Mapping of Research Collaboration, Analysis of Literature Growth and Forecasting Research Progression of IITs

Hiteshkumar Harshadbhai Solanki

Miteshkumar Yashvadan Pandya

Aakansha

ABSTRACT-

The paper gives a holistic overview of growth in research papers, degree of collaboration and future publications of newly established IITs based on Scopus data. It aims to analyse the growth performance and efficiency of the IITs, overall research output by measuring relative growth rate, doubling time and collaboration index. The study reveals that the cumulative publications contributed by newly established IITs during 2010-2021 is 31393. The study also noted the persistence increase in the citations by 526671 against publications contributed by all newly established IITs. The study used Exponential Smoothing (ETS) algorithms as a univariate forecasting method to predict the number of publications that can be published altogether by the IITs in the next five years.

Keywords: Bibliometric Analysis, Doubling Time of Publications, Forecasting Research Growth, Relative Growth Rate, Trend Analysis.

Hiteshkumar Harshadbhai Solanki

Scientist - C (CS)
Information and Library Network Centre
Infocity,Gandhinagar 382007
hitesh@inflibnet.ac.in
ORCID: http://orcid.org/0000-0001-6151-7538

Miteshkumar Yashvadan Pandya

Deputy Librarian
National Forensic Sciences University
Sector-9, Gandhinagar - 382007
miteshkumar.pandya@nfsu.ac.in
ORCID: https://orcid.org/0000-0001-7592-7175

Aakansha

Library and Information Assistant
Indian Institute of Technology Indore
Khandwa Rd, Simrol,
Madhya Pradesh 453552
aakansha@iiti.ac.in

INTRODUCTION

Research is carried out to prove the existence of something and evaluation of research is done to improve the quality of research. Knowledge generation shows the progress of research, and its evaluation helps to generate valuable information for taking appropriate decisions. Research remains discipline-oriented, and it may lead towards interdisciplinary and collaborative research within the research community. The collaborative research contributes towards narrowing the gap in the procedural methodologies followed while conducting research. Collaborative research would prove to be an alliance among the national and international institutions. Focusing on research conducted by the IITs helps to identify areas of substantive interest of the researchers. This will increase the knowledge base and create intellectual standards among the researchers.

Technical advancements in various fields of knowledge have shown phenomenal growth in the past two decades. Teaching and research are essential for stable & progressive overarching growth of the country.

Growth can be achieved by conducting impactful research in various fields that can help in improving social life. The Union Government has focused on technical education and research in the recent past. As a result of this, to foster technical education & research, the Union Government has established new IITs across the country. The growth and development of an academic institute can be achieved by enhancing the research base in collaboration with academics and industries. The research should be focused on the exploration of problems in society and industry. It is important to evaluate and examine the research to curtail the academic, administrative, and technical problems at a regular interval.

Growth of research publication and improvement in the citation count is one of the criteria to evaluate research in the given domain. Growth in various aspects plays a key role in the overarching development of an institute. It shows substantial efforts in achieving the predefined goals and targets. Growth in the research publication shows the progress of individual researchers as well as an institution.

Nowadays, there is an exponential growth in research publications in various fields. Many predatory journals also exist in the publishing industry that take advantage of authors who are trying to publish an article and looking for the journals that are charging fees by eliminating peer-review or editing procedures. Due to technological advancements, the research in technical and scientific publications has grown manifold. Based on the research publications trend for the last eleven years, an attempt has been made to forecast the growth in research publications.

The forecasting research growth of the newly established IITs using bibliometric data would help institutes to set their research goals by presuming future growth in research publications.

