Evaluating the Effectiveness and Performance of Research of Degree Awarding Institutes of Pakistan: By Using Parametric and non-Parametric Stochastic Dominance Approach

Mamoona Anam
Faculty of Computing and Information Technology, Department of Computer Science
International Islamic University, Islamabad Pakistan 192122 mamoona.anam.vt@iiu.edu.pk

Jamal Abdul Nasir
Department of Computer Sciences National University of Ireland H91TK33 Jamal.nasir@nuigalway.ie

Malik Sikander Hayyat Khayal
Department of Computer Science, Preston University, Islamabad Pakistan 192122
drsikandarhayat@preston.edu.pk

Abstract: One of the key academic performance measures for research is the volume of publications and their standardized effect. An international trend towards using bibliometrics study evaluation aids decision-making by policymakers and research administrators. However, due to the assumptions and constraints of bibliometric data compared to the conventional methodologies, a non-parametric methodology is more appropriate for evaluating the research performance. This study assesses researchers' performance using parametric and non-parametric stochastic dominance approaches. The study covers the period from 1992 to 2023, consisting of 126 researchers from the top five public universities in Islamabad, Pakistan named Allama Iqbal Open University, International Islamic University, Quaid-e-Azam University, Comsat, and National University of Science and Technology. The study considers the researchers' publications who worked in two distinct fields: mathematics and computer sciences. Moreover, this study provides crucial insights for different disciplines of researchers by contrasting the advancements in the specialized areas with broader management literature using a variety of alternative methodologies. Researchers and policymakers can use this study as a useful resource to comprehend and survey possible advancements.

Keywords: bibliometrics, stochastic dominance, publications, h-index, productivity, citations, researchers

1. Introduction

Over time higher education governance emphasized quality research output. Therefore, more control mechanisms have been established to assess the research performance of the researchers. Due to performance evaluation, there is a high tendency to foster competitive research markets and outwardly alter the foundations in the academic field (Hamann, 2016). The assessment of research performance through bibliometrics is of considerable importance. With the drastic advancement of computer technology in the 1960s, bibliometrics instruments were created based on different statistical and mathematical models for information mining and literary analysis(Ho, Mantello, Nguyen, & Vuong, 2021). Bibliometric instruments such as Pajek, UNINET, Cite Space, and VOS viewer have made significant importance to analyze the current research status and technology to understand future 'hot' research directions, grasp the direction of the discipline, and clarify the evolution of a discipline. More than ten different types of bibliometric instruments have been created so far and are internationally available(Gao et al., 2019).Among them, Cite Space is the most frequently used program by domestic researchers. It has the advantages of supporting various data types, having a wide range of capabilities, and having good visualization effects(Zhu, Lin, & Li, 2022).By offering a reliable way to visualize and analyze the overall research activity, development dynamics, and development trends of a discipline (Liu et al., 2023), bibliometrics makes up for the shortcomings of traditional literature review research, such as low reference volume, use of only qualitative induction and analysis, conclusions and recommendations that are heavily influenced by researchers' knowledge, and weak objectivity.

Assessment of research performance can be done for multiple purposes, for instance, informing the organizational strategies, research output enhancement, and making policy decision-making processes etc. Research performance assessment is more important for research administrators and policymakers. It should come as no surprise that various bibliometric indicators have increased recently, frequently with additional modifications and more complex computation techniques. It is suggested that a non-parametric approach should be used for measuring research performance along with the parametric one because of the bibliometrics' application assumptions and limitations, as well as those of the various indicators, both alone and in combination, in study evaluation (Abramo, Costa, & D'Angelo, 2015).

Nowadays, researchers' assessment extends beyond merely quantifying publications and citations. In addition to these traditional metrics, a wide range of bibliometric indicators is employed, as highlighted byRafols and Stirling (2021). This approach enables a more comprehensive understanding of scientific accomplishment's multifaceted nature. Onodera and Yoshikane (2015) propose utilizing a variety of metrics encompassing eight different measures, including but not limited to the average annual publication count and cumulative citation counts. According toSahel (2011),individual metrics possess distinct advantages and may potentially offset the drawbacks associated with other metrics. The comprehensive evaluation of research performance necessitates the consideration of multiple metrics.

Nevertheless, it is important to acknowledge that numerous metrics selected for a study exhibit a strong correlation with one anotherdespite potential variations in outcomes at a more granular level(Abramo, D'Angelo, & Di Costa, 2023; Bornmann & Williams, 2017). Hence, the metrics employed in an evaluation study must be designed in such a way as to minimize the occurrence of redundant outcomes.Despite bibliometrics being a distinct area of study for a considerable period, there remains a lack of consensus regarding the standardized application of bibliometric analyses to individual scholars(Sahel, 2011).This study also aims to establish recommendations for the meaningful evaluation of individual researchers in formal sciences and engineering. This study is deemed essential within this specific domain. Assessing the performance of individual scientists is a crucial aspect of evaluating research, and the results of these evaluations can significantly influence institutional research strategies, encompassing funding initiatives, personnel recruitment and termination, and promotions. The focus of our study pertains to the process of selecting data to conduct an evaluation, the subsequent analysis of the collected data, and how the results are presented. The study has been restricted to the fundamental tool and non-parametric methodology. This implies that we selectively suggest options for evaluation based on their perceived necessity and significance, considering the wide range of available choices.

