https://doi.org/10.51408/issi2025_067

Boost Formalism- A New Framework to Assess the Impact of Collaborations at Institutional Level

Prashasti Singh[1], Vivek Kumar Singh[2], Abhirup Nandy[3], Hiran H Lathabai[4]

[1]*prashasti.singh8@gmail.com*
Department of Computer Science, Shri Ram College of Commerce, University of Delhi, Delhi-110007 (India)

[2]*vivekks12@gmail.com*
Department of Computer Science, University of Delhi, Delhi-110007 (India)
NITI Aayog, Government of India, Delhi-10007 (India)

[3]*abhirupnandy.online@gmail.com*
Institute of Informatics and Communication, University of Delhi, Delhi-110067 (India)

[4]*hiranhl007@gmail.com*
Amrita CREATE, Amrita Vishwa Vidyapeetham, Amritapuri-690525, Kerala (India)

Abstract

Research collaboration at the international level has increased manifold during the last two decades. In addition to mutual benefits in the form of infrastructure sharing and knowledge flows, technology development and transfer, complementary and common solutions for shared problems, etc., research collaboration has also been associated with higher research productivity and impact. There are several previous studies that tried to measure and analyze international research collaboration for different countries and regions, and in the process developed different indicators and formalisms. However, there is no well-defined indicator to quantify the possible impact of international research collaboration on research output and citations of an institution. Recently, a set of boost indicators was introduced to reflect the effect of collaboration on productivity, impact, etc., of countries. This paper explores the possibility of adopting the boost formalism at an institutional level. The formalism is deployed and evaluated on research output data of 1000 Indian institutions. Different boost indicators are computed and validated through correlation studies. Results indicate that the proposed boost formalism can act as a suitable measure for assessing the possible impact of international research collaboration on research output and citations of institutions.

Introduction

Research collaboration is often defined as a group of researchers working together to solve complex scientific problems (Katz & Martin, 1997). It has also been defined as a social phenomenon where researchers pool their knowledge, experience, skills, and technology, intending to produce new scientific knowledge (Bozeman & Boardman, 2014). Research collaboration provides researchers with numerous mutual advantages in the form of knowledge transfer and training of researchers, resource sharing, access to complex and costly equipment, infrastructure, expansion and diversification of research network, funding etc. (Katz & Martin, 1997, Beaver, 2001; Birnholtz, 2007; D'Ippolito & Ruling, 2019). With the ICT revolution, the distances to interactions and collaborations have decreased, and as a result, the research collaboration is now transcending institutional and geographical boundaries. Several studies have analyzed collaboration at the international level and have observed that it has risen linearly during the last two 2-3 decades, as measured in terms of the number of internationally co-authored papers published (Glanzel, 2001; Persson, Glänzel & Danell, 2004; Lee & Bozeman, 2005; Wagner & Leydesdorff, 2005; Leydesdorff & Wagner, 2008; Mattsson et al., 2008; Adams, 2012). Considering the benefits of International Research Collaboration (IRC), policymakers of different countries see it as a valuable tool and are designing various programs to foster such collaboration (Katz & Martin, 1997; Wagner et al., 2001; Boekholt et al., 2009).

Some studies have postulated that scientific collaboration is strongly associated with research productivity and economic growth, along with a significant impact on citation (Glänzel, 2001; Abramo, DÁngelo, & Solazzi, 2011; Abramo, DÁngelo & Murgia, 2017; Inglesi-Lotz & Pouris, 2013; Ntuli et al., 2015). Many previous studies have focused their attention on measuring and characterising IRC trends, patterns, and impacts in different countries. Various indicators to measure the association strength in terms of propensity, intensity, and affinity in international collaboration have been proposed. Initially, the key focus of the analysis of IRC was the size of the country and related geographical, socioeconomic, and historical factors (Price, 1969; Frame & Carpenter, 1979) shaping research collaboration. As the research work on IRC grew, some indicators like the 'cooperation index' (Schubert & Braun, 1990) and 'exclusive strategy' (Luukkonen, Persson & Sivertsen, 1993) were introduced. The weighted affinity index was introduced thereafter (Leclerc & Gagné, 1994) to weigh the measured links between two countries based on the observed/expected ratio. For calculating absolute strength between pairs of countries, the Salton measure was proposed (Schubert & Braun, 1990; Glanzel, 2001). In addition to the affinity index, similarity measures such as cosine similarity (van Eck & Waltman 2009), inclusion index (van Eck & Waltman 2009; Luukkonen, 1993), Jaccard similarity (van Eck & Waltman 2009; Luukkonen, 1993), and multilateral similarity (Goodman's quasi-independence) (Luukkonen, 1993) were applied to bibliometric data. Three major algorithms have been proposed to define the Probability Affinity Index (PAI), namely non-overlapping (Leclerc and Gagné, 1994), overlapping (Zitt et al, 2000), and self-exclusive methods (Luukkonen, Persson & Sivertse, 1992; Schubert & Glänzel, 2006). The partnership probability index was developed by Yamashita & Okubo (2006) and was applied in combination with PAI as the Salton-Ochiai index on inter-sectoral organizational collaboration. Recently, some variants of the relative intensity of collaboration were studied by Fuchs, Sivertsen & Rousseau (2021).

Though many of the previous studies proposed indices to measure and characterize international research collaboration, there has not been a development towards a suitable indicator to measure the impact of international research collaboration on the research output and citations of an institution, a country, or any other actor in the scientific research landscape. There lies the research gap that this study attempts to bridge. The study proposes a Boost formalism consisting of different boost indicators that can be used to measure what effect or impact the international research collaboration may have on the productivity (research output is the proxy taken) or impact (citations are the proxy taken) of an institution. The formalism is described in detail, and thereafter its applicability in an institutional context is demonstrated on research publication data of 1000 Indian institutions. The suitability and relevance of boost indicators are evaluated. Finally, the usefulness, applicability, and further extension possibilities of formalism are discussed.

Related work

The investigation of international research collaboration (IRC) through co-authorship patterns began with efforts to characterize the interaction between the scientific output of a nation and its large-scale determinants. Price's (1969) contribution was a path-breaking effort in this regard, where he analyzed the correlations between a nation's scientific activities and socioeconomic determinants such as economic scale and technological capability. This initial effort brought to the forefront the role of national resources in shaping the dynamics of scientific collaboration. Frame and Carpenter (1979) took these findings further by examining the 1973 Science Citation Index (SCI) data, which included over 100 subfields categorized into nine scientific fields across 167 countries. The study found a positive correlation between a nation's scientific capability, measured by publication output, and internationally co-authored publications, indicating that larger scientific communities engage more in global collaborations. These early studies formed the foundation for systematic methodologies in IRC measurement, with the role of national scale in shaping collaboration behavior being a central theme.
To measure collaboration strength and trends, researchers have come up with various indicators to quantify IRC. Schubert and Braun (1990) came up with the cooperative index, a percentage difference between actual international co-authorships and expected values, adjusted for country size. The index made it possible to compare collaboration tendencies between countries, taking into consideration differences in scientific output. They also used Salton's measure (Salton & McGill, 1983), which measures the relative intensity of co-authorship relationships between countries. The measure was used by Glänzel and Schubert (2001) to analyze collaboration between 36 countries. Though useful for symmetric collaboration patterns, Salton's measure is difficult to use to capture asymmetric relationships, where a country dominates the partnership, and thus is of limited use in various collaboration scenarios. Luukkonen, Persson, and Sivertsen (1992) responded to size-dependency with the Probabilistic Affinity Index (PAI), which attempts to quantify collaboration strength regardless of country size. PAI cross-checks actual co-authorships against expected ones, and values above 1 represent stronger-than-expected collaboration. PAI, however, overestimates the importance of countries with skewed collaboration distributions and those with dominant partners. To counteract this, Schubert and Glänzel (2006) created the preference index of co-authorship, which is an enhancement on PAI in the sense that it accounts for specific country collaboration preferences and removes size effects. This index generates a more advanced measure of bilateral scientific connections, reflecting the country's affinity more precisely. Luukkonen et al. (1993) also suggested other measures of collaboration intensity, such as bilateral similarity

