Mapping excellence and diversity of research performance in India

Gangan Prathap

We propose a framework score to see how Indian universities and research-focused institutions fare in the world of high-end research in terms of excellence and diversity of their research base. For this we use a web application available in the public domain which visualizes scientific excellence worldwide in several subject areas. Only in 15 subject areas does India have a presence among global institutions in the 22 areas in which there are at least 50 institutes globally that have published more than 500 papers. The country has no institution which can be counted globally in seven areas: arts and humanities; business, management and accounting; health professions; neuroscience; nursing; psychology, and social sciences. India's research base is completely skewed towards the physical sciences and engineering with very little for biological sciences and medicine, and virtually none in social sciences, and arts and humanities when excellence at the highest level is considered.

Keywords: Bibliometrics, framework score, research diversity, scientific excellence.

THE Confederation of Indian Industry (CII) and Nature Index have joined hands to report on India's place in global science[1]. An optimistic picture of Indian science emerges showing significant growth in high-quality scientific publication and also reveals a particular strength in the broad discipline of chemistry. Indeed, India fares well when compared with countries that have similar economic size and conditions (including Australia, Brazil, Italy, Russia, Singapore, South Korea and Taiwan). Nature Index tracks only the affiliations of research articles published in a select group of 68 superior science journals. Although these are arguably the journals one would choose to publish his/her best research, the restriction leads to an imagined 'ascent to world-class science' and a 'historic love affair with chemistry'. We shall demonstrate that by using a much larger aggregator like *Scopus*, some of these statements may have to be qualified.

We shall work with a particular web application which visualizes scientific excellence worldwide in 22 major subject areas[2–5]. Table 1 lists the 22 subject areas covered by *Scopus* data collected for the *SCImago* Institutions Ranking[6], which is the basis for the web application. The latest and fourth release is based on articles during the publication period 2008–2012. Only those institutions (universities or research-focused) that have published at least 500 articles, reviews and conference papers in each category within the publication period are covered. Also, only subject categories where globally at least 50 institutions are found meeting this criteria are included in the web application. The full counting method was used to attribute papers from the *Scopus* database to institutions: if an institution appears in the affiliation field of a paper, it is fully attributed to this institution (with a weight of 1). In the Indian context, this considerably inflates the performance of those institutions which participate in global programmes like that of CERN, etc.

Table 1. Twenty-two subject areas covered by *Scopus* data collected for the *SCImago* Institutions Ranking

Subject area
Agricultural and biological sciences
Arts and humanities
Biochemistry, genetics and molecular biology
Business, management and accounting
Chemical engineering
Chemistry
Computer science
Earth and planetary sciences
Energy
Engineering
Environmental science
Health professions
Immunology and microbiology
Materials science
Mathematics
Medicine
Neuroscience
Nursing
Pharmacology, toxicology and pharmaceutics
Physics and astronomy
Psychology
Social sciences

Gangan Prathap is in Vidya Academy of Science and Technology, Thrissur 680 501, India and A. P. J. Abdul Kalam Technological University, Thiruvananthapuram 695 016, India.
e-mail: gangan@vidyaacademy.ac.in

Of the 22 subject areas in which the web application performs worldwide comparison[2], India has no presence in seven areas – in Arts and humanities; business, management and accounting; health professions; neuroscience; nursing; psychology, and social sciences. The country does not even have a single institution that meets the required threshold of being able to publish 500 papers during 2008–2012 in these areas (i.e. an average of 100 papers for each year in the period covered).

In this article, we look closely at all Indian institutions which appear on the list in the remaining 15 areas to see how they perform relative to each other in each area and also how the aggregate performance in one area compares with another. There are 70 institutions that make the cut. As many as 34 institutions appear in only one area each, whereas one (CSIR) appears in 14 out of the 15 areas. The only area in which CSIR does not play a prominent role is mathematics. In all, there are 213 individual entries across 15 subject areas. Note that CSIR, ISRO, DRDO, ICMR, etc. are counted as single entities. Only one institution from the corporate sector, namely Tata Sons Ltd, makes it to this list.

