Methods employed in the *Web of Science* and *Scopus* databases to effect changes in the ranking of the journals

D. Gnana Bharathi

Databases of Web of Science and Scopus show significant variations in the number of articles published by some journals. The number of articles published in a given journal may not be reflected as such in terms of numbers in both the databases. Such variations are often found in those journals that publish articles such as editorials, news, letter, etc. in addition to the research articles and reviews. In certain journals, the databases selectively categorize some of the articles as citable articles which the journals themselves specify as editorial, news, policy forum and book review. These variations lead to inappropriate evaluation of journals. This article highlights the occurrence of such discrepancies and variations in the databases and calls for accuracy and efficiency in compilation of the data.

Keywords: Citation, evaluation, impact factor, *Scimago Journal Rank*, *Scopus*, *Web of Science*.

Thomson Reuters' *Web of Science* (*WoS*), Elsevier's *Scopus* and *Google's Scholar* are the major online databases that serve as platforms to evaluate the impact of the scientific journals. *WoS* and *Scopus* are fee-based services, while *Google Scholar* is an open access resource. The number of articles and citation counts in *WoS* and *Scopus* are used to assess the impact of scientific journals as Impact Factor (IF) and *Scimago Journal Rank* (SJR), respectively. Evaluation of journals, in general, is made based on two parameters, viz. citations from different journals for the articles published in a journal, as the numerator and number of articles published from the corresponding journal, as the denominator.

Articles that receive more citations are considered to have high impact in the scientometrics studies. Along with peer review, it has over the last three decades been increasingly used to judge and quantify the importance of scientists and scientific research[1]. It is used to estimate the quality, impact, originality, penetration or visibility of individual and corporate performance within and across disciplines[2]. Similarly, identification and classification of articles play a crucial role in the evaluation of journals. Journals publish articles and classify them under different categories. The databases compile different types of journals into selected categories of their own for standardization of the database. This may lead to misrepresentation of articles in different categories.

Variations in the databases are widely studied by several authors[3–5]. These studies have mainly concentrated on the field of specialization[6,7]. In-depth comparison of the databases is carried out on specific usages[8–10].

D. Gnana Bharathi is in Central Leather Research Institute, Chennai 600 020, India. e-mail: dgbharathi@yahoo.com

Indexing of journals by databases

Journals are indexed based on various criteria by different databases. *WoS* and *Scopus* follow well-defined selection criteria. Thomson Reuters' selection criteria is based on the Bradford law and requires minimum years of publication, basic publishing standards, its editorial content, the international diversity of its authorship and the citation data associated with it are all considered[11]. Similarly, *Scopus* delineation of journals is based on 'peer-reviewed content, regular publication, relevancy, publication ethics, etc.'[12]. *WoS* includes around 12,000 journals and *Scopus* about 18,850 journals in their database. Citations received by these journals are measured for calculation of IF and SJR.

The number of articles published by a journal varies from year to year. However, the databases follow different methodologies to measure the content of the journals[13,14]. While *WoS* indexes cover-to-cover of the journals, *Scopus* excludes book reviews and conference meeting abstracts of the journals in its coverage of total articles. Citable articles in IF include 'article', 'review' and 'proceedings paper' and SJR includes 'article', 'review' and 'conference paper'[15,16]. Citable articles of IF and SJR are chosen by specialists of the respective databases based on their critical evaluation of articles of the journals. However, there are significant variations among the databases in selecting the articles.

Effect of selection and classification of articles

There is a vast difference in the selection and classification of the journals by the databases. Detailed studies on the databases have been carried out by a number of authors[17,18]. However, influence of the citations and articles by the databases is less studied. Authors and journals benefit only by increasing the number of citations. Databases, on the other hand, can influence the journals by means of increasing or decreasing number of citations and articles.

Experts in the databases decide what is 'citable articles'. Inclusion of every article decreases the IF and SJR by means of increasing the denominator in the calculation. Exclusion of the articles as citable articles can increase or maintain the IF and SJR of the journals by not increasing the denominator. If these articles get citations, IF and SJR of the journals will increase.

Any biased addition or omission of articles by the databases can influence the ranking of the journals. As a result, a new evaluation method which avoids classification of articles, but only the cited articles, proven by at least one citation, was proposed[19].

Methodology

Journals for the study are selected representing the entire range of IF. Selection of journals is based on selective and random methods. Selective journals include top two and bottom two journals according to the IF ranking in 2011. Journals such as *Nature* and *Science* are also selectively chosen. Journals such as *Curr Sci* and *Indian J Med Res* are also included, as are non-English journals *Veliger* and *Electrochemistry*. Other journals are randomly selected representing the different ranges of IF-2011 ranking.

