71	5	7%	12	17%	30	42%	42	59%	25	35%
Netherlands	46	38	83%	44	96%	44	96%	38	83%	36	78%
Germany	42	24	57%	32	76%	26	62%	30	71%	23	55%
Poland	39	12	31%	16	41%	35	90%	26	67%	13	33%
Russian Fed	38	1	3%	5	13%	26	68%	11	29%	2	5%
Switzerland	36	11	31%	16	44%	28	78%	12	33%	3	8%
Romania	31	9	29%	7	23%	16	52%	24	77%	12	39%
Turkey	27	3	11%	6	22%	20	74%	20	74%	15	56%
Pakistan	25	3	12%	9	36%	16	64%	7	28%	4	16%
Colombia	23	10	43%	9	39%	23	100%	17	74%	15	65%
Italy	23	7	30%	12	52%	15	65%	13	57%	7	30%
Ukraine	23	2	9%	2	9%	22	96%	3	13%	1	4%
Australia	22	10	45%	10	45%	9	41%	9	41%	7	32%
South Korea	22	5	23%	11	50%	19	86%	10	45%	3	14%
France	21	7	33%	9	43%	11	52%	12	57%	6	29%
Japan	21	5	24%	6	29%	15	71%	12	57%	6	29%
China	20	7	35%	12	60%	12	60%	8	40%	4	20%
Argentina	18	4	22%	7	39%	11	61%	14	78%	10	56%
UAE	18	1	6%	3	17%	4	22%	1	6%	1	6%
Czech Republic	16	8	50%	12	75%	12	75%	13	81%	11	69%
Bangladesh	15	0	0%	0	0%	12	80%	6	40%	3	20%

Egypt	15	1	7%	1	7%	12	80%	6	40%	2	13%
Spain	15	5	33%	7	47%	10	67%	8	53%	4	27%
Cuba	14	2	14%	0	0%	0	0%	7	50%	4	29%

First concluding insights to follow up in future research

This exploratory analysis shows that the coverage of agricultural research by 'traditional' bibliometric databases such as WoS, Scopus or CAB Abstracts is limited. Even coverage by the new more comprehensive databases such *Dimensions* also has gaps, but on the other hand, it may include journals with poor or problematic editorial practices.

Discussion

The preliminary analysis clearly indicates that popular bibliometric databases such as WoS, Scopus, and CAB Abstracts have limited coverage of agricultural research. Some agricultural journals published in developing countries may not meet the selection criteria set by these databases, resulting in exclusion from their indexes. As a consequence, research from these countries, particularly in the field of extension, is likely to be overlooked in these databases' coverage.

While CAB Abstracts offers better coverage compared to the other two, it may not be sufficient for a comprehensive literature search. CAB Abstracts, being a subject-specific database focusing on agriculture and public health, and the interface between the two, has a limitation. It only concentrates on the literature within this specialty and overlooks other relevant literature that could be valuable to researchers in the field. Moreover, the database lacks the capability to link citations beyond the literature within its specialty. As early as 1975, Garfield analysed the references in papers published in core agricultural journals and pointed out, "agricultural scientists use and cite the same hard core of frequently cited basic research journals used by all other research workers in the life sciences." Agricultural researchers would find using a multidisciplinary citation database with broad coverage more advantageous for their literature searches compared to relying on subject-specific databases like CAB Abstracts.

The rise of new multidisciplinary citation databases like OpenAlex, Dimensions and Lens.org, which include data from open information sources and open metadata repositories, and do not base their content recruitment policies on journal performance, brings hope for improved coverage, as evident from this analysis.

While this approach shows promise, it has some problems, as it could potentially include papers published in journals with varying levels of editorial and academic rigour. However, these issues might be mitigated as databases continue to refine their algorithms and with the help of catalogues of editorial practices such as DOAJ or Latindex and platforms such as MIAR. Moreover, for users, conducting searches and filtering results from comprehensive databases is easier than navigating through dozens of interfaces.

National Agricultural Research Systems (NARS) should take the role of effective data providers by developing interoperable directories of important journals that help facilitate research communication at the national levels and set up open access repositories of papers written by their scientists. Also NARS should encourage the agricultural research community to embrace emerging models of research communication, such as preprint servers, peer-reviewed preprints, and overlay journals. All these will help database service providers improve and maintain content quality.

Overall, results of this analysis show that the benefits the databases with comprehensive coverage offer would far outweigh the problems related to the notions of quality of contents. This approach has the potential to overcome limitations observed in traditional bibliometric databases, resulting in a more comprehensive and inclusive coverage of agricultural research, particularly from developing countries.

Conducting a comprehensive analysis of agricultural journals across various databases and compiling a list of journals from all regions of the world would be highly beneficial, especially in the absence of such a comprehensive list. Additionally, conducting detailed analyses of journal coverage at the regional and country levels can help identify gaps in data provision and service offerings by various agencies.