measures (e.g., Jaccard, Salton), multilateral similarity by Goodman's quasi-independence model, and multidimensional scaling for graphical representation of IRC networks. Such methods, though pioneering, remain size-dependent, overestimating the contribution of large countries compared to small ones, which makes equitable comparisons difficult.

Later studies developed new indicators to overcome earlier limitations. Leclerc and Gagné (1994) developed the proximity index (PRI), a quantifier of the strength of collaboration against the number of co-authored outputs. The PRI is aimed at symmetric relationships between nations, with greater values signifying stronger collaborative relations; however, its focus on symmetry limits its use. Zitt, Bassecoulard, and Okubo (2000) developed a publication-level probabilistic affinity index, in contrast to the co-authorship-level PAI, to measure the strength of collaboration between five major scientific nations: France, Germany, Japan, the UK, and the USA. Their approach overcame the impact of self-co-authorship through iterative margin recalibrations, thus ensuring a fair assessment of international relations. Yamashita and Okubo (2006) examined inter-sectoral collaboration between France and Japan through the combination of PAI with Salton's measure, a modification of the Ochiai coefficient (Ochiai, 1957; Zhou & Leydesdorff, 2016). They also developed the Probabilistic Partnership Index (PPI), measuring the infrequency of observed partnership links against predicted distributions. The PPI complements the PAI by identifying the statistical significance of partnerships, thus introducing a new dimension to collaboration processes.

Recent advances have focused on improving IRC measures to address contemporary challenges. Fuchs, Sivertsen, and Rousseau (2021) introduced the Relative Intensity of Collaboration (RIC), an improvement over earlier asymmetric indices, such as Luukkonen's PAI, which failed to capture relative increases in co-authored papers (Rousseau, 2021). RIC provides a robust measure of collaboration intensity by considering total collaboration volumes and pairwise interactions, thus improving its performance in asymmetric cases. Chinchilla-Rodriguez et al. (2021) explored differences in the use of PAI, such as differences in the handling of co-authorship matrix diagonals (e.g., setting to zero, as in Luukkonen et al., 1992; Leclerc & Gagné, 1994; Schubert & Glänzel, 2006; Fuchs et al., 2021) and normalization methods (Zitt et al., 2000; Yamashita & Okubo, 2006). These differences show the complexity of standardizing IRC measures across different research environments.

Counting methods have also been included in IRC analysis, providing authorship credit in collaborative research. Full counting provides equal credit to all authors, while fractional counting provides proportionate credit (Frandsen & Nicolaisen, 2010; Harsanyi, 1993; Lindsey, 1980; Waltman, 2016). Gauffriau (2017) outlined these approaches, highlighting their strengths and weaknesses in bibliometric studies. Most PAI-based analyses employed full counting, except for Leclerc and Gagné (1994) and Zitt et al. (2000), which explored fractional alternatives. Braun, Glänzel, and Schubert (1991) and Okubo, Miquel, Frigoletto, and Doré (1992) also dealt with the implications of counting methods for fair collaboration assessment.

As much as IRC indicators are prevalent across the world, there is an urgent gap: there is no measure among the current ones that reflects the impact of IRC on institutional productivity (publication output) or influence (citations). While country-level evidence has been useful, evidence at the institutional level is required to know how collaborations define research landscapes. This paper fills this gap by introducing a boost formalism—a collection of straightforward indicators to approximate the impact of IRC on institutional citations and publications. Using publication data from 1,000 Indian institutions, this framework offers a new approach to guide institutional strategies and policymaking, complementing traditional bibliometric measures.

Boost formalism: A discussion

Dua et al. (2023) introduced a set of indicators, viz. the boost indicators, to reflect the effect of collaborations on productivity, impact, etc., of countries. The idea of a boost in productivity and citation provides a way to quantify the impact of collaboration on productivity, citations, and altmetrics for different countries. The boost measures can be extended to the institutional context as follows:

Productivity boost **(β_P):** It can be defined as the ratio of the total number of publications (TP) to the total number of indigenous publications (TIP) of an institution, expressed in percentage. It can be expressed as follows,

$$\beta_P = \left[\frac{TP}{TIP} - 1\right] \times 100\ \%$$

The expression suggests that if an institution does not engage in collaboration, then $\beta_P = 0\ \%$. The value of β_P is directly proportional to the boost in productivity due to collaborations. A higher value of β_P indicates a higher reliance of the institution on international research collaboration. The ideal value of β_P is difficult to determine. As per the rule of thumb, if $\beta_P > 50\ \%$, then it indicates that the institution is more dependent on international collaboration than the indigenous ecosystem. On the other hand, if $\beta_P > 100\ \%$, it indicates that the institution is highly dependent on collaboration. If an institution has an infinite β_P (TIP=0 and a TP value of 1 or above), it signifies absolute dependence on collaboration.

Citations boost **(β_C):** It is defined as the ratio of total citations (TC) to the total citations received by indigenous publications (TIC) of an institution.

$$\beta_C = \left[\frac{TC}{TIC} - 1\right] \times 100\ \%$$

As per the rule of thumb, if $\beta_C > 50\ \%$, then it indicates the institution is more reliant/dependent on international collaborations for citation or impact than the indigenous scholarly system. On the other hand, if $\beta_C > 100\ \%$, it indicates that the institution is highly dependent. In other words, this indicates that the indigenous scholarly ecosystem is drawing very low relative impact and reach. Therefore, the institution should choose some impactful platforms or sources to disseminate its scientific research and improve the visibility of the indigenous scholarly research outputs.

***Boost ratio of impact per unit boost in productivity* (γ_c):** It is the net boost of citation per unit boost of productivity due to international research collaborations.

$$\gamma_c = \frac{\beta_C}{\beta_P}$$

If the value of $\gamma_c < 1$, international research collaborations are less rewarding and if $\gamma_c > 1$, such collaborations are rewarding. The benefit of research collaboration depends on the value of γ_c. This means the higher the value of γ_c, the greater the benefit of collaboration.