Information about the total articles and citable articles is collected from *WoS* and *Scopus* databases. Total articles include all the content of the journal representing all categories. Citable articles include article, review and proceedings or conference paper. Only those journals where significant or no variations exist between the databases are analysed in detail.

According to the *Scopus* database, *CA A Cancer Clin* has published 9, 1, 0, 7 and 0 articles in the years 2007, 2008, 2009, 2010 and 2011 respectively. This appeared to be incorrect. Therefore, the information on the number of total articles is obtained from *Scimago* website, which ranks the SJR of the journals. Citable articles are derived using the data table of the journal in the website.

Assuming that citations reported by both the databases are equal, how the variation in selection of articles by the databases effect changes in IF ranking of the journals for the years 2001, 2006 and 2011 is given in Table 1. For example, estimated citations for the journals (2010) are calculated by multiplying the IF (of 2010) with citable articles (of 2008 and 2009).

Recalculation for SJR is not carried out due to the complexities in calculation of SJR in which each citation can differ from another based on its prestige, similar to Google's Page Rank. Further, SJR is calculated based on 3 years of publications.

Variation in the criteria of the selection of articles

There are variations between *WoS* and *Scopus* on the total number of articles published from the same journals. The difference also exists in citable articles in the databases. Figure 1 shows the number of articles published by six journals, viz. *Lancet*, *Nature*, *Science*, *New Eng Med*, *Current Science* and *Cell* from 1996 to 2009.

Figure 1(Ia) shows the total number of articles published by *Lancet* between 1996 and 2009 as recorded in *WoS* and *Scopus*. In the 1990s, *Scopus* counts show close to 4000 articles published by *Lancet* in comparison to less than 3000 articles by *WoS*. In the last decade, *WoS* and *Scopus* counts are more or less equal to the number of *Lancet* articles published. Similarly, the difference in the total number of articles between the two databases is given for journals, viz. *Nature* (IIa), *Science* (IIIa), *New Eng Med* (IVa), *Curr Sci* (Va) and *Cell* (VIa).

Counting number of articles published by a journal, as stated in the selection criteria of the databases, should result in comparable number of articles for a given journal in both the databases. However, there are significant differences in the databases. These differences in calculation of the number of articles clearly demonstrate the erroneous counting of the total number of articles published by a given journal by *WoS* and *Scopus*. The total number of articles of a journal does not play any role in the calculation of IF, but is included in the SJR calculation.

Figure 1(Ib) shows the number of 'citable articles' published every year by *Lancet* according to *WoS* and *Scopus* databases. The difference is to the extent of 300–400 articles, implying preferential selection by one or both the databases. Similarly, the difference between the two databases is given for *Nature* (IIIb), *Science* (IIIb), *New Eng Med* (IVb), *Curr Sci* (Vb) and *Cell* (VIb). A neutral and proper evaluation of articles should give more or less equal number of 'citable articles' among the databases. These differences assume significance as the number of 'citable articles' is used as a denominator for IF and SJR calculation.

Discrimination through classification

Each journal has its own type of classification of the content it publishes. As the type of articles differs from one journal to another, it is necessary for the databases to have common type of classifications to cover all the

Table 1. Estimated citations for the journals are calculated by multiplying the number of citable articles in the *WoS* with the IF to get estimated citations. The estimated citations are divided by the citable articles as per the *Scopus* database to obtain IF^s of *Scopus* database. IF^s value is compared with the IF value to see the variation of error in percentage