The current study's core objective is to examine individual researchers' research performance at the departmental level by applying parametric and non-parametric stochastic dominance (SD) approaches. However, the existing studies documented the research performance analysis at the aggregate level, for instance, research group levels. Therefore, the current study documents the performance assessment at the departmental levelof researchers. To our knowledge, no study has been found on assessing researcher performance by considering the assessment on the departmental level for Pakistan.

This study's significance is determining how to distribute the limited resources and which institutions or areas of research need to be focused on. The current study also has importance froma policy-making point of view. For instance, academic institutions can use the RPA to inform policydecisions related to financing, promotions, research directions, policy development,collaborations and quality control. It also encourages superior research, supports decisions, and improves the academic atmosphere.

The following section provides the existing literature to explain the current core factor in research performance measurement. The next section, Methods and Data, explains the considered performance indicators the studyuses and details the SD approach with its first, second and third orders. After that, the next section, named Results and Discussion, documents the core findings from the study with the help of tables and graphs. The final section serves as the concluding segment of the study.

2. Indicators in Research Performance Assessment

The assessment of research performance is a very challenging task when several indicators are considered. Several key factors create uncertainty in research performance assessment. The list includes subjective evaluation, time lag, subjective evaluation, Metrics and Indicators, Multi- and interdisciplinarity, Data Accessibility, Research Ethics and Reproducibility etc. Research having multiple outputs is one of the main areas where quantifying performance is challenging. Any new knowledge that is not codified cannot be measured by bibliometrics, and even when new information is codified, bibliometrics still struggles to recognize and quantify its different forms(Abramo et al., 2015).

According to Moed's research from 2005, the most popular way to formalize research output in the so-called "hard sciences" is through publication in academic journals. Databases like Scopus and Web of Science (WoS) are recognized as sufficiently transparent about their content and coverage due to their extensive use and testing in bibliometric investigations. As a result, bibliometrics merely use papers included in WoS or Scopus as a proxy for overall hard sciences output. The immediate outcome is that those not counted outputs will eventually be ignored.

While the humanities, arts, and a sizable percentage of the social sciences reject this approximation, the hard sciences do. Publications that use new information have various values depending on how they advance knowledge. Contrary to limitations on the indicator, such as "network" citations and negative citations, bibliometrics employs the number of publications' citations as a proxy for this influence(Torres-Salinas, Robinson-García, & Gorraiz, 2017).

Bibliometricians employ a scaling factor derived from the citation distribution of publications within the same year and subject area to standardize the citations of each publication. This is necessary due to diverse citation standards across different domains. Several scaling factors, including measures such as the average, median, and z-score of normalized distributions, have been suggested and utilized to normalize citations in the field. Few scholars have suggested normalizing citations by the number of bibliographic references of the citing paper since interdisciplinary work may easily suffer in the evaluation from being lost in a categorical classification system(Kurtz & Henneken, 2017; Leydesdorff, Bornmann, Mutz, & Opthof, 2011). The complexity of interdisciplinary work can be attributed to the challenges posed by categorical classification systems. No parametric field-normalization method is likely to completely overlap the citation distributions in the case of a significant number of fields(Tang, Wan, & Zhang, 2014).Furthermore, evaluation exercises must include a comparison of academics within the same field. This is crucial to mitigate any potential distortions in performance rankings that may arise from variations in publication intensity across different academic disciplines.

In bibliometrics, the main focus is on the quantitative analysis of academic publications and citations. Bibliometrics is an effective tool for assessing research performance and its impact on scientific work. Some important key aspects of bibliometrics assist in research performance assessment. For instance, publication count: in this aspect, bibliometrics facilitates the publications count, which can be used to predict the research productivity and can be helpful in research output assessment. Secondly, through bibliometrics, citation analysis is a good indicator of research impact or indicates that the work has significantly impacted the academic community. The third main aspect of bibliometrics is the h-index. This popular bibliometric measure that combines citation analysis and publication counts is the h-index. It evaluates the effectiveness and productivity of a researcher's efforts. A researcher who has written h articles that have all received at least h citations is said to have an h-index of h.The fourth is the evaluation of journal impact factors is also included in the journal impact factor as bibliometrics. This metric measures the number of citations obtained by articles published in a specific publication over time. It is common practice to assess the value and reputation of scholarly papers using journal impact factors.To thoroughly evaluate research performance, bibliometrics should also be used with the non-parametric technique, such as the SD technique.