***Citedness boost* (β_{rc}):** It is the ratio of total citedness (total cited ratio) to the citedness ratio of the indigenous publications.

$$\beta_{rc} = \left[\frac{r_T}{r_{TI}} - 1\right] \times 100\ \%$$

where

$$r_T = \frac{total\ number\ of\ cited\ publications}{total\ number\ of\ publications} = \frac{TP_{cited}}{TP}$$

&

$$r_{TI} = \frac{total\ number\ of\ cited\ indigenous\ publications}{total\ number\ of\ indigenous\ publications} = \frac{TIP_{cited}}{TIP}$$

Citedness boost value greater than but close to 1 indicates that indigenous publications also have considerably good citedness. β_{rc} and β_c can be used together to determine whether an institution's indigenous works are making enough impact. β_{rc} value closer to 1 (like < 1 %), but considerably high β_c (like > 50 %) can indicate that despite the potential of indigenous works to gain citations, a considerable amount of work is remaining under-cited or not getting enough citations.

***Boost ratio of impact per unit boost in citedness* (δ_c):** It is the net boost of impact per unit boost of citedness due to international collaborations.

$$\delta_c = \frac{\beta_C}{\beta_{rc}}$$

The effectiveness of collaboration depends on the value of δ_c. The higher the value of δ_c, the higher the effectiveness of foreign collaboration. If the value of δ_c is very high with $\beta_{rc} < 1\ \%$, it indicates that the majority of collaboration is of good quality and rewarding as well. On the other hand, a high value of δ_c with $\beta_{rc} > 1\,\%$, indicates that there are some less rewarding collaborations. The reason for this could be that the collaboration can be a new tie or maybe the collaboration was formed long back but working on obsolete themes. Therefore, such collaboration should be reviewed to strengthen the collaboration by working on trending themes, to stop weaker ties and search for new ties or to minimize emphasis on such collaboration.

Demonstration of the Formalism

Data

In order to demonstrate the formalism of Boost in productivity and citations, research publication data for a large set of 1,000 Indian institutions collected from the Dimensions for an earlier work (Singh *et al.*, 2022) was used. The top 1000 Indian Institutions were selected on the basis of the total research output of those institutions during 2010-2019. The data comprised all document types and corresponded to the time period 2010 to 2019. The metadata fields that were accessed included the year of publication, DOI, citations, author(s) country affiliation, etc. The query formulated was as follows:

Search Query
search publications where year in [2010:2019] and research_orgs.id="{GRIDID}" and type in ["article"] return publications [research_org_countries+type+authors+year+abstract+open_access_categories_v2+research_orgs+authors_count+concepts_scores+field_citation_ratio+publisher+times_cited+altmetric_id+category_for+doi+title+category_sdg+journal+reference_ids+id+altmetric+issn+funder_countries+funders+relative_citation_ratio+supporting_grant_ids]

In the search query above, "GRIDID" corresponds to a unique ID assigned to each institution and these IDs for the top 1000 Indian Institutions were taken from the database. This was then passed one by one in the search query post which data for each of the Institutions was downloaded and processed.

Methodology

Post data download, different scientometric measures were computed by processing the appropriate metadata fields in the processed data. *Firstly*, the values of TP (total papers) and TC (total citations) were computed. TP was obtained from the total data count for each institution, while TC was obtained by summing up the values under the "times_cited" field for each institution. *Secondly*, in order to get the count of ICP (internationally collaborated papers) the "research_org_countries" field was investigated. This field contained the names of countries that collaborated to publish a record. Thus, for each institution, ICP comprised the total number of records that had more than country (India) listed in this field; while the records that had only one country (India) listed in this field comprised the share of TIP (total number of indigenous publications). Similarly, TIC (total number of indigenous citations) was obtained by summing up the values under the “times_cited” field that corresponded to only one country (India) listed in the “research_org_countries” field. *Thirdly*, for each institution, the computed values of productivity and citations were then used to compute the different boost indicators mentioned above. *Finally*, to better realise the nature of the different computed boost indicators, their values were correlated with the NIRF (National Institution Ranking Framework) ranks of each Indian Institution. A brief overview on NIRF is provided in the *Evaluation* section of this paper.

Results

The different boost indicators were computed for all the 1,000 institutions considered. The values for a set of 50 such institutions having high research output are presented in **Table 1**. The file containing the complete list of the 1,000 institutions considered, along with their relevant values and computations would be provided on request.

Table 1. Different Productivity Indicators of selected 50 Institutions.

S. No.	Institution Name	Acronym	TP	TIP	ICP	ICP %	TC	TIC	βp	βc	ϒc	βrc
1	Anna University, Chennai	AU Chennai	29698	25995	3703	12.47	316029	243145	14.245	29.976	2.104	1.918
2	All India Institute of Medical Sciences, Delhi	AIIMS Delhi	20545	17869	2676	13.03	225624	124171	14.976	81.704	5.456	2.498
3	Indian Institute of Science Bangalore	IISC	20257	15004	5253	25.93	308491	198107	35.011	55.719	1.591	1.516
4	Indian Institute of Technology Kharagpur	IIT KGP	18329	14621	3708	20.23	274172	200985	25.361	36.414	1.436	1.003
5	Indian Institute of Technology Bombay	IITB	17384	12861	4523	26.02	235472	149501	35.168	57.505	1.635	1.98
6	Indian Institute of Technology Madras	IITM	16650	12877	3773	22.66	207338	145541	29.3	42.46	1.449	1.415
7	Indian Institute of Technology Delhi	IITD	15402	12211	3191	20.72	232485	165499	26.132	40.475	1.549	1.114
8	University of Delhi	DU	15134	11288	3846	25.41	235350	123865	34.072	90.005	2.642	4.684
9	Bhabha Atomic Research Centre	BARC	13752	10443	3309	24.06	207521	121162	31.686	71.276	2.249	1.132
10	Post Graduate Institute of Medical Education and	PGIMER Chandigarh	13712	12224	1488	10.85	142646	91401	12.173	56.066	4.606	2.009

