DESIDOC Journal of Library & Information Technology, Vol. 34, No. 3, May 2014, pp. 191-196

A Bibliometric Evaluation of Research on the Monsoon

Gangan Prathap
National Institute for Interdisciplinary Science and Technology, CSIR, Thiruvananthapuram-695 019
E-mail: gp@niist.res.in

ABSTRACT

With concerns about climate change and its impact on the monsoon cycle over South Asia and its ramifications elsewhere on the globe, there has been renewed interest in the science behind the forces that drive the annual cycle and its variability from year to year and within the season. This is reflected in the growing research output in this area, particularly from India and China. In this paper, a comprehensive and in-depth bibliometric analyses that breaks down scholarly performance into three components - quantity, quality and consistency have been conducted. The citation data is retrieved from the Web of Science. The most productive organisations, countries, authors and also the most influential journals in which this newly emerging area is published using these criteria, with particular emphasis on the Indian contribution have been identified.

Keywords: Bibliometrics, three-dimensional evaluation, indicators, quality, quantity, consistency, citation, monsoon

1. INTRODUCTION

The monsoon system determines the weather and climate over most parts of Asia, mainly the Indian sub-continent, and over South-east Asia and large parts of China. Directly, it influences the lives of nearly one-third of the world population and indirectly the lives of nearly three-quarters of the world population.

At one time, India accounted for nearly half of the academic research in this area. This is not surprising, as ever since the name originated from the Arabic (‘mausam’ for seasons became ‘monsoon’), the monsoon, characterised by seasonally reversing winds and rains was associated mainly with the Indian and South Asian regions. It is only recently with the emergence of global satellite data and images and computational models that integrate the global earth and atmospheric systems that the monsoon began to be seen as a very complex global phenomenon.

As predictable as the annual cycle of the monsoon is in bringing wet summers and dry winters to the South Asian region, equally unpredictable is its variability from year to year and even within a season. A small negative departure from the expected long term average rainfall can mean a significant drop in food production and this understandably has a considerable impact on the economies of this region. Because of global warming and climate change, changes in the monsoon rainfall may have a large impact on agricultural productivity in the coming years[1]. This is one reason why there is now increased attention given to monsoon forecasting. This is clear from the bibliometric evidence from 1987 to 2012, that the number of records in the *Web of Science* database has increased from 48 to 1833.

The *Web of Science* database allows us to refine the results in terms of publication years, countries, organisations, authors and journals (source titles), etc. The focus in this paper will be mainly on the research activity in India in this area. We examine this area of research in terms of the top countries, top organisations, top authors and finally the most influential journals in which these papers have appeared. The 3-D evaluation recently proposed by Prathap[2,3] is used. A simple heuristic model[4,5] using 2-D quantity (productivity in terms of number of papers published) and quality (specific impact as defined by citations per paper) are complemented with a third dimension, called consistency η. This enables a better 3-D evaluation of the information production process. If the number of papers is P, the quality (or impact si), is measured by the ratio C/P, where, C is the total number of citations received by P papers. The third dimension, consistency η is a measure of the variability in the quality of the individual papers in the publication set, or in other words, the shape of the distribution curve.

Using all three components together, a z-index can be computed from an energy-like term:

$Z = \eta X = \eta 2E$ as $z = Z^{1/3}$,

Received 2 April 2014, Online published 20 May 2014

which has the same dimensions as the number of publications, and therefore also the *h*-index. Where *X* is Exergy and E is Energy[4,5], it is possible to imagine a composite indicator named Zynergy for $Z = \eta X = \eta 2E$. This index combines quantity, quality, and consistency (or efficiency) to give a 3-D bibliometric evaluation. Thus, the *h*- and *z*- are secondary single number bibliometric indicators of performance.