REVIEW OF LITERATURE

Qualitative evaluation of publication patterns could be one of the ways for measuring the impact of research work. Mahapatra (1994) conducted a study on finding out correlation between the growth of publications and citations. The study analyzes three types of growth curves namely exponential, linear, and logistics patterns of relative growth and doubling time for both articles and citations. The results revealed that with an increasing number of publications, the citations will also increase sufficiently. The results also revealed that the citation may be used as standard to study growth of literature in the subject. Analysis to find out the growth of literature on the basis of growth rate analysis was done by Sundararajan and Ponnudurai (2011). An attempt has been made to study the growth of literature using growth rate analysis for the literature published in marine biology. The study considered 25 years of bibliometric data for the evaluation of growth and doubling time of the literature published in Marine Biology of fauna. The results revealed that there was a phenomenal increase in the number of publications at the end of the year 2008. However, the relative growth rate has shown a declining trend that decreased gradually from 0.57 in the year 1980 to 0.01 in the year 2004. Baby and Kumaravel (2012) conducted a study to find out the pattern of the relative growth rate and priority index of the Journal of Clinical Microbiology for a period of five years. The results revealed that the doubling time for publications is increasing. The subject Bacteriology showed the highest number of articles and an increasing trend in collaborative research was noticed.

Priya & Ponnudurai, (2011) analysed Relative Growth Rate and Doubling Time of articles published in conference on neural network research across the world in general and in India specific. The results reveled that at the global level conference articles share is 49.88 percent whereas in India remains on the second position with 41.88 percent. The study also revealed that the doubling time of publications on neural network research output increased and recorded at 4.78 years for the world whereas doubling time of publications for India is 2.89 years. A study predicting research impact using Citation Time Series data has been conducted by Jiang et al., (2021). Researchers have given the HINTS (Heterogeneous Information Network to Time Series) framework for predicting the impact of new publications.

The review of literature has presented various aspects of mapping research using Relative Growth Rate for different disciplines and institutions. The literature survey reveals that mapping research performance is not a new process and was adopted by individual researchers for journals and subjects. It is noted for the review that the coverage and scope of different studies varies. This literature review depicts that the considerable studies have been published on RGR, DT impact of research papers and prediction. However, there is no similar study carried out for predicting or forecasting the growth of research publications. Also scope and coverage of different study varies.

OBJECTIVES OF THE STUDY

The growth and developmen t of an academic institute can be achieved by enhancing the research base in collaboration with academics and industries. The present study contrives to analyze the growth of research on the basis of relative growth rate, doubling time and other collaborative parameters. The major objectives of the study are as follows:

- To analyze the growth performance and efficiency of the IITs;
- To find out year-wise Doubling Time (DT);
- Degree of Collaboration (DC), Collaboration Index (CI), Collaboration Coefficient (CC);
- The relative growth rate of publications and citations;
- Year-wise annual growth rate and compound average growth rate; and
- Prediction in the growth of research work for the next five years.

SCOPE AND LIMITATION OF THE STUDY

The present study is an attempt to find out year wise Doubling Time (DT), degree of collaboration, the relative growth rate of publications and citations, year-wise annual growth rate, and compound average growth rate of the IITs established during 2009-10. The study is limited to only eight IITs namely Indian Institute of Technology Gandhinagar, Indian Institute of Technology Bhubaneswar, Indian Institute of Technology Indore, Indian Institute of Technology Jodhpur, Indian Institute of Technology Hyderabad, Indian Institute of Technology Patna, Indian Institute of Technology Ropar, and Indian Institute of Technology Mandi. The data of newly established IITs who have completed 12 years of their existence in teaching and research have only been considered for the study. The period of the study is twelve years i.e., 2010-2021 and data have been retrieved using affiliation id from the Scopus database on November 30, 2022. The old institutions that were elevated as IITs during the period 2009-10 have not been considered in the present study.

METHODOLOGY

The bibliometric data extracted from Scopus was tabulated in MS-Excel. Various well-established formulae have been applied to find out results keeping in view the objectives of the study. Mahapatra (1994) has devised formulae for calculating Relative Growth Rate (RGR) and Doubling Time (DT) of publications and citations. The same formula has been used to find out RGR and DT for publications and citations.

$$\bar{R}(P) = \frac{\log_e 2^P - \log_e 1^P}{2^p - 1^p}$$

Where,

$\bar{R}(P)$ is the relative growth rate of publications for a specific time interval.