Methodology

We use the web application to build-up India-specific indicators[2]. For each institution in a specific subject area, we count the number of papers published, P, and the associated best paper rate (BPR). BPR is the proportion of publications from an institution which belongs to the 10% most cited publications in their respective subject area and publication year. We can then use the indicator i = BPR/10 to be a measure of quality. BPR corresponds to PP (top 10%) used in the Leiden Ranking and the excellence rate used in the *SCImago* Institutions Ranking[6]. The excellence rate is a field-normalized, size-independent indicator which serves as a measure of the high-quality output of research institutions. We can then compute a single-valued composite outcome indicator for the research performance of each institution in each area by introducing the second-order indicator[7] called the exergy term from the quantity (size) and quality (excellence) indicators, $x = i^2P$.

Within an area, we will find several institutions that have P and i varying considerably. Thus, the size-dependent proxy for research performance may vary by orders of magnitude. Similarly, when we take within an institution, a subject-wise cross-section, P, i and X vary considerably, as we shall see later in the text. There is therefore a huge variation in performance. This issue of diversity was addressed recently[8]. It was argued that structural diversity – the diversity of disciplines, institutions and support mechanisms is needed as 'it is a property of a "strong" research base that not only produces great research today but also has the capacity to address new challenges flexibly and responsively tomorrow. It is distinct from the contribution made by social diversity – the diversity of gender, nationality and ethnicity – to productivity, innovation and social cohesion'[8]. We argue that in a system or set of j categories or sources (that is, institutions within a discipline or area, or disciplines or areas within an institution), if x_j is the exergy of each source of a total of S sources, then we can have a measure of consistency or evenness of distribution η defined as follows[7]

$$X = \Sigma x_j,$$

$$E = \Sigma x_j^2,$$

$$x = X/S,$$

and, $$\eta = X^2/(SE).$$

We now need a measure that combines performance as measured by x_j and X with diversity[8]. The Stirling approach to diversity[9] adopted in Bornmann *et al.*[5] combines three basic properties: variety, balance and disparity. In our case, S is the measure of variety as it is the number of categories into which system elements (institutions in an area or areas within an institution) are apportioned. For example, we have 34 institutions in India in physics and astronomy that have published more than 500 papers during 2008–2012. CSIR has 14 subject areas in which it has published more than 500 papers during the same period. All else being equal, the greater the variety, the greater the diversity[8]. In the present case, we interpret balance as a function of the variation of x_j elements across categories. It performs the same role as statistical variance. We find that η as defined above is a natural candidate for measuring this and $\eta = 1$ is the ideal condition when all elements perform at the same level. Again, all else being equal, the more even the balance, the greater the diversity[5].

Since disparity has no role to play in the present context[8] (unlike in say ecology or economics), we can propose a framework score (F), which combines the number of elements in a system S, the total exergy X within the system (institutions within an area or areas within an institution), and the balance as the product $F = \eta X$. We shall use this framework score to see how the Indian science ecosystem is faring.

Results

As there are 70 unique entities and 15 subject areas, there is a large number of tables and figures that can be generated. We can therefore only give illustrative examples that show how the framework score can provide a picture of the excellence and diversity of the research base in the country. Of the 70 unique entities that appear in the list of universities and research-focused institutions, CSIR emerges with the highest framework score. Table 2

Table 2. Subject-wise performance of the Council of Scientific and Industrial Research during 2008–2012

Subject area	No. of papers	Best paper rate	X	Size S	Average x	Consistency η	F-score
Engineering	4088	16.7	11,401.0	14	4477.0	0.69	43,460.78
Materials science	5695	12.5	8898.4				
Physics and astronomy	4130	12.9	6872.7				
Chemistry	8266	8.9	6547.5				
Pharmacology, toxicology and pharmaceutics	3115	12.0	4485.6				
Medicine	2362	13.7	4433.2				
Energy	849	22.2	4184.2				
Chemical engineering	3990	10.0	3990.0				
Environmental science	2402	12.6	3813.4				
Biochemistry, genetics and molecular biology	5833	7.2	3023.8				
Agricultural and biological sciences	2642	9.6	2434.9				
Computer science	821	12.6	1303.4				
Earth and planetary sciences	1623	7.2	841.4				
Immunology and microbiology	1289	5.9	448.7				

Table 3. Institution-wise performance in the area of materials science during 2008–2012.