The precise computation of η requires the knowledge of the complete citation sequence (i.e., the distribution curve) for each individual scientist (or aggregation like institute, journal or country). This is obtained directly from the *Web of Science* for each country, organisation, author and journal taken up in the present analysis and the methodology to obtain this is as follows:

2. SECOND-ORDER INDICATORS FROM THE PRIMARY QUALITY, QUANTITY AND CONSISTENCY INDICATORS

The journal impact factor is now increasingly accepted as a proxy or indirect measure of the quality or scholarly influence of a journal[6]. In the same way, the scientific output of an individual or an entity can be measured using the following parameter space:

2.1 Quantity

No. of papers/articles '*P*' published during a prescribed window which will be called the publication window (in this case, the window is from 1987 to the date of access of Web of Science database).

2.2 Quality

The impact *'i'* computed as *C/P*, where *C* is the number of citations during a prescribed citation window of all the articles *P*. Note that the definition of i needs two distinct windows to be identified, the publication window and the citations window. The famous JIF is based on the use of a publication window of two years immediately preceding a single year citation window[6-9]. In the present case, these are taken as identical, i.e., from 1987 till date of access.

Once the quantity *P* and quality *i* parameters are defined, it is possible to postulate the following sequence of indicators of performance[5]:

Zeroth order indicator: $P = i^0P$

First order indicator: $C = i^1P$

Second order indicator: $X = i^2P = i^1C = i$.

C is derived from the citation sequence, ck of the citations of each paper in a publication portfolio of *P* papers as the total number of citations, $C = \sum c_k$, $k = 1$ to P. Note that both *P* and *C* serve as indicators of performance in their respective ways. One can think of $C = iP$ as the first order indicator for performance. Prathap[4,5] showed that the exergy indicator $X = i^2P$, is an energy like quantity which can be thought of as a second order indicator of performance. This paradigm then leads to a trinity of energy like terms[4,5]:

$X = i^2P$

$E = \sum c_k^{\,2}$

$S = \sum(c_k - \mathrm{i})^2 = E - X$

where

$P = \sum 1$

$C = \sum c_k$

$i = C/P$.

Ever since the *h*-index was proposed, it is a common practice to rearrange the citation sequences in a monotonically decreasing order. Very high skews are noticed because the highly cited articles are found in a small core, implying a possible huge variation in the quality of each paper in the publication set. Thus, two different sets can have the same *C*, and one could have achieved this with far fewer papers, with a higher quality of overall performance, or with the same number of papers (i.e., same quality) but a higher degree of consistency or evenness. Thus, *C* by itself, which is a first-order indicator may not be the last word on the measurement of performance. The product $X = iC = i^2P$ is a robust second-order performance indicator[4,5] is arguably a better proxy for performance. Apart from *X*, an additional indicator *E* also appears as a second-order indicator. The coexistence of *X* and *E* allows us to introduce a third attribute that is neither quantity nor quality. In a bibliometric context, the appellation 'consistency' may be more meaningful. The simple ratio of *X* to *E* can be viewed as the third component of performance, namely, the consistency term $\eta = X/E$. Perfect consistency ($\eta = 1$, i.e., when $X=E$) is a case of absolutely uniform performance; that is, all papers in the set have the same number of citations, $ck = c$. The greater the skew, the larger is the concentration of the best work in a very few papers of extraordinary impact. The inverse of consistency thus becomes a measure of concentration.

Thus, for a complete 3-D evaluation of publication activity, one needs *P, i,* and η. These are the three primary components of a quantity–quality–consistency landscape.

3. METHODOLOGY

Consider the scientific output in the area described by Topic=(monsoon) as indexed in the *Web of Science* (a Thomson-Reuters product). The period 1986-All Years (updated 27-12-2013) was chosen for which subscription was available. All articles *P*, and citations *C* gathered by these *P* articles, are counted. Then the impact '*i*' is computed for this period. From the citation sequence for each entity (author, country, organisation or journal), consistency η can be computed using simple Excel spreadsheet functions.