$\log_e 1^P$ is the natural logarithm of initial number of publications

$\log_e 2^P$ is the natural logarithm of final number of publications

$$DT(P) = \frac{\log_e 2}{\bar{R}(P)} = \frac{0.693}{\bar{R}(P)}$$

Where, DT(P) is the Doubling Time of publication growth

$$\bar{R}(C) = \frac{\log_e 2^c - \log_e 1^c}{2^c - 1^c}$$

Table 1: Year Wise Growth of Publication and Relative Growth Rate

Year	Publications	Cumulative Publications	Annual Growth Rate (in %)	Log (W1P)	Log(W2P)	RGR(P)	Mean RGR (P)	DT(P)	Mean DT(P)
2010	216	216	-	-	5.38	-	-	-	-
2011	458	674	212.04	5.38	6.51	1.13	0.67	0.61	1.17
2012	764	1438	113.35	6.51	7.27	0.76		0.91	
2013	1137	2575	79.07	7.27	7.85	0.58		1.19	
2014	1697	4272	65.9	7.85	8.36	0.51		1.36	
2015	2043	6315	47.82	8.36	8.75	0.39		1.78	
2016	2425	8740	38.4	8.75	9.08	0.33	0.28	2.1	2.52
2017	3075	11815	35.18	9.08	9.38	0.30		2.31	
2018	4003	15818	33.88	9.38	9.67	0.29		2.39	
2019	4510	20328	28.51	9.67	9.92	0.25		2.77	
2020	5203	25531	25.6	9.92	10.15	0.23		3.01	
2021	5862	31393	22.96	10.15	10.35	0.20	-	3.47	-

Where,

$\overline{R}(C)$ is the relative growth rate of citations received against publications over the specific time interval,

$\log_e 1^C$ is the natural logarithm of citations received for the publications at the start of the year

$\log_e 2^C$ is the natural logarithm of citations received for the publications at the end of the year

$$DT(C) = \frac{\log_e 2}{\overline{R}(C)} = \frac{0.693}{\overline{R}(C)}$$

Where,

DT(C) is the Doubling Time of citation growth against publications

ANALYSIS AND INTERPRETATION OF DATA

The publications and citations data retrieved from Scopus database have been plotted into excel for the purpose of calculating the annual growth rate, relative growth rate and doubling time. Table 1 contains year wise growth of publications, relative growth rate and doubling time of publications whereas table 2 contains year wise growth of citations, relative growth rate and doubling time of citations received by publications. Publications are considered to be one of the instruments for showcasing the research productivity of IITs. Table 1 shows year wise publications of newly established IITs with annual growth, relative growth rate and doubling time of the publications. It may be inferred from the data that the publications are increasing steadily over the years. However, the cumulative annual growth rate of publications is fluctuating.