Institution	No. of papers	Best paper rate	X	Size S	Average x	Consistency η	F-score
Council of Scientific and Industrial Research	5695	12.5	8898.44	24	1503.8	0.40	14,263.13
Indian Institute of Science	2463	12.5	3848.44				
Indian Association for the Cultivation of Science	991	19.2	3653.22				
Shivaji University	568	22.5	2875.50				
Indian Institute of Technology, Kharagpur	2145	10.5	2364.86				
Bhabha Atomic Research Centre	1965	9.8	1887.19				
Indian Institute of Technology, Madras	1716	10.3	1820.50				
Indian Institute of Technology, Delhi	1648	9.2	1394.87				
Indian Institute of Technology, Bombay	1393	8.9	1103.40				
University of Delhi	835	11.3	1066.21				
Indian Institute of Technology, Roorkee	1018	9.9	997.74				
Banaras Hindu University	1061	9.4	937.50				
Jadavpur University	1076	9.1	891.04				
University of Hyderabad	600	11.5	793.50				
Indian Institute of Technology, Kanpur	1310	7.2	679.10				
Indian Institute of Technology, Guwahati	642	10.0	642.00				
National Institute of Technology, Tiruchirappalli	545	9.8	523.42				
University of Calcutta	518	8.4	365.50				
Defence Research and Development Organisation	1158	5.4	337.67				
Anna University	1389	4.9	333.50				
Indira Gandhi Centre for Atomic Research	902	5.8	303.43				
University of Madras	596	5.3	167.42				
VIT University	513	5.7	166.67				
Mangalore University	899	2.1	39.65				

provides a glimpse of how the indicators are computed from the number of publications and BPR for each subject area in which CSIR has published more than 500 papers. Its strongest research base is in engineering and the physical sciences. It also shows a healthy presence in the life and medical sciences.

Table 3 shows how the institution-wise performance can be demonstrated for a chosen subject area – in this case it is materials science during 2008–2012. We see a huge range of the X-score, from 8898.44 for CSIR to 39.65 for Mangalore University.

Table 4 provides a list of the 70 Indian universities and research-focused institutions which have published more than 500 papers in the respective areas during 2008–2012. Only one institution from the corporate sector, namely Tata Sons Ltd makes it to this list. Again, we see a use range of variation of all indicators – S ranges from 1 (34 institutions) to 14 (CSIR); x from 19.9 (University of Mysore) to 5474.6 (Panjab University), η from 0.39 (University of Delhi) to 1 (34 institutions), and F from 19.9 (University of Mysore) to 43460.8 (CSIR).

Table 5 shows the cumulative F-scores for the 22 subject areas covered by *Scopus* data collected for the *SCImago* Institutions Ranking. India's strongest research base is in engineering, thanks to the role played by CSIR and the Indian Institutes of Technology. Taken together with chemical engineering, computer science, energy, and materials science, it is clear that India's love affair is with engineering in the broadest sense. Table 6 reveals this clearly, where the F-scores are arranged in broader

Table 4. Indian universities and research-focused institutions which have published more than 500 papers in the respective areas during 2008–2012