4. DATA, RESULTS AND DISCUSSIONS

The analysis was refined using the topic option adopting the following strategy to get the world output:

Topic=(monsoon)
Timespan=All years,
Databases=SCI-EXPANDED, *CPCI-S*, *CPCI-SSH*, *CCR-EXPANDED, IC*

A total of 19,603 records are obtained. This is followed up with the address option to identify India's share in this area:

Topic = (monsoon)
Refined by: Countries/Territories = (INDIA)
Timespan = All years,
Databases = *SCI-EXPANDED*, *CPCI-S, CPCI-SSH, CCR-EXPANDED, IC*

Now, 5,167 records are found, i.e., during the 1987-date period, India accounted for 26.35 % of the academic research in this area. It is only recently with the emergence of global satellite data and images and computational models that integrate the global earth and atmospheric systems that the monsoon began to be seen as a very complex global phenomenon. Further details are displayed in Table 1 and Fig. 1. It can be seen that since 1991, India's share of the world activity had come down and remained at a stable level of approximately a quarter of the world output. Figure 2 shows where the competition is coming from China's presence in this area has risen rapidly from less than 5 % global share in 1987 to more than 30 % in 2012.

Table 2 shows that India ranks high among the countries publishing in this area. A default ranking of the top ten countries in this area of research put India at the second place in terms of publications alone. The US is the leading player in terms of numbers of papers published during the whole period, but most recently, China has overtaken the US as the most prolific contributor to

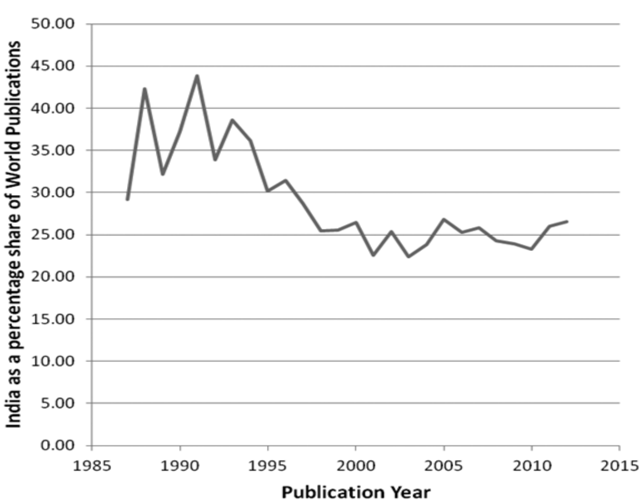

Figure 1. India's publications in monsoon research as a percentage share of world output.

Table 1. India's publications in monsoon research as a percentage share of world output

Year	No. of papers		
	World	India	% share
1987	48	14	29.17
1988	52	22	42.31
1989	56	18	32.14
1990	86	32	37.21
1991	219	96	43.84
1992	239	81	33.89
1993	277	107	38.63
1994	260	94	36.15
1995	321	97	30.22
1996	375	118	31.47
1997	380	109	28.68
1998	475	121	25.47
1999	504	129	25.60
2000	578	153	26.47
2001	572	129	22.55
2002	646	164	25.39
2003	813	182	22.39
2004	821	196	23.87
2005	869	233	26.81
2006	1056	267	25.28
2007	1206	311	25.79
2008	1365	332	24.32
2009	1444	345	23.89
2010	1616	377	23.33
2011	1651	430	26.04
2012	1833	487	26.57

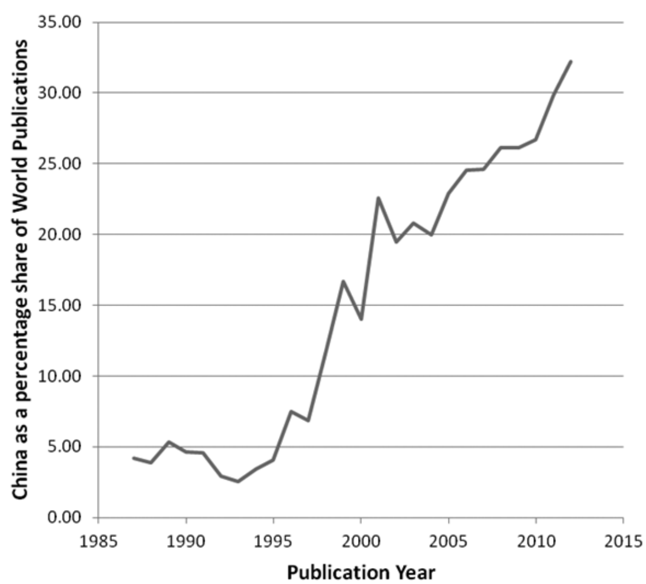

Figure 2. China's publications in monsoon research as a percentage share of world output.