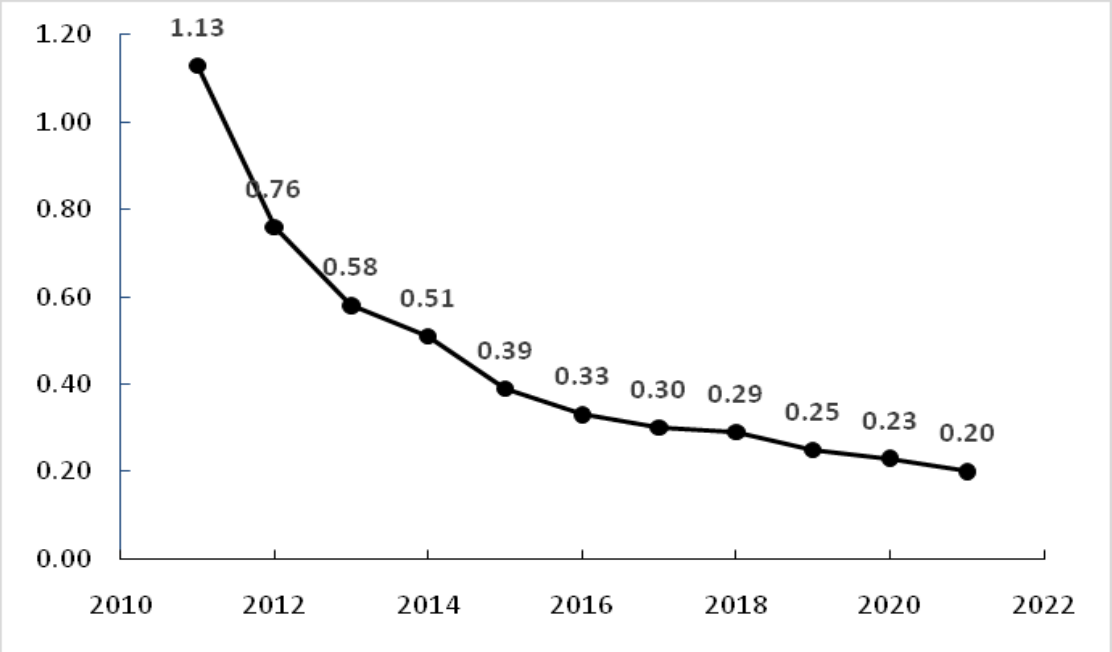

Figure 1: Relative Growth Rate of Publications

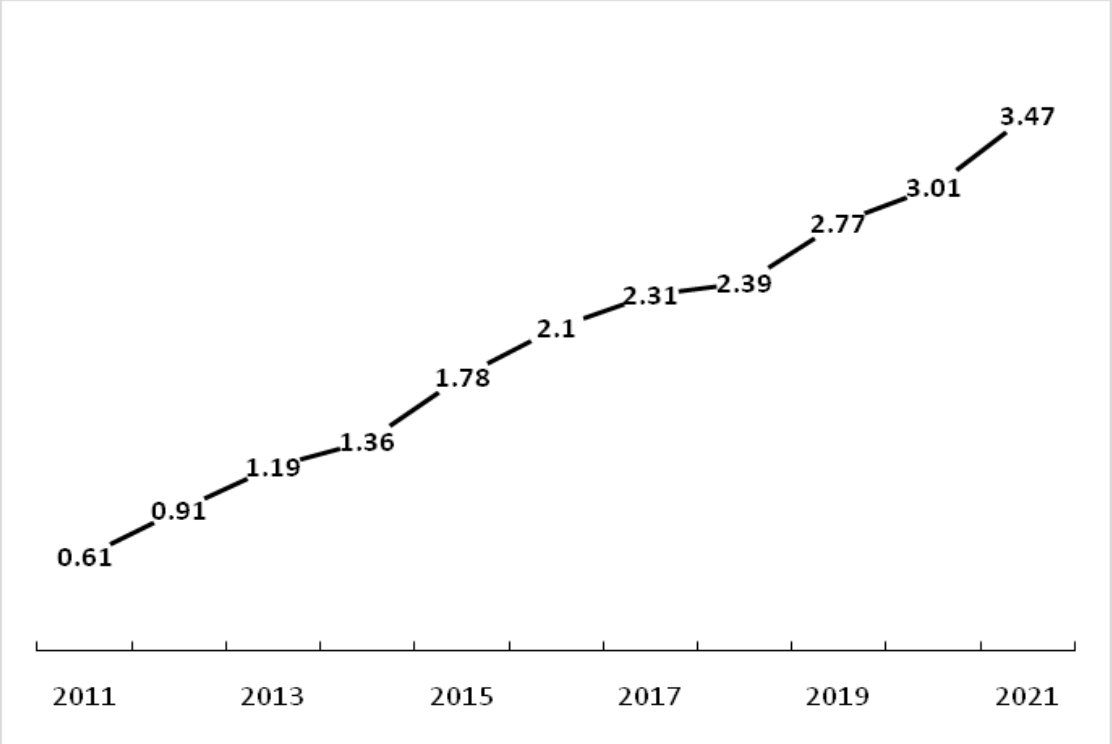

Figure 2: Doubling Time of Publications

Table 2: Year Wise Growth in Citations and Relative Growth Rate

Year	Citations	Cumulative Citations	Annual Growth Rate (in %)	Log (W1C)	Log (W2C)	RGR (C)	Mean RGR (C)	DT (C)	Mean DT (C)
2010	4842	4842	-	-	8.49	-	-	-	-
2011	10574	15416	218.38	8.49	9.64	1.15	0.68	0.6	1.2
2012	17627	33043	114.34	9.64	10.41	0.77		0.9	
2013	32098	65141	97.14	10.41	11.08	0.67		1.03	
2014	35805	100946	54.97	11.08	11.52	0.44		1.58	
2015	44574	145520	44.16	11.52	11.89	0.37		1.87	
2016	75920	221440	52.17	11.89	12.31	0.42	0.24	1.65	3.7
2017	97213	318653	43.9	12.31	12.67	0.36		1.93	
2018	64900	383553	20.37	12.67	12.86	0.19		3.65	
2019	59769	443322	15.58	12.86	13	0.14		4.95	
2020	49614	492936	11.19	13	13.11	0.11		6.3	
2021	33735	526671	6.84	13.11	13.17	0.06	-	11.55	-