		Calculation for 2001					
			Citable articles (1999 + 2000)		Est. citations	IF 2001 on *Scopus* data	
Journal	ISSN No.	IF 2001	*WoS*	*Scopus*	in 2001	IF^s	Diff.
CA A Cancer Clin	0007-9235	35.933	30	46	1,078	23.435	–34.8%
New Engl Med	0028-4793	29.065	759	1,159	22,060	19.034	–34.5%
Chem Rev	0009-2665	21.044	252	250	5,303	21.212	0.8%
Lancet	0140-6736	13.251	1,929	2,349	25,561	10.882	–17.9%
Adv Phys	0001-8732	16.200	20	20	324	16.200	0.0%
Nature	0028-0836	27.955	2,331	2,569	65,163	25.365	–9.3%
Cell	0092-8674	29.219	692	560	20,220	36.106	23.6%
Science	0036-8075	23.329	1,891	2,333	44,115	18.909	–18.9%
Cancer Cell	1535-6108	–	–	–	–	–	–
Circulation	0009-7322	10.517	2,205	2,205	23,190	10.517	0.0%
Trends Biochem Sci	0968-0004	14.329	240	199	3,439	17.281	20.6%
Diabetes Care	0149-5992	5.404	648	641	3,502	5.463	1.1%
Faraday Discuss	1364-5498	3.261	134	158	437	2.766	–15.2%
Earth Planet Sci Lett	0012-821X	2.700	616	624	1,663	2.665	–1.3%
Tetrahedron	0040-4020	2.276	2,238	2,230	5,094	2.284	0.4%
Neural Networks	0893-6080	1.431	188	181	269	1.486	3.8%
Indian J Med Res	0971-5916	0.340	150	150	51	0.340	0.0%
Electrochemistry	1344-3542	0.606	398	290	241	0.832	37.3%
Curr Sci	0011-3891	0.600	910	779	546	0.701	16.8%
Ophthalmologe	0941-293X	0.552	268	287	148	0.515	–6.7%
Mech Eng	0025-6501	0.109	202	0	22	–	–
Metallurgist+	0026-0894	0.008	243	93	2	0.021	162.5%
Plast Eng	0091-9578	0.136	81	68	11	0.162	19.1%
J Jpn Soc Tribologis	0915-1168	0.038	320	140	12	0.087	128.9%
Nav Archit	0306-0209	0.015	457	208	7	0.033	120.0%
		Calculation for 2006					
			Citable articles (2004 + 2005)		Est. citations	IF 2006 on *Scopus* data	
Journal	ISSN No.	IF 2006	*WoS*	*Scopus*	in 2006	IF^s	Diff.
CA A Cancer Clin	0007-9235	63.342	38	43	2,407	55.977	–11.6%
New Engl Med	0028-4793	51.296	624	1,414	32,009	22.637	–55.9%
Chem Rev	0009-2665	26.054	315	316	8,207	25.972	–0.3%
Lancet	0140-6736	25.800	912	1,496	23,530	15.728	–39.0%
Adv Phys	0001-8732	9.389	19	19	178	9.389	0.0%
Nature	0028-0836	26.691	1,979	2,766	52,821	19.097	–28.5%
Cell	0092-8674	29.194	607	594	17,721	29.833	2.2%
Science	0036-8075	30.028	1,857	2,877	55,762	19.382	–35.5%
Cancer Cell	1535-6108	24.077	201	157	4,839	30.825	28.0%
Circulation	0009-7322	10.940	2,120	2,416	23,193	9.600	–12.2%
Trends Biochem Sci	0968-0004	13.863	172	200	2,384	11.922	–14.0%
Diabetes Care	0149-5992	7.912	956	1,047	7,564	7.224	–8.7%
Faraday Discuss	1364-5498	4.731	134	164	634	3.866	–18.3%
Earth Planet Sci Lett	0012-821X	3.887	893	873	3,471	3.976	2.3%
Tetrahedron	0040-4020	2.817	2,425	2,429	6,831	2.812	–0.2%
Neural Networks	0893-6080	2.000	221	223	442	1.982	–0.9%
Indian J Med Res	0971-5916	1.224	277	230	339	1.474	20.4%
Electrochemistry	1344-3542	0.574	270	266	155	0.583	1.6%
Curr Sci	0011-3891	0.737	1,013	1,100	747	0.679	–7.9%
Ophthalmologe	0941-293X	0.762	311	309	237	0.767	0.7%
Mech Eng	0025-6501	0.107	149	96	16	0.166	55.1%
Metallurgist+	0026-0894	0.062	208	196	13	0.066	6.5%
Plast Eng	0091-9578	0.101	69	63	7	0.111	9.9%
J Jpn Soc Tribologis	0915-1168	0.012	257	216	3	0.014	16.7%
Nav Archit	0306-0209	0.001	696	196	1	0.004	300.0%

(*Contd*)

Table 1. (*Contd*)