	Research, Chandigarh										
11	Vellore Institute of Technology University	VITU	12526	10072	2454	19.59	150495	107739	24.365	39.685	1.629	2.141
12	Jadavpur University	JU	12502	10216	2286	18.29	167053	120836	22.377	38.248	1.709	1.423
13	Indian Institute of Technology Roorkee	IITR	12470	10089	2381	19.09	207482	144100	23.6	43.985	1.864	0.756
14	Indian Institute of Technology Kanpur	IITK	11583	8776	2807	24.23	154648	102439	31.985	50.966	1.593	1.785
15	Indian Institute of Technology Guwahati	IITG	10364	8432	1932	18.64	146145	103298	22.913	41.479	1.81	1.252
16	Banaras Hindu University	BHU	10214	7959	2255	22.08	176674	117377	28.333	50.518	1.783	1.716
17	University of Calcutta	CU	9703	7850	1853	19.1	113064	81394	23.605	38.91	1.648	1.791
18	University of Pune	SPPU	9510	8046	1464	15.39	99732	72156	18.195	38.217	2.1	2.311
19	Visvesvaraya Technological University, Belgaum	VTU Belgaum	8959	7937	1022	11.41	64177	50249	12.876	27.718	2.153	2.058
20	Panjab University	PU	8469	5616	2853	33.69	171093	78320	50.801	118.454	2.332	3.538
21	Manipal Academy of Higher Education, Manipal	MAHE	8307	6575	1732	20.85	74606	48989	26.342	52.291	1.985	3.054
22	Aligarh Muslim University	AMU	8025	5744	2281	28.42	123656	73717	39.711	67.744	1.706	2.388
23	Maulana Azad National Institute of Technology, Bhopal	MANIT Bhopal	7866	6859	1007	12.8	107564	85646	14.681	25.591	1.743	0.824
24	University of Madras	UNOM	7017	5573	1444	20.58	96803	64557	25.911	49.95	1.928	2.616
25	University of Hyderabad	HCU	6651	5288	1363	20.49	91303	62711	25.775	45.593	1.769	2.999
26	Indian Institute of Chemical Technology, Hyderabad	IICT	6519	5485	1034	15.86	103422	75896	18.851	36.268	1.924	4.643
27	Jawaharlal Nehru University	JNU	6363	5068	1295	20.35	91933	54243	25.552	69.484	2.719	4.308
28	Amity University, Noida	AUUP	6325	5036	1289	20.38	57284	37570	25.596	52.473	2.05	2.663

29	Indian Institute of Technology (ISM) Dhanbad	ISM	6322	5552	770	12.18	78416	64986	13.869	20.666	1.49	0.977
30	Bharathiar University, Coimbatore	BU Coimbatore	6194	4197	1997	32.24	88737	46165	47.582	92.217	1.938	4.166
31	Tata Institute of Fundamental Research	TIFR	6152	2564	3588	58.32	143789	26907	139.938	434.393	3.104	6.908
32	University of Kerala	UK	5834	5058	776	13.3	54766	42393	15.342	29.186	1.902	1.993
33	Annamalai University	AU Tamil Nadu	5376	4447	929	17.28	80872	63250	20.89	27.861	1.334	1.443
34	Christian Medical College & Hospital, Vellore	CMCH Vellore	5334	4067	1267	23.75	56797	24848	31.153	128.578	4.127	4.76
35	Pondicherry University	Pondicherry University	5064	4257	807	15.94	60962	44344	18.957	37.475	1.977	2.271
36	King George's Medical University, Lucknow	KGMU Lucknow	5050	4622	428	8.48	37116	29030	9.26	27.854	3.008	1.713
37	Thapar University, Patiala	TIET Patiala	4987	4245	742	14.88	74715	56080	17.479	33.229	1.901	0.975
38	Bharathidasan University	Bharathidasan University	4954	3577	1377	27.8	72566	45478	38.496	59.563	1.547	2.058
39	Jamia Milia Islamia	JMI	4923	3569	1354	27.5	77521	50081	37.938	54.791	1.444	3.066
40	National Institute of Technology Rourkela	NITR	4897	4363	534	10.9	60716	48926	12.239	24.098	1.969	0.448
41	Birla Institute of Technology and Science, Pilani	BITS Pilani	4774	3913	861	18.04	55937	39727	22.004	40.803	1.854	1.292
42	Indian Statistical Institute, Kolkata	ISI Kolkata	4751	3124	1627	34.25	52270	29198	52.081	79.019	1.517	2.104
43	Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow	SGPGI Lucknow	4652	4201	451	9.69	70416	33205	10.736	112.064	10.439	2.218
44	National Chemical Laboratory, Pune	NCL Pune	4598	3794	804	17.49	92440	65842	21.191	40.397	1.906	1.997
45	Indian Association for the Cultivation of	IACS Kolkata	4477	3501	976	21.8	83081	61666	27.878	34.727	1.246	1.131

	Science, Kolkata											
46	Indian Institute of Engineering Science and Technology, Shibpur	IIEST Shibpur	4438	3838	600	13.52	44685	36046	15.633	23.967	1.533	0.992
47	National Institute of Mental Health and Neurosciences, Bengaluru	NIMHANS	4416	3648	768	17.39	45522	28897	21.053	57.532	2.733	2.068
48	National Institute of Technology Tiruchirappalli	NIT-T	4353	3642	711	16.33	59008	46617	19.522	26.58	1.362	1.816
49	West Bengal University of Technology, Kolkata	MAKAUT WB	4351	3694	657	15.1	41509	32230	17.786	28.79	1.619	1.577
50	Amrita Vishwa Vidyapeetham University	AMRITA	4052	3479	573	14.14	41548	27121	16.47	53.195	3.23	1.816

Note: **TP**-> Total Publications, **TIP**-> Total Indigenous Publications, **TC**-> Total Citations, **TIC**-> Total Indigenous Citations, **ICP**-> Internationally Collaborated Publications (ICP=TP-TIP, ICP%=(ICP/TP)*100).

From **Table 1**, it can be observed that among the top 50 productive institutions, except for the Tata Institute for Fundamental Research, all institutions have βp values <50%, indicating a self-reliant research ecosystem. Also, 9 Institutions have βp values <15% which indicates that the institutions have achieved a much higher productivity boost through domestic publications without much need for collaborations which is also seen owing to the fact that their Internationally collaborated publications (ICP=TP-TIP) comprise a share of <15% of their Total Publications (TP). These institutions are, namely, AU Chennai, AIIMS Delhi, PGIMER Chandigarh, VTU Belgaum, MANIT Bhopal, ISM Dhanbad, KGMU Lucknow, NITR and SGPGI Lucknow. It is to be noted that among the 7 IITs appearing in the top 50 list, IIT Gandhinagar (Rank 15), IIT Roorkee (Rank 13), IIT Kharagpur (Rank 4), IIT Delhi (Rank 7) and IIT Madras (Rank 6) have βp values <30% while IIT Kanpur (Rank 14) and IIT Bombay (Rank 5) display βp values of approx 32% and 35% respectively while they rank much higher in terms of TP. Moreover, the minimum value of βp is observed for King George Medical University (KGMU Lucknow, 9.26%) while the maximum value of βp is observed for Tata Institute for Fundamental Research (TIFR, 139.938). However, in terms of ranking by TP, KGMU (Rank 36, TP 5050) ranks lower than TIFR (Rank 31, TP 6152). According to the interpretation of βp values, this indicates that TIFR, despite having published a greater number of research publications than KGMU Lucknow, is more dependent on collaborative research than indigenous research, as the βp value for KGMU is >100%.