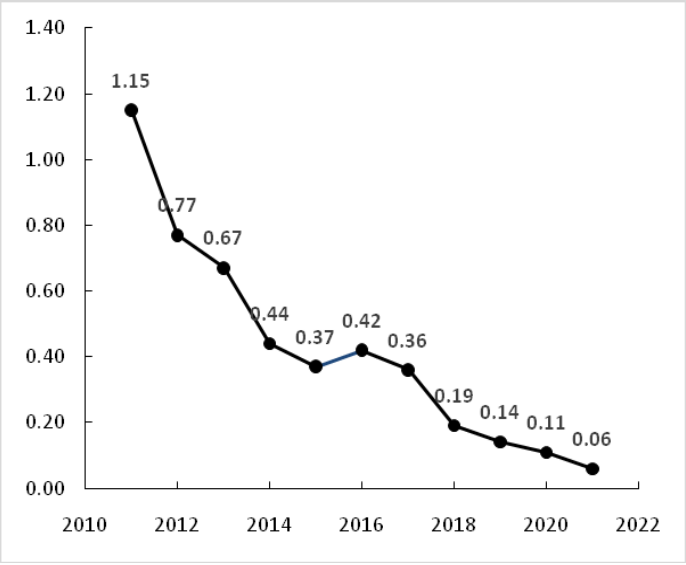

Figure 3: Relative Growth Rate of Citations

Table 1 shows the cumulative annual growth rate of publication for the newly established IITs during the period 2010-2021. The highest growth 212.04 was shown in 2011 followed by 113.35 in 2012. It may be noted that the annual growth rate fluctuates during the study period and does not record in negative. The mean doubling time of publications during the 2011-2015 is 1.17 that is lesser as compared to the last five-year slab i.e., 2016-2020 which is recorded at 2.52. The annual growth rate of publications has been calculated using the following formula.

$$AGR(P) = \frac{2^P - 1^P}{1^P} x\ 100$$

Where,
AGR(P) is the Annual Growth Rate of publications,

1^P is the initial number of publications over specific time interval,

2^P is the final number of publications over specific time interval

$$CAGR = \left(\frac{V_{final}}{V_{begin}} \right)^{1/t} - 1$$

Where,

CAGR= Compound Annual Growth Rate

V_{begin} = Beginning Value

V_{final} = Final values

T= Time in years

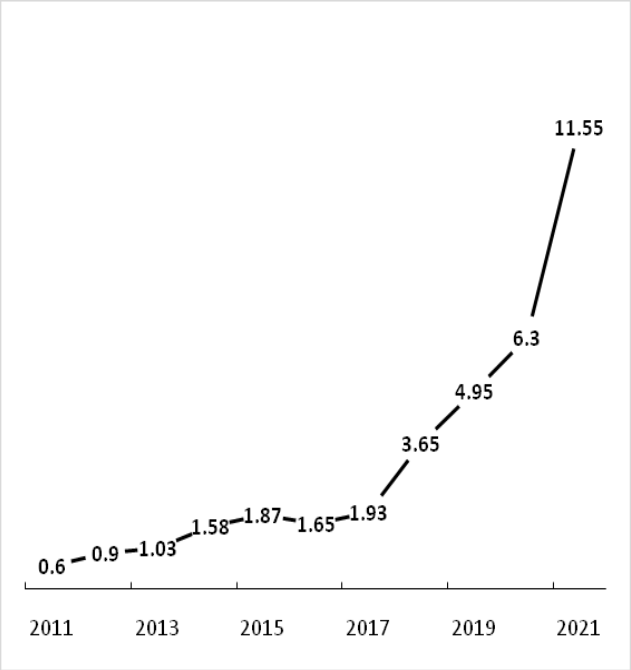

Figure 4: Doubling Time of Citations

Compound Annual Growth Rate (CAGR) is the mean annual growth rate of publications over a specified period of time for more than a year. It is one of the most accurate ways to determine the ascent or descent in publications over time. CAGR of publications for the newly established IITs recorded at 31.66 % over a period of 12 years.