the world of monsoon research. It is clear from the table that India has the lowest impact (here taken as a proxy for quality) among the countries in this list. England has the highest quality or impact of research. Arguably, the best measure for comparison would be the exergy term *X* and Table 2 ranks the countries using this measure.

The analysis according was further refined to the organisations option adopting the following strategy:

Topic = (monsoon)
Refined by: Organisations-Enhanced = (xxx)
Timespan = All years,
Databases=*SCI-EXPANDED*, *CPCI-S, CPCI-SSH, CCR-EXPANDED, IC*.

Table 3 shows the leading organisations in monsoon research ranked according to the default

Table 2. Top ten countries publishing in this area ranked by exergy *X*

Country	P	C	i	X
USA	5790	166094	28.69	4764631.58
England	1145	33308	29.09	968928.27
Germany	1364	34959	25.63	895990.97
Peoples Republic of China	4524	61775	13.65	843534.62
France	1212	28738	23.71	681413.07
Japan	1549	27308	17.63	481424.70
India	5167	45210	8.75	395576.56
Australia	962	19427	20.19	392316.35
Netherlands	362	9406	25.98	244400.10
Canada	351	7187	20.48	147159.46

quantity parameter *P*. The Chinese Academy of Sciences leads this list. Four Indian entities appear prominently in this list, –Council of Scientific and Industrial Research (CSIR); National Institute of Oceanography (NIO) which is a constituent laboratory of the CSIR;, Indian Institute of Tropical Meteorology (IITM) which belongs to the Ministry of Earth Science; and the various institutes belonging to the Indian Institute of Technology system. Table 4 shows more of the leading organisations from India in this area of research. Figures 3 and 4 summarise the performance of these institutions in a *h-z* 2-D map.

The analysis according was further refined to the authors option adopting the following strategy:

Topic = (monsoon)
Refined by: Countries/Territories = (INDIA) AND Authors=(xxx)
Timespan = All years,
Databases = *SCI-EXPANDED*, *CPCI-S, CPCI-SSH, CCR-EXPANDED, IC*

Table 3. Leading organisations in monsoon research in the world ranked according to default quantity parameter *P*.

Organisations	P	*i*	*η*	*h*	z
Chinese Academy of Sciences	2702	16.11	0.13	78	45.12
Council of Scientific & Industrial Research (CSIR), India	966	10.61	0.18	40	26.75
National Institute of Oceanography, India	617	13.30	0.20	40	27.99
Indian Institute of Tropical Meteorology	619	11.19	0.14	39	22.11
National Oceanic Atmospheric Admin NOAA, USA	565	37.20	0.24	74	57.47
China Meteorological Administration	550	8.79	0.21	34	20.71
University of California System	549	32.55	0.19	64	47.67
Indian Institute of Technology (IIT), India	528	8.96	0.16	31	18.77
National Aeronautics Space Administration (NASA), USA	512	33.20	0.30	69	55.58
University of Hawaii System	439	32.90	0.29	66	51.40

Table 4. Leading organisations in monsoon research in India ranked according to parameter *P*.