Journal	ISSN No.	Calculation for 2011					
			Citable articles (2009 + 2010)			IF 2011 on *Scopus* data	
		IF 2011	*WoS*	*Scopus*	Est. citations in 2011	IF^s	Diff.
CA A Cancer Clin	0007-9235	101.78	41	44	4,173	94.840	–6.8%
New Engl Med	0028-4793	53.298	697	1,230	37,149	30.202	–43.3%
Chem Rev	0009-2665	40.197	373	369	14,993	40.633	1.1%
Lancet	0140-6736	38.278	551	1,304	21,091	16.174	–57.7%
Adv Phys	0001-8732	37.000	14	15	518	34.533	–6.7%
Nature	0028-0836	36.280	1,728	1,931	62,692	32.466	–10.5%
Cell	0092-8674	32.403	679	675	22,002	32.595	0.6%
Science	0036-8075	31.201	1,759	1,860	54,883	29.507	–5.4%
Cancer Cell	1535-6108	26.556	175	172	4,647	27.019	1.7%
Circulation	0009-7322	14.739	1,150	1,416	16,950	11.970	–18.8%
Trends Biochem Sci	0968-0004	10.847	163	152	1,768	11.632	7.2%
Diabetes Care	0149-5992	8.087	1,001	1,029	8,095	7.867	–2.7%
Faraday Discuss	1364-5498	5.000	162	146	810	5.548	11.0%
Earth Planet Sci Lett	0012-821X	4.180	1,073	1,083	4,485	4.141	–0.9%
Tetrahedron	0040-4020	3.025	2,466	2,470	7,460	3.020	–0.2%
Neural Networks	0893-6080	2.182	269	268	587	2.190	0.4%
Indian J Med Res	0971-5916	1.837	344	357	632	1.770	–3.6%
Electrochemistry	1344-3542	0.954	218	218	208	0.954	0.0%
Curr Sci	0011-3891	0.935	733	703	685	0.975	4.3%
Ophthalmologe	0941-293X	0.617	298	316	184	0.582	–5.7%
Mech Eng	0025-6501	0.209	153	176	32	0.182	–12.9%
Metallurgist+	0026-0894	0.110	210	212	23	0.109	–0.9%
Plast Eng	0091-9578	0.050	60	62	3	0.048	–4.0%
J Jpn Soc Tribologis	0915-1168	0.009	221	233	2	0.009	0.0%
Nav Archit	0306-0209	0.005	364	116	2	0.016	220.0%

journals in specific categories. This enables databases to categorize and group different types of articles into editorial, article, letter and review, news, notes, to name a few. *WoS* classifies articles of the journals mainly as editorial material, article, news item, letter, review, biography, correction, book review, note and conference paper. *Scopus* classifies articles of the journals mainly as editorial, article, letter, note, short survey, conference paper, erratum, review, report, book and undefined.

Many journals publish articles as editorial, news, views and forum, which the journals do not treat as original research output. Yet, some of these articles are classified as 'citable articles' in *WoS* and *Scopus*. This mismatching of articles by the databases is found in journals such as *Nature* and *Science*, and journals with low impact factor such as *Inter Dis Sci*, *Curr Sci* and *Sci Am*.

Some articles in *Nature* that are classified as Concepts, News Feature, Careers and Recruitment, Product Review, News and Views, and Feature are grouped as article in *WoS*. Similarly, some articles in *Science* classified as News Focus, Essays on Science and Society, Tech Sight, News, Policy Forum, Association Affairs and Editorial are categorized as articles in *WoS*. Categorization of a manuscript published in a journal as a citable article is arbitrary in *WoS* and does not follow any pattern. Table 2 shows few examples of how various types of articles such as Editorial, Commentary and Career are treated as citable articles.

Table 3 shows how citable articles in *Nature* are derived from the total number of articles by *WoS* and *Scopus*. In *Nature*, citable articles are derived from 24 types of articles for *WoS* and 25 types of articles for *Scopus*. Only eight types of articles in both databases contributed completely to the citable category comprising 918 articles. The matching citable articles in these two databases are 995. However, *WoS* reports 1042 articles and *Scopus* reports 1457 articles as citable articles of *Nature* for the year 2005. As a result, one-tenth of articles in *WoS* and one-third of articles in the *Scopus* are added to the denominators for calculation of IF and SJR, reducing their ranking.

Deviation by inclusion of public information

In *Scopus*, there are unjustifiable additions to the number of articles. For example, in *CA A Cancer Clin*, a page by the editors informing its readers about the inability to publish a specific article 'Cancer Survey, 2006' has been included. Likewise in *Curr Sci*, articles such as 'General article' and 'Special section' articles that deal with topics intended for general public are classified as citable articles in *WoS*.