Relative growth rate is generally calculated to find out the increase in the publications and citations over a period. Table 1 portrays the relative growth rate and doubling time of publications of newly established IITs for the period of twelve years i.e., 2010-2021. The highest relative growth rate 1.13 is noted in the year 2011 followed by 0.76 in the year 2012. As depicted in figure 1, the relative growth of publications is decreasing exponentially over the period of first five-year slab. The mean relative growth rate of publications during the 2011-2015 is 0.67 that is higher as compared to the last five-year slab i.e., 2016-2020 which is recorded at 0.28. The highest 3.47 doubling time is recorded in the year 2021 growth followed by 3.01 in the year 2020. As shown in figure 2, the doubling time of publications increases constantly as compared to previous year.

The calculation of doubling time of publications is inversely proportional to the relative growth rate since it is wholly dependent on the relative growth rate. Table 2 shows the annual growth rate of citations received for the newly established IITs during the period 2010-2021. The highest annual growth rate 218.38 was shown in 2011 followed by 114.34 in 2012. It may be inferred that the annual growth rate oscillates during the study period and reported in reduction during the last five years of study period i.e., 2017-2021.

As shown in figure 3, the relative growth rate of citations decreases over the years except 2016 as compared to previous years. The mean relative growth rate during the first five-year block i.e., 2011-2015 is 0.68 and second five-year block i.e., 2016-2020 is recorded at 0.24 which is lesser than first five-year block. It may be noted that the mean doubling time of citations is increasing, and the mean relative growth rate of citations is decreasing. The reason behind decreasing trend in RGR of citations is due to insufficient number of citations received after publication in the same year.

The calculation of doubling time of citations received is inversely proportional to the relative growth rate since it is wholly dependent on the relative growth rate. Compound Annual Growth Rate (CAGR) can also be computed to find out the pattern of citations received by the institutions over a period. It is one of the most precise ways to determine the growth or fall in the citations of publications over the period. The CAGR of the citations was recorded at 17.56% over a period of

Table 3: Institute Wise Publication Collaboration Index

Institute Name	1	2	3	4	5	6	7	8	9	10	> 10	Total	Multi-Authored Publications	Collaboration Index (CI)	Degree of Collaboration (DC)	Collaboration Coefficient (CC)
Indian Institute of Technology (IIT) Gandhinagar	152	850	773	522	311	180	108	69	32	26	213	3236	3084	4.05	0.95	0.66
Indian Institute of Technology (IIT) Bhubaneshwar	80	873	994	609	297	183	105	46	20	11	529	3747	3667	4.55	0.98	0.69
Indian Institute of Technology (IIT) Indore	157	1159	1158	997	601	398	208	134	61	48	440	5361	5204	4.43	0.97	0.69
Indian Institute of Technology (IIT) Jodhpur	74	540	587	456	263	153	77	44	30	18	43	2285	2211	3.85	0.97	0.67
Indian Institute of Technology (IIT) Hyderabad	226	1629	1605	1134	649	352	165	108	53	37	442	6400	6174	4.05	0.96	0.67
Indian Institute of Technology (IIT) Patna	107	1281	1257	726	332	134	83	34	17	16	52	4039	3932	3.35	0.97	0.64
Indian Institute of Technology (IIT) Ropar	126	877	826	630	300	166	103	42	46	34	106	3256	3130	3.79	0.96	0.66
Indian Institute of Technology (IIT) Mandi	83	754	899	540	315	185	104	71	46	28	44	3069	2986	3.78	0.97	0.67

12 years which shows an increase in cumulative citation received by publications of newly established IITs.