Organisations	P	*i*	*η*	*h*	z
Council of Scientific & Industrial Research	966	10.61	0.18	40	26.75
National Institute of Oceanography (NIO)	617	13.30	0.20	40	27.99
Indian Institute of Tropical Meteorology (IITM), Pune	619	11.19	0.14	39	22.11
Indian Institute of Technology (IIT)	528	8.96	0.16	31	18.77
Indian Institute of Science (IISc), Bangalore	273	24.32	0.09	39	24.09
Physical Research Laboratory (PRL), Ahmedabad	232	15.75	0.21	30	23.09
Indian Institute of Technology (IIT), Delhi	215	5.73	0.29	18	12.65
Indian Meteorological Department (IMD)	175	6.54	0.12	17	9.64
Indian Institute of Technol. (IIT), Kharagpur	135	12.23	0.11	19	13.22
Vikram Sarabhai Space Center (VSSC)	126	13.47	0.35	22	20.09
Cochin University of Science & Technology (CUSAT)	122	6.02	0.23	13	10.07
University of Calcutta	105	5.30	0.32	13	9.82

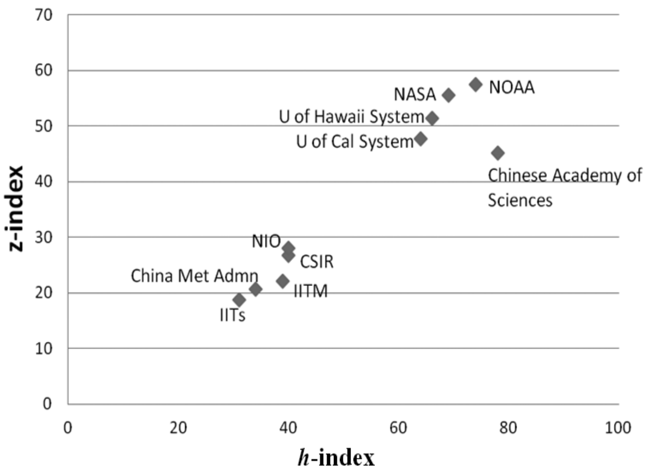

Figure 3. A 2-D *z*-*h* map of leading organisations in the world in monsoon research.

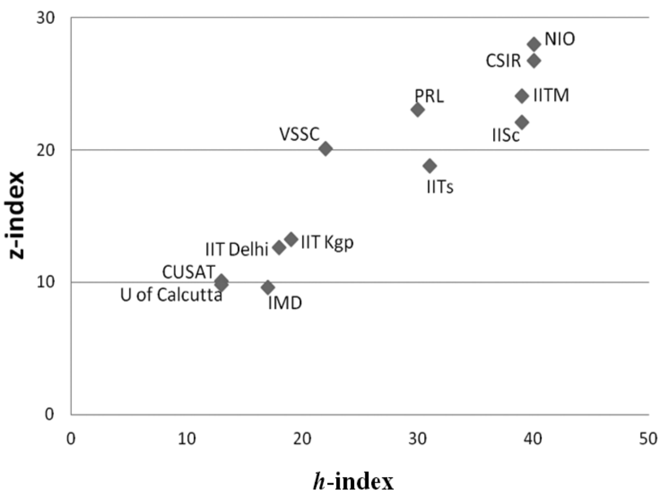

Figure 4. A 2-D *z*-*h* map of leading organisations in India in monsoon research.

Table 5 shows the leading authors in monsoon research in India ranked according to the *z*-index. Since there is poor or negative correlation between *P* and *i* and *η* respectively, ranking by the default quantity parameter *P* would have given a misleading picture about the best performing monsoon researchers.

Table 6 shows the leading journals in which Indian authors have published their results. These are mainly journals originating from India.

5. CONCLUSIONS

In this paper, a 3-D bibliometric analysis has been used to identify the leading countries, organisations and authors and also the most influential journals in the area of monsoon research. Scholarly performance is broken down into three components – quantity, quality, and consistency. The citation data retrieved from the Web of Science is used to categorise the entities according to these quantities. It is noticed

Table 5. Leading authors from India in monsoon research ranked according to the *z*-index