In terms of a Boost in productivity due to Citations, i.e. βc values, 28 Institutions have βc values <50%. A few of these are AU Chennai (~30%), IIT KGP (~36%), IIT Madras (42.5%), IIT Delhi (~40.48%), VITU (~40%), etc. Only 3 Institutions have βc values <25%, namely ISM (20.7%), IIEST Shibpur and NIT Rourkela (~24%). The maximum βc value is observed again for TIFR (434.39%), and the minimum is observed for ISM. It is to be noted that while TIFR (Rank 31, TP 6152) and ISM (Rank 29, TP 6322) differ marginally in terms of ranking due to TP, they lie on extreme ends of βc values. Thus, according to the interpretation of the βc values, the boost in citations achieved for TIFR is largely a result of its collaboration, while for ISM, it indicates a strong domestic research environment. As for the IITs appearing in the top 50 list, IIT Kanpur (50.97%) and IIT Bombay (57.5%) have βc values >50% while the other IITs like IIT Delhi (~40.48%), IIT Gandhinagar (~41.48%), IIT Madras (42.46%) and IIT Roorkee (~43.99%) have βc values <50%.

In terms of citedness boost i.e. βrc, 6 institutions (NITR, IITR, MANIT Bhopal, TIET Patiala, ISM and IIEST Shibpur), achieve values <1% which indicates impactful indigenous work by these institutions. These institutions also have βc values <50% which further supplements this finding. Among the IITs, it is seen that though IIT Roorkee (βrc=0.756, βc=43.98) has a lesser value of βrc than IIT KGP (βrc=1.003, βc=36.41) but has a higher value of βc than IIT KGP. On the other hand, IISC Bengaluru which ranks 3rd in terms of TP has both βrc=1.5% and βc=55.7% which indicates that both the boost in citations and the citedness boost are a result of collaborations. Here also, TIFR demonstrated the highest value of βrc i.e. 6.9%. Lastly, in terms of the Boost ratio of impact per unit boost in productivity (γ_c), almost all institutions have values > 1% which indicates that the international collaborations have been rewarding.

Cutoff values for βp (>50%, >100%), βc (>50%, >100%), and βrc (≈1%) were chosen using previous research patterns (Adams, 2012; Larivière et al., 2015). For βp >50% (TP = 1.5 × TIP, 33% of output) shows notable collaboration help, while >100% (TP = 2 × TIP) means heavy reliance. For βc >50% (TC = 1.5 × TIC) indicates collaboration boosts citations significantly. For βrc ≈1% means local and collaborative papers are cited similarly. **Table 1** shows KGMU's βp = 9.26% (self-reliant), TIFR's βp = 139.94% (TP = 2.4 × TIP), and IIT Roorkee's βrc = 0.756%. Figure 1's weak link ($R^2 = 0.0014$) supports these cutoffs.

To understand the relationship of the boost indicators with publication and citation counts, scatter plots of TP vs. βp, TC vs. βc, and TC vs. βrc are provided in **Figures 1**, **2** and **3**, respectively. **Figure 1** shows a very weak positive correlation ($R^2 = 0.0014$) between total publications (TP) and productivity boost (βp), suggesting that institutions with a high TP do not necessarily have a proportionally high βp. This implies that some institutions maintain strong indigenous publication ecosystems while others rely heavily on international collaborations. Notable institutions such as IISC, IITs, and AIIMS Delhi have a large TP but moderate βp, indicating a well-developed domestic research ecosystem. In contrast, institutions like ISI Kolkata, PU, and BU Coimbatore exhibit a high βp (>40%), signifying substantial reliance on international collaborations. **Figure 2** examines how total citations (TC) relate to βc, which quantifies the citation boost received due to international collaboration. A low R^2 value (0.0115) in the trend line suggests a weak correlation, indicating that while international collaboration generally increases citations, the extent of this boost varies significantly across institutions. Institutions with high TC but moderate βc, such as IITs and IISc, indicate that their indigenous research is also widely cited. On the other hand, CMCH Vellore, SGPGI Lucknow, and BU Coimbatore, with high βc values (>90), heavily rely on internationally collaborated research for citations, implying that their domestic publications receive comparatively lower impact. **Figure 3** explores how citedness boost (βrc) varies with total citations (TC). A weak negative correlation (R^2 = 0.0086) indicates that institutions with high TC do not necessarily have a high βrc, signifying that indigenous research in certain institutions is already well cited. DU, IICT, JNU, and BU Coimbatore show higher βrc values (>4), meaning their internationally collaborated publications receive significantly more citations per paper than indigenous ones. In contrast, institutions like IISc, IITs, and BARC have moderate βrc, suggesting a relatively balanced impact between international and domestic publications. Finally, **Figure 4** shows the strongest correlation (R^2 = 0.3162) between βp and βc, indicating that institutions that achieve higher productivity boosts through international collaborations tend to also receive proportionally higher citation boosts. This moderate positive correlation indicates that the advantages of international collaboration typically appear simultaneously in both increased productivity and citation impact, although the degree of effect differs significantly among institutions. Notably, institutions such as CMCH Vellore, PU, and BU Coimbatore exhibit particularly strong performance in both metrics.

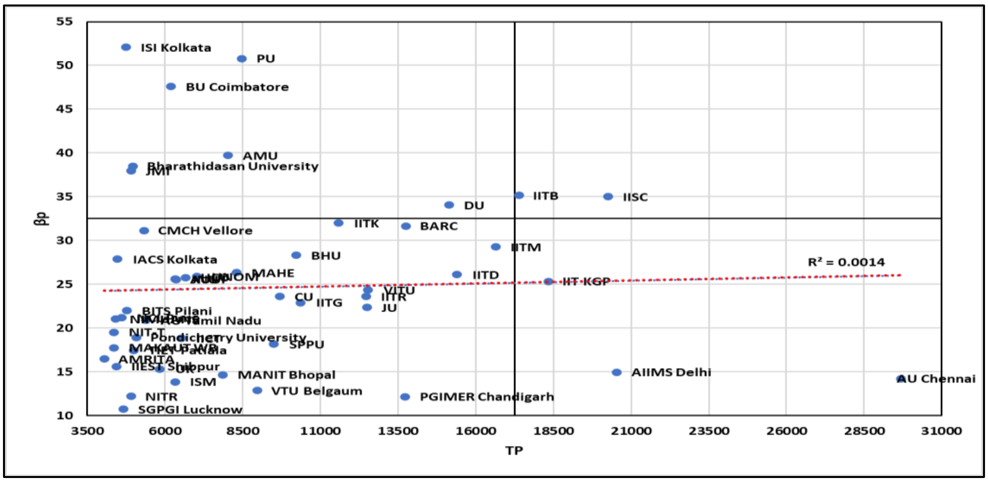

Figure 1. Boost in Productivity vs. Total Publications (TP) of selected 50 Institutions (excluding an outlier- TIFR).