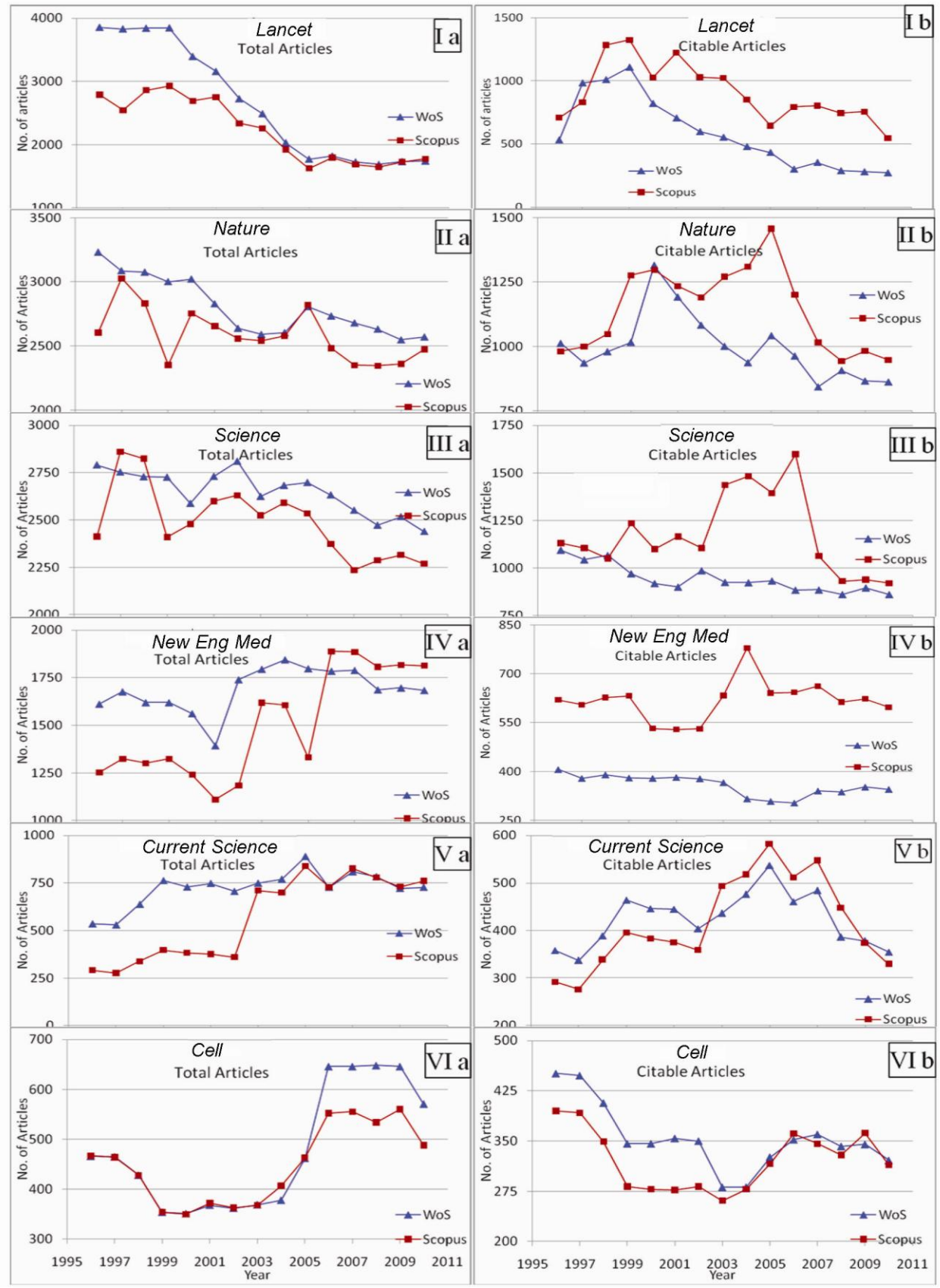

Figure 1. Total number of articles counted every year for the six journals by *WoS* and *Scopus* are shown in Ia to VIa. Calculation of citable articles by these database is given in Ib to VIb.

Table 2. Articles not classified as original research by the journals are categorized as citable articles by *WoS* database

Type of article (as per the Journals)	Title of the article	Journal	Month and Year
Artists on Science: Scientists on Arts	Science in Literature	*Nature*	Mar. 2005
Book Review	Science in Culture (Art Advances in Science)	*Nature*	Sep. 2000
Books and Arts	Science in Culture: Womb with a View?	*Nature*	May 2005
Business	Olympus Finds Market Rival Hard to Swallow	*Nature*	Dec. 2005
Careers and Recruitment	US Science Shocked by Revelations of Sexual Discrimination	*Nature*	Jun. 2000
Concepts	More than Words	*Nature*	Sep. 2001
Editorial	New Climate News	*Science*	Nov. 2000
Essay	Biodiversity and Nature Conservation: Some Common Arguments and Alternative Views	*Interdiscipl Sci Rev*	Dec. 2009
Essay on Science and Society	A Tale of Two Cities: Architecture and the Digital Revolution	*Science*	Aug. 1999
Feature	China's Environment in Globalizing World	*Nature*	Jun. 2005
Futures	The Party's Over (The Party's Over: It was only a Game….)	*Nature*	Jan. 2005
General Article	Foreign R&D Institutes in India: Is there any Positive Impact	*Current Science*	Feb. 2008
Invited Responses	Opposition is True Friendship	*Interdiscipl Sci Rev*	Dec. 2010
News and Views	Medical Technology: Balancing the Unbalanced	*Nature*	Oct. 2003
News and Views Feature	The Cultural Wealth of Nations	*Nature*	Mar. 2004
News Features	What's in a Name?	*Nature*	May 2001
News Focus Neuroscience	Turning Thoughts into Actions	*Science*	Oct. 1999
News Focus Economic Development	A Shifting Equation Links Modern Forming and Forests	*Science*	Nov. 1999
News Focus Scientific community	The Trials of the Butler (Scientific Community – The Trials of the Butler)	*Science*	Dec. 2003
News Frontiers in Optics	Holograms can Store Terabytes. But Where?	*Science*	Nov. 1999
Outlook	Malaria between Hope and a Hard Place (Between Hope and a Hard Place)	*Nature*	Aug. 2004
Policy Forum Medicine	Toward Safe and Effective Medical Abortion	*Science*	July 1998
Policy Forum Science Priorities	An Opportunity-based Science Budget	*Science*	Nov. 2000
Policy Forum Public Health	Is Safe Mail Worth the Price?	*Science*	Feb. 2002
Product Review	Plastic Disposables in the Laboratories	*Nature*	Oct. 1987
Research Commentaries Physics	Chaos Has Come Again	*Science*	Feb. 1998
Research Highlights	Correction (The director of Kew gardens charts the journey towards mapping the variety of plant life)	*Nature*	Dec. 2005
Tech Sight Meta Language	XML is Hatching	*Science*	Aug. 1998