Institution	Size	Average x	η	F-score
Council of Scientific and Industrial Research	14	4477.0	0.69	43,460.78
Indian Institute of Science	10	2424.0	0.61	14,826.37
Indian Institute of Technology, Kharagpur	8	2406.0	0.51	9789.07
Indian Institute of Technology, Delhi	8	2134.3	0.47	8063.80
Indian Institute of Technology, Madras	7	1840.4	0.62	8014.60
Jadavpur University	7	1585.0	0.63	6977.99
Banaras Hindu University	8	1129.6	0.76	6868.27
Indian Association for the Cultivation of Science	3	2503.5	0.90	6743.09
Indian Institute of Technology, Bombay	7	1605.5	0.60	6699.33
Indian Institute of Technology, Kanpur	7	1399.3	0.67	6524.05
Panjab University	2	5474.6	0.51	5622.34
Bhabha Atomic Research Centre	9	1116.5	0.50	5059.31
Indian Institute of Technology, Roorkee	6	1359.9	0.58	4706.59
University of Delhi	8	1472.6	0.39	4601.39
Tata Institute of Fundamental Research	2	3925.2	0.57	4490.47
Indian Institute of Technology, Guwahati	5	923.3	0.84	3883.02
Jawaharlal Nehru Centre for Advanced Scientific Research	2	1744.9	0.95	3302.80
Anna University	7	644.8	0.65	2912.13
Shivaji University	1	2875.5	1.00	2875.50
University of Hyderabad	4	656.5	0.87	2273.48
National Institute of Technology, Tiruchirappalli	3	1227.7	0.59	2165.42
Saha Institute of Nuclear Physics	1	2135.1	1.00	2135.06
University of Calcutta	4	504.5	0.95	1907.73
University of Rajasthan	1	1614.9	1.00	1614.87
Indian Statistical Institute	2	978.3	0.79	1549.64
Defence Research and Development Organisation	4	607.8	0.63	1519.87
All India Institute of Medical Sciences	2	1173.0	0.65	1518.94
Variable Energy Cyclotron Centre	1	1409.9	1.00	1409.92
Indian Council of Medical Research	3	976.4	0.48	1399.54
Indira Gandhi Centre for Atomic Research	3	481.1	0.89	1279.69
Jamia Hamdard	1	1276.0	1.00	1275.95
Institute of Chemical Technology, Mumbai	2	898.5	0.68	1226.92
Annamalai University	5	534.8	0.42	1124.98
Aligarh Muslim University	4	454.2	0.61	1108.82
Indian Council of Agricultural Research	5	383.5	0.51	974.76
VIT University	4	378.5	0.62	944.99
National Institute of Technology, Rourkela	2	614.8	0.72	890.91
Tata Memorial Centre	1	789.7	1.00	789.69
Bengal Engineering and Science University, Shibpur	1	757.9	1.00	757.92
Motilal Nehru National Institute of Technology	1	744.9	1.00	744.93
S.N. Bose National Centre for Basic Sciences	1	719.2	1.00	719.17
Postgraduate Institute of Medical Education and Research	1	629.9	1.00	629.93
Christian Medical College, Vellore	1	614.1	1.00	614.07
Sanjay Gandhi Postgraduate Institute of Medical Sciences	1	608.7	1.00	608.71
University of Madras	4	141.7	0.94	534.96
University of Pune	1	523.5	1.00	523.45
Indian Space Research Organization	2	261.4	1.00	521.67
International Institute of Information Technology, Hyderabad	1	356.5	1.00	356.51
Natoinal Institute of Technology Karnataka	1	305.0	1.00	304.97
Manipal University	3	112.5	0.90	304.51
UGC–DAE Consortium for Scientific Research, Indore	1	288.3	1.00	288.30
PSG College of Technology	1	258.5	1.00	258.48
Tata Sons Ltd	1	246.0	1.00	245.98
Inter-University Accelerator Centre	1	228.3	1.00	228.32
Punjab Agricultural University	1	227.1	1.00	227.07
Physical Research Laboratory	1	216.7	1.00	216.71
King George's Medical University	1	198.1	1.00	198.10
Sree Chitra Tirunal Institute for Medical Sciences and Technology	1	175.6	1.00	175.60
Seth Gordhandas Sunderdas Medical College and King Edward Memorial Hospital	1	147.9	1.00	147.92
Mangalore University	3	51.7	0.90	140.27
Raja Ramanna Centre for Advanced Technology	1	129.4	1.00	129.43
National Institute of Mental Health and Neurosciences	1	126.1	1.00	126.08
Jawaharlal Institute of Postgraduate Medical Education and Research	1	101.4	1.00	101.44
Tamil Nadu Agricultural University	1	100.1	1.00	100.08
Lady Hardinge Medical College	1	62.1	1.00	62.08
Guru Tegh Bahadur Hospital	1	48.0	1.00	48.02
University College of Medical Sciences	1	47.6	1.00	47.59
Govind Ballabh Pant University of Agriculture and Technology	1	26.9	1.00	26.87
Chaudhary Charan Singh Haryana Agricultural University	1	26.2	1.00	26.24
University of Mysore	1	19.9	1.00	19.94

Table 5. *F*-scores for the 22 subject areas covered by *Scopus* data collected for *SCImago* Institutions Ranking