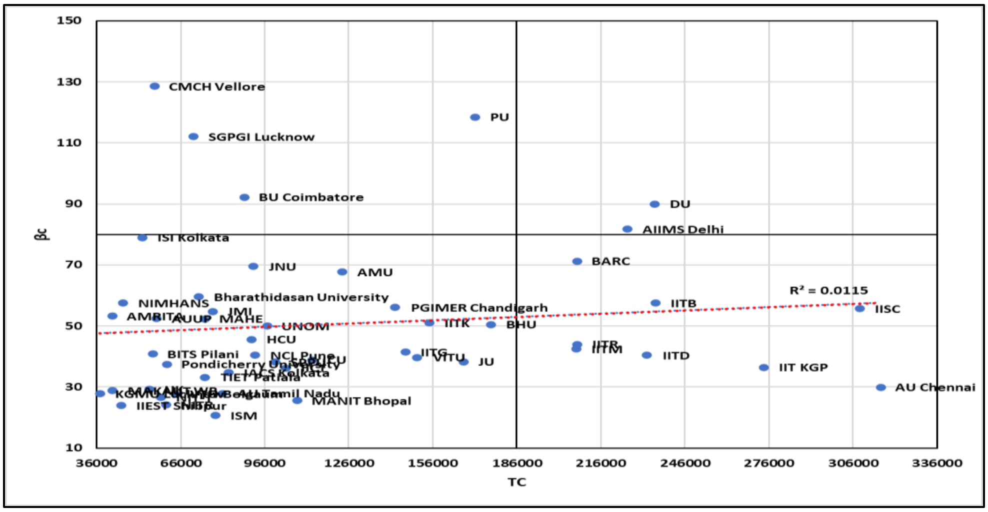

Figure 2. Boost in Citations vs. Total Citations (TC) of selected 50 Institutions (excluding an Outlier- TIFR).

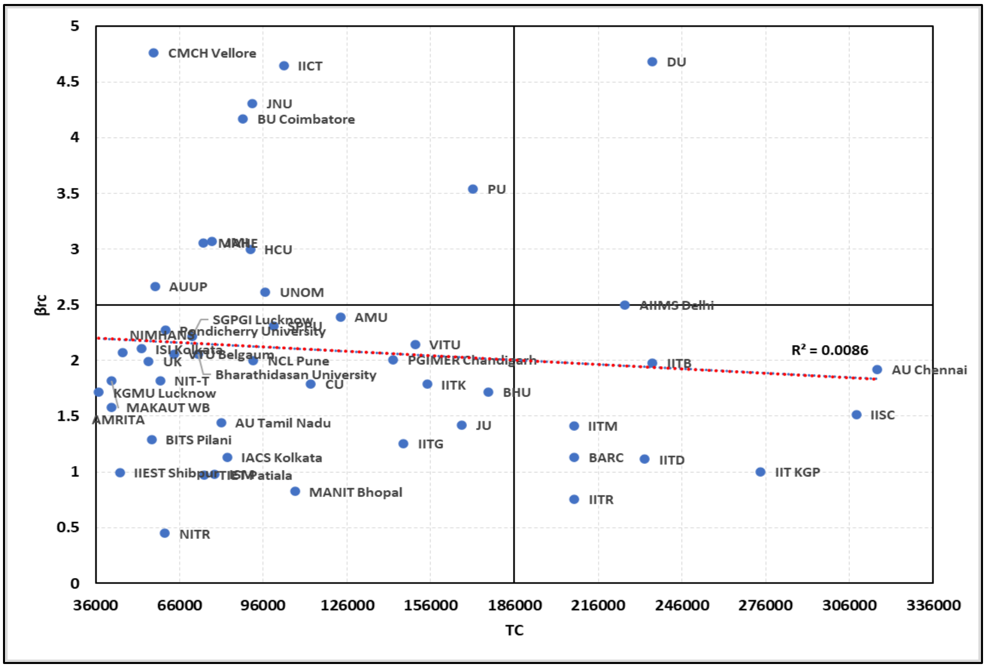

Figure 3. Boost in Citedness vs. Total Citations (TC) of selected 50 Institutions (excluding an Outlier- TIFR).

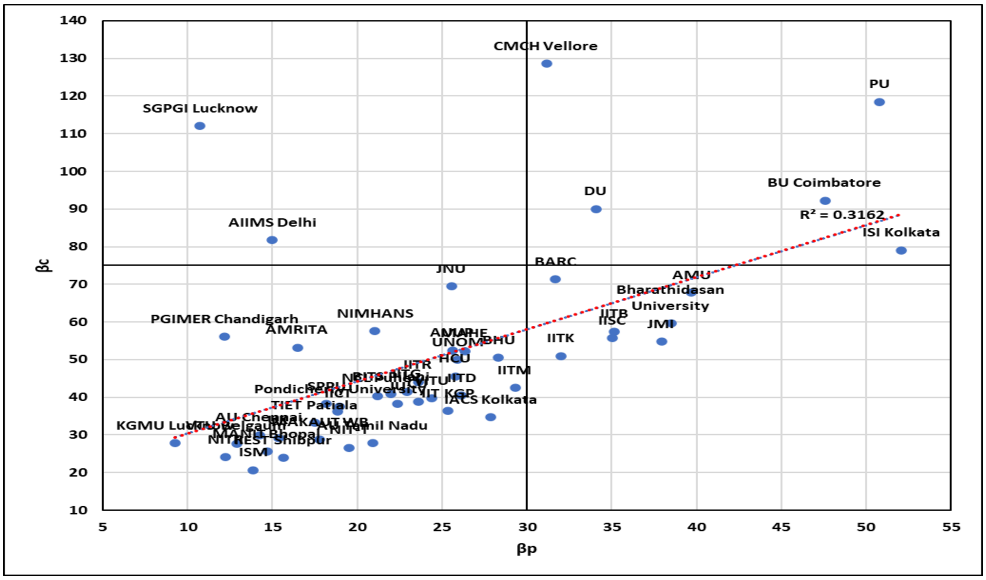

Figure 4. Boost in Citations vs. Boost in Productivity of selected 50 Institutions (excluding an outlier- TIFR).

Evaluation

In order to understand the nature and values of different boost indicators, the values of different boost indicators are correlated with a major national ranking of institutions in India, the NIRF. The National Institutional Ranking Framework (NIRF), established by India's Ministry of Education, ranks higher education institutions based on five weighted parameters: Teaching, Learning, and Resources (30%); Research and Professional Practices (30%); Graduation Outcomes (20%); Outreach and Inclusivity (10%); and Perception (10%). The research component, emphasising publication and citation metrics, aligns closely with the boost indicators' focus on productivity and impact, making NIRF a robust benchmark for validation.

The Spearman's Rank Correlation Coefficient (SRCC) was calculated to compare the rankings of higher education institutions by the NIRF with rankings by different productivity measures. The results from **Table 2** show a positive correlation with NIRF rankings and rankings by Total Publications (TP) and Total Citations (TC) at SRCC values of 0.67 and 0.68, respectively. The high correlation is due to the convergence of TP and TC with NIRF's high weighting of the research and professional practices criterion, which contributes 30% to the overall ranking methodology.

On the other hand, the correlations derived between NIRF rankings and the boost indicators—βp (0.27), βc (0.19), and βrc (0.03)—are significantly lower. These low correlation values indicate that these boost indicators might not be able to capture the complex nature of NIRF's ranking factors, which include Teaching, Learning, and Resources (30%), Graduation Outcomes (20%), Outreach and Inclusivity (10%), and Perception (10%).