Collaboration Coefficient and Degree of Collaboration

Authorship pattern is one of the significant parameters to understand/identify collaboration of authors at various level. The table 3 shows the authorship pattern of the institute wise publications and collaboration index. The collaboration coefficient and degree of collaboration was proposed by Ajiferuke et al., (1988).

$$CI(P) = \frac{\left(\sum_{i=1}^{10} i \times P^i\right) + \left(\sum_{i>10}^{n} 11 \times P^i\right)}{\sum_{i=1}^{n} P^i}$$

Where,

CI(P) is the collaboration index of the publication

P^i is the publication contributed by *i* author(s)

$$DC(P) = \frac{\sum_{i=2}^{n} P^i}{\sum_{i=1}^{n} P^i}$$

Where,

DC(P) is the degree of collaboration of the publications

Table 4: Forecasting Growth of Research Publications for the years 2022-2026

Year	Publications	Forecast (Publications)	Lower Confidence Bound (Publications)	Upper Confidence Bound (Publications)
2010	215	-	-	-
2011	454	-	-	-
2012	753	-	-	-
2013	1129	-	-	-
2014	1693	-	-	-
2015	2027	-	-	-
2016	2395	-	-	-
2017	3053	-	-	-
2018	3940	-	-	-
2019	4445	-	-	-
2020	5048	-	-	-
2021	5862	5862	5862	5862
2022	-	6523	6122	6923
2023	-	7184	6325	8043
2024	-	7845	6433	9257
2025	-	8506	6460	10552
2026	-	9167	6415	11919

P^i is the publication contributed by i author(s)

$$CC(P) = 1 - \left(\frac{\sum_{i=1}^{n} \frac{P^i}{i}}{\sum_{i=1}^{n} P^i} \right)$$

Where,

CC(P) is the collaboration coefficient of the publications

P^i is the publication contributed by *i* author(s)

Table 3 shows institution wise publication collaboration index. Indian Institute of Technology (IIT) Hyderabad has published 6400 articles and Indian Institute of Technology (IIT) Jodhpur has published the lowest 2285 articles during the period of study. It may be inferred from the data that most of the publications are collaborative in nature in all the IITs. IIT Hyderabad has published 6174 collaborative articles out of total 6400 articles. Indian Institute of Technology (IIT) Bhubaneshwar has the highest collaboration index of 4.55 and Indian Institute of Technology (IIT) Patna has the lowest collaborative index of 3.35. Highest degree of collaboration is reported at 0.98 in Indian Institute of Technology (IIT) Bhubaneshwar.