There are articles in *CA A Cancer Clin* (2005, **55**, 195–198; 2004, **54**, 362–365) in the category of 'Patient Pages' informing the patients about the definition of various medical terms and their functions in the human body. *WoS* and *Scopus* include this as citable article, thus increasing numbers in the denominator. For a journal such as *CA A Cancer Clin*, which publishes very limited number of articles (around 20), adding or removing even a single article will substantially change the IF and SJR.

Discrimination in denominator

WoS and *Scopus* also show a decrease in the number of articles published by the journals. The December 2009 issue of *Notes Rec R Soc* published seven articles in the category of research articles. While *Scopus* included all of them, *WoS* included only three in its category of citable articles.

In the year 2010, the same journal published a special issue that contained 16 research articles. While *Scopus* included 14 of them in the citable articles category, *WoS* completely ignored these articles.

Similarly, in the journal *Trends Biochem Sci*, *WoS* did not include articles in the 'opinion' category as citable articles for the years 2004 and 2005. This resulted in decrease in citable articles from 106 to 96 for the year 2004 and 99 to 76 for the year 2005. However, in the subsequent years, articles under the opinion category are included as citable articles.

Duplication of articles with different titles

There are duplications of articles by *Scopus* through partial alteration of the titles. As a result, one article gets two different titles for the same journal with the same volume, issue and page numbers. The duplicate citable articles in the journal *Nature* for the year 2005 by the *Scopus* database are given in Table 4. Citations for the seven articles in *WoS* (432) are close to the combined citations of the

Table 3. Articles published by *Nature* in the year 2005 are classified in 29 categories. *WoS* classifies these articles in six categories and *Scopus* in seven categories. 'Citable articles' come from 24 categories for *WoS* and 25 categories for *Scopus*. Categories of articles that match exactly for *WoS* and *Scopus* are marked in bold. *Scopus* has duplication of articles; those which exceed maximum number of articles in the categories are marked within brackets

	Total articles			Citable articles		
Nature 2005	*Nature*	*WoS*	*Scopus*	*WoS*	*Scopus*	Matching articles
Editorial	128	719	149	0	42	0
Res Highlights	32			1	1	0
News	422	607		0	98	0
News Feature	134			12	39	6
Business	32			2	21	1
Correspondence	185	193	184	0	0	0
Commentary	22			0	10	0
Books and Arts	183	167		2	0	0
Essays	30			4	4	2
News and Views	376			7	21	0
Brief Communication	101			12	102(1)	6
Brief Communication Arising	60			2	14	0
Insights	53			39	43	39
Review	**13**	**56**	**164**	**13**	**13**	**13**
Articles	**131**	**986**	**1289**	**131**	**131**	**131**
Letters	**752**			**752**	**753(1)**	**752**
Technology Feature	24			13	12	9
Mars	**7**			**6**	**6**	**6**
Futures	47			2	0	0
Year of Physics	11			6	6	5
Artists on Science: Scientists on Arts	10			9	5	5
Feature	**2**			**2**	**2**	**2**
Outlook: India	11			10	4	4
News and Views feature	1			1	0	0
Progress	**2**			**2**	**3(1)**	**2**
Physics Detective	5			2	0	0
Huygens Articles	**7**			**7**	**7**	**7**
The Chimpanzee Genome	**7**			**5**	**5**	**5**
Nature Jobs	206				114	0
2005 Gallery					1	0
Short Survey			510			0
Note			467			0
Conference Paper			4			0
Total	**2994**	**2728**	**2767**	**1042**	**1457**	**995**

duplicate articles in *Scopus* (434). The number of citations per article, which should be 62, is reduced to 31 for these 7 duplicated articles. This duplication results in increase of denominator which results in reduction of SJR.