Authors	P	i	*η*	*h*	z
S.R. Shetye	33	37.79	0.49	19	28.38
S. Gadgil	37	32.35	0.43	20	25.47
S.K. Satheesh	32	25.88	0.51	16	22.22
K.K. Moorthy	62	20.19	0.43	22	22.21
S.S.C. Shenoi	37	27.08	0.37	16	21.60
K.R. Kumar	43	22.95	0.44	18	21.57
B.N. Goswami	82	37.34	0.07	21	19.91
D. Shankar	35	24.40	0.32	14	18.89
S.S. Babu	38	18.89	0.38	14	17.22
A.K. Gupta	38	27.37	0.17	13	16.98
V.S.N. Murty	40	15.55	0.27	12	13.71
R. Ramesh	61	14.21	0.21	14	13.69
M. Rajeevan	48	15.29	0.20	14	13.09
J. Srinivasan	42	10.95	0.43	13	12.98
P.C.S. Devara	42	9.07	0.41	11	11.24
U.C. Mohanty	102	4.92	0.37	11	9.67

Table 6. Leading journals in which Indian authors publish their research.

Journal name	No. of papers		
	World	India	% Share
Current Science	350	341	97.43
Indian Journal of Marine Sciences	205	191	93.17
Geophysical Research Letters	707	143	20.23
Mausam	149	131	87.92
Journal of Geophysical Research Atmospheres	834	129	15.47
Proceedings of the Indian Academy of Sciences Earth and Planetary Sciences	142	127	89.44
Environmental Monitoring and Assessment	143	125	87.41
International Journal of Climatology	459	125	27.23
Journal of Earth System Science	130	119	91.54
Meteorology and Atmospheric Physics	247	114	46.15
Journal of the Geological Society of India	126	110	87.30
International Journal of Remote Sensing	126	92	73.02
Atmospheric Environment	205	89	43.41
Theoretical and Applied Climatology	214	73	34.11
Climate Dynamics	502	71	14.14

that China has become a leading player in monsoon research, while relatively, India's output has remained stable. The most productive organisations and authors from India and also the most influential journals in which they publish have also been identified.

REFERENCES

1. Turner, A.G. & Annamalai, H. Climate change and the south asian monsoon. *Nature Climate Change*, 2012, **2**, 587-95.
2. Prathap, G. Quantity, quality, and consistency as bibliometric indicators. *J. Asso. Inf. Sci. Technol.*, 2013, **65**, 214.
3. Prathap, G. The zynergy-index and the formula for the *h*-index. *J. Asso. Inf. Sci. Technol.*, 2013, 65, 426-27.
4. Prathap, G. The energy–exergy–entropy (or EEE) sequences in bibliometric assessment. *Scientometrics*, 2011, **87**, 515-24.
5. Prathap, G. Quasity, when quantity has a quality all of its own: Toward a theory of performance. *Scientometrics*, 2011, **88**, 555-62.
6. Pendlebury, D.A. & Adams, J. Comments on a critique of the Thomson Reuters journal impact factor. *Scientometrics*, 2012, **92**, 395-401.
7. Garfield, E. Citation indexes to science: A new dimension in documentation through association of ideas. *Science*, 1955, **122**, 108-11.
8. Garfield, E. Journal impact factor: A brief review. *Canadian Med. Asso. J.*, 1999, **161**, 979-80.
9. Garfield, E. The agony and the ecstasy: The history and meaning of the journal impact factor. International Congress on Peer Review and Biomedical Publication, 2005. http://garfield.library.upenn.edu/papers/jifchicago2005.pdf/.

About the Author

Dr Gangan Prathap is working as an Outstanding Scientist, *National Institute for Interdisciplinary Science and Technology* (NIIST-CSIR), Thiruvananthapuram. He was trained as an Aerospace Engineer and specialised in mathematical modeling and computer simulation of complex problems in aerospace engineering. For more than thirty years, he has also pursued a parallel interest in research evaluation, bibliometrics, and scientometrics and the application of physical and mathematical insights for research assessment. In recent years, he has proposed a thermodynamic basis for bibliometric sequences which can lead to better indicators for research evaluation like the *p*-index and the EEE-sequences. He has served at many premier institutions in India like the National Aerospace Laboratory (NAL), Bangalore, Centre for Mathematical Medelling and Computer Simulation (C-MMACS) (now CSIR 4PI), Bangalore, Cochin University of Science & Technology (CUSAT), Cochin, and National Institute of Science Communication and Information Resources (NISCAIR), Delhi.