This is explored further through **Figure 5**, which graphically verifies these findings. This is utilized to further support the argument that institutions with high publication and citation values rank better in NIRF's research-oriented evaluation. Conversely, the lower correlations of the boost indicators (βp, βc, and βrc) are evident from the less intense shading and the lower SRCC values, ranging from 0.03 to 0.27. This difference suggests that, while TP and TC are effective indicators of NIRF's focus on research productivity, the boost indicators might be measuring different aspects of institutional performance that are less aligned with NIRF's integrative approach.

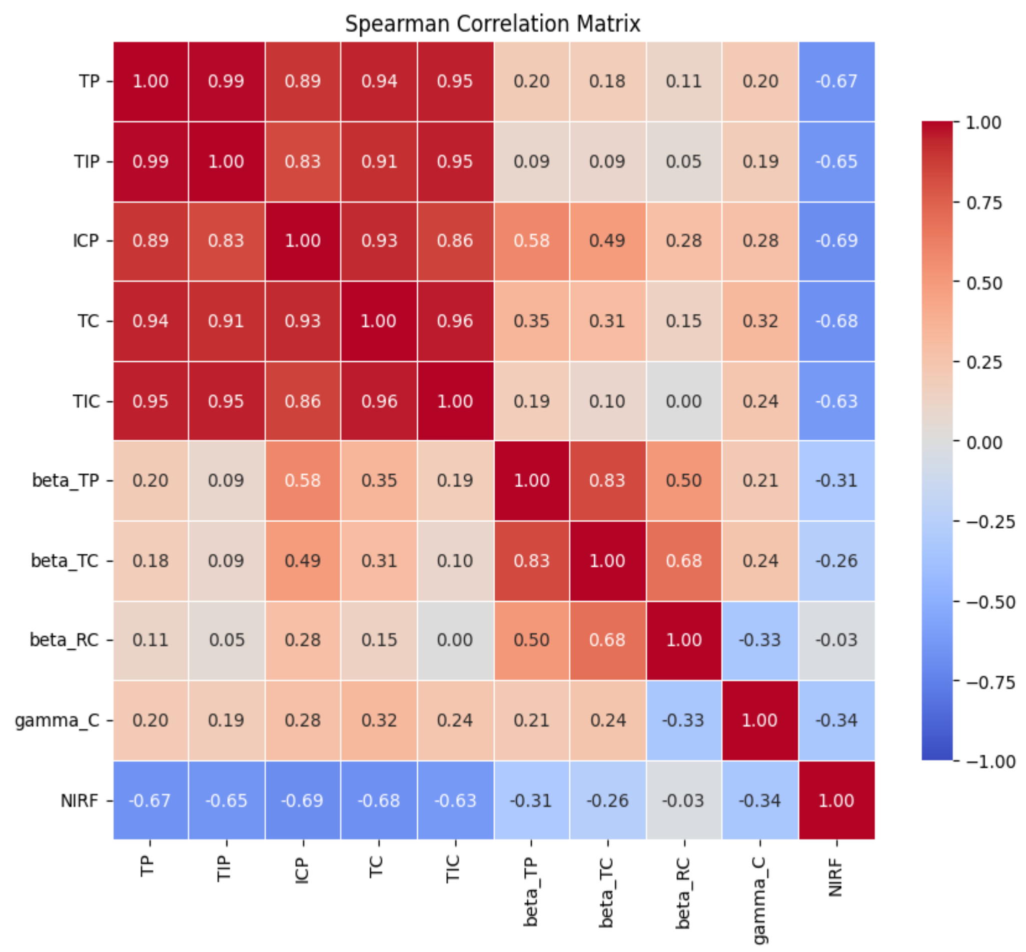

Figure 5. Spearman Correlation between all the parameters, boost parameters, and the NIRF rankings.

Table 2. Values of Spearman Rank Correlation Coefficients for different variables.

Variables	Value of SRCC
NIRF Ranking vs TP	0.67
NIRF Ranking vs TC	0.68
NIRF Ranking vs βp	0.27
NIRF Ranking vs βc	0.19
NIRF Ranking vs βrc	0.03

~~5.~~ Discussion

This study analyses research publication data from 1,000 Indian institutions to assess the impact of research Collaboration using the Boost formalism. By adopting the boost indicators – βp (productivity boost), βc (citation boost), βrc (citedness boost), and ɣc – to an institutional level, this work offers a novel approach to quantifying collaboration effects beyond traditional bibliometric measures like Total Publications (TP) and Total Citations (TC).

The findings of the study reveal that collaboration influences institutions differently. While IITs and AIIMS Delhi maintain strong indigenous research ecosystems with moderate βp values, institutions like TIFR exhibit high βp and βc, suggesting greater dependence on collaborations. Weak correlations between TP and βp, as well as TC and βc, indicate that high publication volume does not always correspond to significant collaborative impact. However, a moderate positive correlation between βp and βc suggests that well-integrated collaborations enhance both productivity and citation impact.

The study has practical implications for institutional research profiling, academic planning, and policymaking. This is especially important because though (i) national scholarly ranking initiatives like NIRF provide a sense about their relative performance (ii) recently proposed indicators like x and x_d provides an idea about the scholarly research portfolio of institutions, these are not capable of providing an idea about the role and extent of influence collaborations have in determining the institutions' current stature. As a boost indicator, such as input, it can complement the information provided by NIRF and other useful indicators like x, x_d, and many others for institutions to plan their way forward and shape their research policy and formulate strategies. Institutions with high βc but moderate βp benefit from selective, high-impact partnerships, while those with high βp but low βrc may need to improve the visibility of their indigenous research. Policymakers can use these insights to allocate resources and design policies that foster meaningful collaboration. The profile of collaboration's impact on institutions highlights its dual role in enhancing productivity and global visibility. Institutions with strong international ties often gain access to cutting-edge knowledge, advanced methodologies, and prestigious networks, contributing to their academic standing. Additionally, IRC enables researchers to tackle complex, multidisciplinary problems requiring diverse expertise, boosting institutional research output and reputation. Such collaborations also help institutions attract better funding, international faculty, and students, creating a virtuous cycle of growth and recognition in the global academic landscape.

Despite its contributions, the study has certain limitations. The analysis remains correlational, making it difficult to establish causal links between IRC and research performance. Future research could incorporate longitudinal studies and subject-specific analyses to refine these metrics further. Additionally, examining external factors such as funding, institutional size, and subject area specializations could improve our understanding of collaboration dynamics. Further, the work has demonstrated computation of the proposed indicators on data downloaded from Dimensions database, a major reason being the larger coverage of Dimensions database (Singh et al., 2021). However, these values for the institutions may vary if data from a different database is used. In this sense, the proposed indicators, like all the bibliometric indicators in existence, are also sensitive to the database used, and indicator quality will also be related to the quality of the database.

By introducing a structured framework to evaluate collaboration's impact, this work provides a valuable perspective on institutional research productivity. The boost formalism offers a scalable and robust model for assessing the effectiveness of international collaborations, guiding institutions and policymakers toward data-driven research strategies.