Subject area	*F*-score
Engineering	38,887.78
Physics and astronomy	25,925.64
Computer science	14,778.86
Materials science	14,263.13
Chemistry	8741.92
Mathematics	7281.23
Medicine	5685.57
Chemical engineering	5208.73
Energy	4251.51
Pharmacology, toxicology and pharmaceutics	3077.66
Agricultural and biological sciences	2324.30
Biochemistry, genetics and molecular biology	2108.51
Environmental science	1951.05
Earth and planetary sciences	1760.32
Immunology and microbiology	613.72
Arts and humanities	0
Business, management and accounting	0
Health professions	0
Neuroscience	0
Nursing	0
Psychology	0
Social sciences	0

Table 6. *F*-scores arranged in broader groupings and according to the panels used in the Research Excellence Framework of the UK in 2014

Subject area	*F*-score
Engineering	77,390.00
Physics and astronomy	25,925.64
Medicine and life sciences	11,485.45
Chemistry	8741.92
Mathematics	7281.23
Agricultural and biological sciences	2324.30
Environmental science	1951.05
Earth and planetary sciences	1760.32
Research Excellence Framework 2014 panel	
Biological sciences and medicine	13,809.76
Physical sciences and engineering	123,050.17
Social sciences	0.00
Arts and humanities	0.00

groupings or according to the four panels used in the Research Excellence Framework of the UK in 2014. India's research base is completely skewed towards the physical sciences and engineering (nearly 90%) with very little for biological sciences and medicine (the remaining 10%), and virtually none in social sciences, and arts and humanities when excellence at high levels implied by the use of BPR is considered.

Concluding remarks

Thanks to the generosity of its creators, we now have web applications available in the public domain which visualize scientific excellence worldwide in several subject areas[2]. We have now proposed a framework score to see how Indian universities and research-focused institutions fare in the world of high-end research in terms of excellence and diversity. Only in 15 subject areas does India have a presence among global institutions. It has no institution which can be counted globally in seven areas: arts and humanities; business, management and accounting; health professions; neuroscience; nursing; psychology, and social sciences. India's research base is completely skewed towards the physical sciences and engineering with very little for biological sciences and medicine, and virtually none in social sciences, and arts and humanities when excellence at the highest level is considered.

It would seem from this that India concentrates its strengths and its research institutions in the physical sciences and engineering sectors, and only a token presence is seen in the life sciences, medical and biotechnology sectors. There seems to be no visible output at the highest levels regarding the attention it needs to give to various social and economic challenges.

1. Kogleck, L., Priyadarshini, S., Pincock, S. and Bocquet, A., Indian science ascending. December 2015; http://www.natureindex.com/news/indian-science-ascending
2. http://www.excellencemapping.net/#/view/measure/top10/calculation/a_ohne_kovariable/field/materials-science/significant/false/org/
3. Bornmann, L., Stefaner, M., de Moya Anegón, F. and Mutz, R., Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualization of results from multi-level models. *Online Inf. Rev.*, 2014, **38**(1), 43–58.
4. Bornmann, L., Stefaner, M., de Moya Anegón, F. and Mutz, R., What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. *J. Inform.*, 2014, **8**(3), 581–593.
5. Bornmann, L., Stefaner, M., de Moya Anegón, F. and Mutz, R., Ranking and mappping of universities and research-focused institutions worldwide: the third release of excellencemapping.net. *COLLNET J. Scientometrics Inf. Manage.*, 2015, **9**(1), 61–68.
6. http://www.scimagoir.com/
7. Prathap, G., The Energy–Exergy–Entropy (or EEE) sequences in bibliometric assessment. *Scientometrics*, 2011, **87**(3), 515–524.
8. Anon, The value of structural diversity: Assessing diversity for a sustainable research base. Digital Science and the Science Policy Research Unit, University of Sussex, UK, December 2015.
9. Stirling, A., A general framework for analysing diversity in science, technology and society. *J. R. Soc. Interface*, 2007, **4**, 707–719.

ACKNOWLEDGEMENTS. I thank the creators of the web applications; Lutz Bornmann, Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Munich, Germany; Rüdiger Mutz, Professorship for Social Psychology and Research on Higher Education, ETH Zurich, Switzerland; Moritz Stefaner, Truth & Beauty Operator, and Felix de Moya Anegón, CSIC/CCHS/IPP, SCImago Group (Spain), Communication and Information Science Faculty, University of Granada, Spain for making them available in the public domain. I also thank Lutz Bornmann for useful discussions.

Received 26 January 2016; accepted 31 March 2016

doi: 10.18520/cs/v111/i3/470-474