If *Scopus* gives citations only to one of the duplicated articles, it could be an error. However, sharing of the citations in duplicate articles clearly indicates that it cannot be an error. Authors, in general, read the articles from original publications or from citations given in articles. Therefore, it is unlikely that errors are made by the authors.

Discussion

Addition of citations by authors and journals is designed to increase citations. Increase in citations benefits author, as it may lead to improved visibility, *H*-index and advancement in career. Journals can increase their citations either by asking the authors to cite the previous articles or citing their own publication in the editorials. Increase in citations of journals improves their IF, SJR and other such ranking methods.

Databases, on the other hand, can influence citations as well as the number of 'citable articles' published by the journals. The total number of research articles of the journals counted by *WoS* and *Scopus* is either arbitrary or a preferred selection. The difference between the databases increases for those journals that publish more informative articles. It remains more or less equal when the journals publish original research output or reviews; for example, *Chem Rev* and *Tetrahedron*.

The selection of citable articles, according to the selection criteria, is based on critical evaluation by experts. However, the inclusion of 'Cultural wealth of nations',

Table 4. Addition of articles by duplicating with modified titles by *Scopus*. Citations of the articles in *Scopus* are given in brackets. Combined citations of duplicate articles of *Scopus* are more or less equal to *WoS* citations

Title of the article	Common features for both articles			Title of the repeated article by *Scopus*	*WoS*
	Journal	Issue and page	Author		
Environment: early ant plagues in the new world (0)	*Nature*, 2005, vol. 433	7021, pp. 32	Wilson, E. O.	Early ant plagues in the new world (5)	5
Burrowing mechanics: burrow extension by crack propagation (11)	*Nature*, 2005, vol. 433	7025, pp. 475	Dorgan, K. M., Jumars, P. A., Johnson, B., Boudreau, B. P., Landis, E.	Burrow extension by crack propagation (27)	35
Insect behavior: migratory bands give crickets protection (14)	*Nature*, 2005, vol. 433	7027, pp. 703	Sword, G. A., Lorch, P. D., Gwynne, D. T.	Migratory bands give crickets protection (10)	24
Materials chemistry: a synthetic enamel for rapid tooth repair (39)	*Nature*, 2005, vol. 433	7028, pp. 819	Yomagichi, K., Onuma, K., Suzuki, T., Okada, F., Tagami, J., Otsuki, M., Senawange, P.	A synthetic enamel for rapid tooth repair (11)	50
Indonesian earthquake: earthquake risk from co-seismic stress (62)	*Nature*, 2005, vol. 434	7031, pp. 291	McCloskey, J., Nalbant, S. S., Steacy, S.	Earthquake risk from co-seismic stress (24)	83
Developmental technology: dogs cloned from adults somatic cells (67)	*Nature*, 2005, vol. 436	7051, pp. 641	Lee, B. C., Kim, M. K., Jang, G., Oh, H. J., Yuda, F., Kim, H. J., Shamim, M. H., (…), Hwang, W. S.	Dogs cloned from adults somatic cells (82)	154
Protein glycosylation: chaperone mutations in Tn syndrome (32)	*Nature*, 2005, vol. 437	7063, pp 1252	Ju, T., Cummings, R. O.	Chaperone mutations in Tn syndrome (51)	81

'What's in a name', 'Opposition is true friendship', 'The trails of the butler' and 'Is the safe mail worth the price', to name a few, as citable articles by the databases clearly indicates that those who evaluate whether an article is citable do not seem to pay attention to the content and even to the title of the articles.

Altering the number of citable articles adversely affects the ranking of the poorly cited journals. IF and SJR of these poorly cited journals can be altered by either increasing or decreasing the number of citable articles by the databases. This causes high level of fluctuations as in *J Jpn Soc Tribologis* in 2001 and *Nav Architect* in 2001, 2006 and 2011; where errors exceed more than 100% of reported IF.

Alteration or changes in the number of 'citable articles' can significantly change the status of the well-cited journals that publish a limited number of articles per year. Increase or decrease of a few articles will lead to significant changes as in *Adv Phys*, in which an increase in the number of citable articles from 14 to 15 can result in change of IF from 37.000 to 34.533. Similarly, for *CA A Cancer Clin*, if the number of citable articles is 44 instead of 41, IF can change from 101.780 to 94.840.