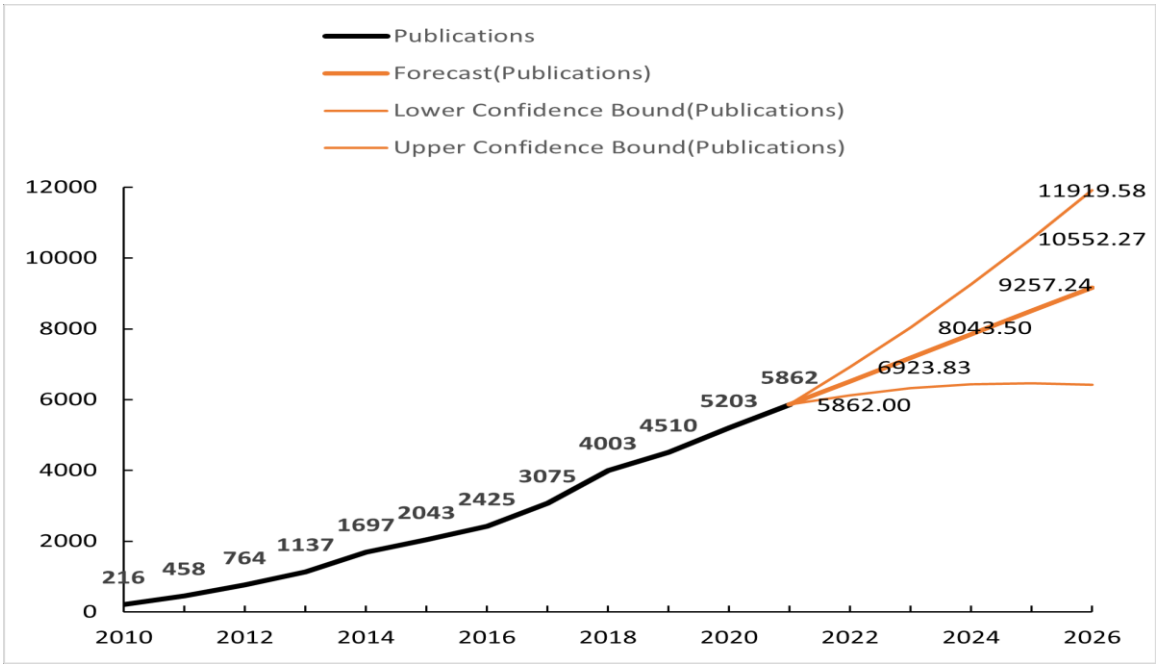

Figure 5: Forecasting Growth of Research Publication for the years 2021-2026

Lowest degree of collaboration reported at 0.95 in Indian Institute of Technology (IIT) Gandhinagar. Highest collaboration coefficients have been noted for Indian Institute of Technology (IIT) Bhubaneshwar and Indian Institute of Technology (IIT) Indore at 0.69 and the lowest 0.64 have been noted for Indian Institute of Technology (IIT) Patna.

Forecasting Research Progression using Bibliometric Data

The bibliometric data retrieved from Scopus have been tabulated in the MS- Excel and method of least square was applied to calculate the forecast of research publications based on data. Exponential Smoothing (ETS) algorithms have been used as a univariate forecasting method to predict the number of publications that can be published by the newly established IITs in the next five years. Forecast is a method of estimation that gives 95% confidence interval using ETS algorithm and the upper confidence bound is the estimation of maximum number as well as lower confidence bound is the minimum number of publications that will be expected during the period of next five years.

As shown in the figure 5 above, the trend of research publications in the newly established IITs is increasing constantly. The upper confidence bound, and the lower confidence bound have been calculated with 95% of confidence interval. The lower confidence bound shows estimated publications of IITs will have a minimum 6523, 7184, 7845, 8506, 9167 publications for the upcoming five years i.e., between 2022 and 2026 respectively. Whereas the upper confidence bound shows estimated publications of IITs will have a maximum 6923, 8043, 9257, 10552, 11919 publications for the upcoming five years i.e., between 2022 and 2026 respectively.

RESULTS AND DISCUSSIONS

- The cumulative publications contributed by all considered newly established IITs during 2010-2021 is 31393, so on an average approximately 2616 publications per annum are contributed by all newly established IITs together.
- The cumulative citations received against publications contributed by all considered newly established IITs during 2010-2021 is

526671, precisely it is noted that approximately 17 citations received per publication is achieved by newly established IITs.

- The relative growth rate of publications as well as citations is drastically decreased over the years.
- The collaboration index of the IITs varies from 3.35 to 4.55 for the period of study.
- The degree of collaboration remains between 0.95 to 0.98 that shows the all the IITs under study are close to each other in collaboration.
- The collaboration coefficient ranges between 0.64 to 0.69 which shows authorship pattern of all the IITs have significant collaboration among peers.
- The growth of research publications is increased exponentially for the considered study period and the forecasting research publications for the newly established IITs would be linear in the next five years.

REFERENCES

1. Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. Scientometrics, 14(5), 421–433. h*ttps://doi.org/10.1007/BF02017100*
2. Baby, K., & Kumaravel, J. P. S. (2012). Relative growth rate and priority index of Journal of Clinical Microbiology during 2006-2010. International Journal of Library and Information Studies, 2(2), 1–9.
3. Jiang, S., Koch, B., & Sun, Y. (2021). HINTS: Citation Time Series Prediction for New Publications via Dynamic Heterogeneous Information Network Embedding. Proceedings of the Web Conference 2021, 3158–3167. *https://doi.org/10.1145/3442381.3450107*
4. Mahapatra, G. (1994). Correlation between growth of publications and citations: A Study based on growth curves. Annals of Library Science and Documentation, 41(1), 8–12.
5. Priya, N., & Ponnudurai, R. (2011). Studies on relative growth rate and doubling time of conference articles productivity of neural network research. International Journal of Development Research, 1(4), 1–4.
6. Sundarajan, D., & Ponnudurai, R. (2011). Bibliometric study of marine biology publication output at international level. International Journal of Recent Scientific Research, 2(11), 276–278.