Declarations

Funding and/or Conflicts of interests/Competing interests: The authors declare that no funding was received for this work. Further, the manuscript complies with the ethical standards of the conference and there is no conflict of interest whatsoever.

References

Abramo, G., DÁngelo, A.C. & Murgia, G. (2017). The relationship among research productivity, research collaboration, and their determinants. *Journal of Informetrics*, 11(4), 1016-1030.

Abramo, G., DÁngelo, A.C. & Solazzi, M. (2011). The relationship between scientists' research performance and the degree of internationalization of their research. *Scientometrics*, 86(3), 629-643.

Adams, J. (2012). "Collaborations: The rise of research networks." *Nature*, 490(7420), 335–337. DOI: 10.1038/490335a.

Braun, T., Glänzel, W., & Schubert, A. (1991). The bibliometric assessment of UK scientific performance—some comments on Martin's "reply". *Scientometrics, 20*(2), 359-362.

Beaver, D. D. (2001). Reflections on scientific collaboration, (and its study): past, present, and future. *Scientometrics, 52*(3), 365-377. doi:10.1023/a:1014254214337

Birnholtz, J. P. (2007). When do researchers collaborate? Toward a model of collaboration propensity. *Journal of the American Society for Information Science and Technology*, 58(14), 2226-2239.doi:10.1002/asi.20684

Boekholt, P., Edler, J., Cunningham, P., & Flanagan, K. (2009). Drivers of international collaboration in research. *European Commission*. https://doi.org/10.2777/81914
Bozeman, B., & Boardman, C. (2014). *Research collaboration and team science: A state-of-the-art review and agenda* (Vol. 17). New York: Springer.
Chinchilla-Rodríguez, Z., Bu, Y., Robinson-García, N., & Sugimoto, C. R. (2021). An empirical review of the different variants of the probabilistic affinity index as applied to scientific collaboration. *Scientometrics*, *126*, 1775-1795.
Davidson Frame, J., & Carpenter, M. P. (1979). International research collaboration. *Social studies of science*, *9*(4), 481-497.
D'Ippolito, B., & Ruling, C. C. (2019). Research collaboration in Large Scale Research Infrastructures: Collaboration types and policy implications. *Research Policy*, 48(5), 1282-1296. doi:10.1016/j.respol.2019.01.011
Dua, J., Singh, V. K., & Lathabai, H. H. (2023). Measuring and characterizing international collaboration patterns in Indian scientific research. Scientometrics, 128(9), 5081-5116.
Frandsen, T. F., & Nicolaisen, J. (2010). What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics, 4*(4), 608–617.
Fuchs, J. E., Sivertsen, G., & Rousseau, R. (2021). Measuring the relative intensity of collaboration within a network. *Scientometrics*, *126*(10), 8673-8682.
Gauffriau, M. (2017). A categorization of arguments for counting methods for publication and citation indicators. Journal of Informetrics, 11(3), 672-684.
Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, *51*(1), pp.69-115.
Glänzel, W., & Schubert, A. (2001). Double effort= double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, *50*(2), 199-214.
Harsanyi, M. A. (1993). Multiple authors, multiple problems—bibliometrics and the study of scholarly collaboration: A literature review. *Library and Information Science Research*, 15, 325–354
Inglesi-Lotz, R., & Pouris, A. (2013). The influence of scientific research output of academics on economic growth in South Africa: an autoregressive distributed lag (ARDL) application. *Scientometrics*, *95*, 129-139.
Katz, J.S. & Martin, B.R. (1997). What is research collaboration? *Research policy*, *26*(1), pp.1-18.
Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). "Team size matters: Collaboration and scientific impact since 1900." *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332. DOI: 10.1002/asi.23266.
Leclerc, M., & Gagné, J. (1994). International scientific cooperation: The continentalization of science. *Scientometrics*, *31*(3), 261-292.
Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, 35(5), 673-702.
Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics, 2*(4), 317–325.
Lindsey, D. (1980). Production and citation measures in the sociology of science: the problem of multiple authorship. Social Studies of Science, 10, 145–162
Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology, & Human Values*, *17*(1), 101-126.
Luukkonen, T., Tijssen, R., Persson, O., & Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, *28*(1), 15-36.
Mattsson, P., Laget, P., Nilsson, A., & Sundberg, C. J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, *75*(3), 555-574.

Ntuli, H., Inglesi-Lotz, R., Chang, T. Y., & Pouris, A. (2015). Does research output cause economic growth or vice versa? Evidence from 34 OECD countries. Journal of the Association for Information Science and Technology, 66(8), 1709–1716

Ochiai, A. (1957). Zoogeographical studies on the solenoid fishes found in Japan and its neighboring Regions - II. *Bulletin of the Japanese Society of Scientific Fisheries, 22*(9), 526-530.

Okubo, Y., Miquel, J. F., Frigoletto, L., & Doré, J. C. (1992). Structure of international collaboration in science: Typology of countries through multivariate techniques using a link indicator. *Scientometrics, 25*, 321-351.

Persson, O., Glänzel, W. & Danell, R. Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics* 60, 421–432 (2004). https://doi.org/10.1023/B:SCIE.0000034384.35498.7d

Price, D. J. de Solla. (1969). Measuring the size of science. *Proceedings of Israel Academy of Sciences and Humanities*

Rousseau, R. (2021). Going back in time: Understanding patterns of international scientific collaboration. *Journal of Scientometric Research, 10*(1), 126-129.

Salton G., McGill M.J., *Introduction to Modern Information Retrieval,* McGraw-Hill, N.Y. 1983

Schubert, A., & Braun, T. (1990). International collaboration in the sciences 1981–1985. *Scientometrics, 19*(1-2), 3-10.

Schubert, A., & Glänzel, W. (2006). Cross-national preference in co-authorship, references and citations. *Scientometrics, 69*, 409-428.

Singh, V. K., Nandy, A., Singh, P., Karmakar, M., Singh, A., Lathabai, H. H., ... & Kanaujia, A. (2022). Indian Science Reports: a web-based scientometric portal for mapping Indian research competencies at overall and institutional levels. *Scientometrics*, 127(7), 4227-4236.

Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. Scientometrics, 126, 5113-5142.

van Eck, N. J., & Waltman, L. (2009). How to normalize co-occurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology, 60*(8), 1635–1651.

Wagner, C. S., Brahmakulam, I., Jackson, B., Wong, A., & Yoda, T. (2001). Science and technology collaboration: Building capacity in developing countries. *Document Number MR-1357.0-WB, RAND Corporation, Santa Monica, CA.*

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, 34(10), 1608–1618.

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365-391

Yamashita, Y., & Okubo, Y. (2006). Patterns of scientific collaboration between Japan and France: Inter-sectoral analysis using Probabilistic Partnership Index (PPI). *Scientometrics, 68*(2), 303-324.

Zhou, Q., & Leydesdorff, L. (2016). The normalization of occurrence and Co-occurrence matrices in bibliometrics using Cosine similarities and O chiai coefficients. *Journal of the Association for Information Science and Technology, 67*(11), 2805-2814.

Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics, 47*(3), 627-657.