Duplication of articles in the journal *Nature* in the year 2005, in which titles of some articles were modified, made them as additional articles by the *Scopus*. Further, it distributes the citations received by these articles. These activities clearly indicate lack of professional evaluation and assignment of articles of the journals by the databases.

Citations from the general articles such as editorials and letter, increase IF and SJR of the journals as they do not alter denominator. This can be curtailed by the databases by categorizing them as citable articles. Being unsure of citations or inclusion as citable articles by *WoS* and *Scopus* may force journals to reduce publishing the editorials, news, letter, etc. – thereby, losing the originality and characteristic of the journals.

Misrepresentation of an article in a different classification leads to ups and downs in the status of some of the journals. Terming those articles which cannot be considered as a research output, as original research articles, and selecting the citations from articles such as editorials, news, letters and others but not making them accountable in measuring the quality of the journals reflects poorly on the validity of the IF and SJR calculations. On the other hand, if the databases define criteria clearly and disclose the methodology adopted for the selection and evaluation of the journals to the public debate, the ranking of journals can be true and realistic.

1. Meho, L. I., The rise and rise of citation analysis. *Phys. World*, 2007, **20**, 32–36.
2. Egghe, L. and Rousseau, R., *Introduction to Informetrics: Quantitative Methods in Library, Documentation and Information Science*, Elsevier Science Publishers, Amsterdam, 1990.

3. Bar-Ilan, J., Citations to the 'Introduction to informetrics' indexed by *WoS*, *Scopus* and *Google Scholar*. *Scientometrics*, 2010, **82**, 495–506.
4. Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R. and Karageorgopoulos, D. E., Comparison of SCImago journal rank indicator with journal impact factor. *Faseb J.*, 2008, **22**, 2623–2628.
5. Klavans, R. and Boyack, K. W., Is there a convergent structure of science? A comparison of maps using the ISI and scopus databases, 2007.
6. Gorraiz, J. and Schloegl, C., A bibliometric analysis of pharmacology and pharmacy journals: *Scopus* versus *Web of Science*. *J. Inform. Sci.*, 2008, **34**, 715–725.
7. Kulkarni, A. V., Aziz, B., Shams, I. and Busse, J. W., Comparisons of citations in *Web of Science*, *Scopus*, and *Google Scholar* for articles published in general medical journals. *J. Amer. Med. Assoc.*, 2009, **302**, 1092–1096.
8. Huh, S., Citation analysis of *The Korean Journal of Internal Medicine* from KoMCI, *Web of Science*, and *Scopus*. *The Korean J. Internal Med.*, 2011, **26**, 1–7.
9. Lopez-Illescas, C., Anegon, F. D. M. and Moed, H. F., Comparing bibliometric country-by-country rankings derived from the *Web of Science* and *Scopus*: the effect of poorly cited journals in oncology. *J. Inform. Sci.*, 2009, **35**, 244–256.
10. Sember, M., Utrobicic, A. and Petrak, J., Croatian Medical Journal Citation Score in *Web of Science*, *Scopus*, and *Google Scholar*. *Croatian Med. J.*, 2010, **51**, 99–103.
11. Reuters, T., The Thomson Reuters journal selection process, 2011; http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/
12. Content Selection, SciVerse, 2012; http://www.info.sciverse.com/scopus/scopus-in-detail/content-selection
13. *Scopus*, Content Coverage Guide, 2012; http://www.info.sciverse.com/UserFiles/sciverse_scopus_content_coverage_0.pdf
14. *Web of Science*, Fact Sheet, 2012; http://thomsonreuters.com/content/science/pdf/Web_of_Science_factsheet.pdf
15. McVeigh, M. E. and Mann, S. J., The journal impact factor denominator: Defining citable (counted) items. *Jama–J. Amer. Med. Assoc.*, 2009, **302**, 1107–1109, 1972.
16. SCImago, *SJR* – SCImago Journal and Country Rank, 2007 (20 June 2011); www.scimagojr.com
17. Gavel, Y. and Iselid, L., *Web of Science* and *Scopus*: a journal title overlap study. *Online Inform. Rev.*, 2008, **32**, 8–21.
18. Moed, H. F. and Van Leeuwen, T. N., Improving the accuracy of institute for scientific information's journal impact factors. *J. Amer. Soc. Inform. Sci.*, 1995, **46**, 7.
19. Bharathi, D. G., Methodology for the evaluation of scientific journals: Aggregated Citations of Cited Articles. *Scientometrics*, 2011, **86**, 563–574.

ACKNOWLEDGEMENTS. I thank the Director, Central Leather Research Institute, Chennai for his permission to publish the paper. I also thank Aparna Srinivasan for editing the manuscript. I also thank to Dr T. P. Sastry, CLRI, Chennai for his support.

Received 12 October 2012; re-revised accepted 4 June 2013