In conclusion, bibliometrics has become a potent instrument for evaluating the quality of research. It offers quantitative measures to assess research productivity, impact, and visibility by analyzing publication and citation data. However, to thoroughly evaluate research excellence, it is necessary to utilize bibliometric indicators cautiously, considering their limitations and supplementing them with other evaluation techniques.

Further restrictions and approximations are believed to apply to each bibliometric indicator. The new crown indication, the h-index, is the most well-liked performance indicator thatAbramo et al. (2015) examine the advantages and disadvantages of. When selecting indicators and measurement techniques, it is essential to consider the objectives and context of the assessment exercise. Still, most bibliometrics agrees that a combination of indicators must be used to evaluate the research performance rather than a single instrument.

It is inappropriate to recommend a single performance metric due to the diversity of objectives among research organizations and over time. However, this does not justify the widespread use of multiple indicators. The present research also presents a collection of five bibliometric indicators. Two of them are in areas where the co-authors' diverse contributions are acknowledged by their respective places in the paper's byline. Some many assumptions and simplifications must be used to evaluate research performance using bibliometric measurement.

As in other studies, a significant difficulty in the current study is the fair accessibility of resources across scientists within a certain profession. Research productivity is commonly regarded as the primary and essential performance indicator in most evaluation exercises. The current study examines this phenomenon by utilizing the FSS metric, which evaluates published works' quantity and quality. Following is the formula of FSS:

$$FSS_R = \frac{1}{t}\sum_{i=1}^{N}\frac{c_i}{\bar{c}}f_i$$

The professor's work history is indicated in the observed period by the letters t, N, and c_i, respectively, while N denotes the number of publications and c_i denotes the number of citations each article has received. The citation distribution means for all referenced works from the same year and subject area as the publication in question is represented by the variable "c." The variable "f_i" denotes the fractional contribution of the researcher to the publication labelled as "i". As opposed to the method of "full counting," which is used in this case, the methodology of "fractional counting" of research contributions is used. It is believed to be more in line with the microeconomic theory of production. The methodology employed in this study allows for the systematic assessment of individual author contributions, considering their respective positions in the byline.

In certain academic disciplines, the fractional contribution is determined by arranging authors in alphabetical order and is mathematically defined as the reciprocal of the total number of authors involved. However, the fractional contribution is assigned different weights in various other disciplines. In life sciences, it is customary, both within Italy and internationally, for authors to indicate their contributions to published research by arranging their names in a specific order within the byline. Within these academic disciplines, the relative importance assigned to individual co-authors is contingent upon their position within the byline and the type of co-authorship involved, namely intra-mural or extra-mural. A publication's first and last authors affiliated with the same university receive 40% of the citations, with the remaining 20% going to the other authors.

When the first two and last two authors of a publication are affiliated with different universities, the citation distribution can be described as follows: According to the data, it is observed that 30% of citations are attributed to the authors listed first and lasted in the publication. Additionally, 15% of citations are ascribed to the authors listed second and second-to-last. The remaining 10% of citations are distributed among all other authors involved in the publication. Abramo et al. (2015) suggest that neglecting to consider the number and arrangement of authors in the paper's byline may result in notable distortions in individual rankings.

Productivity is an essential indicator for assessing the efficiency of production systems. Nevertheless, measuring research excellence is a crucial metric in evaluating performance, as it relates to the ability to generate pioneering findings. As a result, we quantify the number of articles attributed to the top 1%
($HCA_{1\%}$) and5%($HCA_{5\%}$) of global publications based on their citation count, concerning each professor. Within the realm of life sciences, it is customary for the initial author's byline entry in a publication to indicate the individual who originated the primary idea and the researcher who made the most substantial contributions to the study and composition. The final author's position typically denotes the role of team leader in the research project in a corresponding manner. The assignment of primary or final authorship is considered a mark of prestige and is widely acknowledged in the scientific community. Following this, the number of scholarly articles in which a professor assumes the position of either the primary author or the last author is denoted as$First_A$or $Last_A$, respectively. In general, it can be deduced that the Pakistan context is improbable to yield distorted performance measures due to variable returns to scale, which could be attributed to disparities in university sizes.

3. Data and Methodology

This section explains the data collection, data preparation, variables and their measurement and analytical technique the current study uses. Following are the details for data and sample, and methodology, respectively.

3.1. Data and Sample

The data of individual professors are takenWeb of Sciences. The data consists of 126 researchers from five universities in Islamabad, Pakistan. These are Allama Iqbal Open University, International Islamic University, Quaid-e-Azam University, Comsat, and National University of Science and Technology. In this sample, we considered two disciplined professors, data mathematics and computer sciences. The sample considered from 2010 to 2023. We also consider five other performance indicators to assessthe research performance of these two departments. These include the FSS index, $HCA_{1\%}$, $HCA_{5\%}$, $First_A$, and $Last_A$ indicatorsfor each of these, we measure the five indicators described in the ‘‘Indicators in Research Performance Assessment’’ section.

3.1.1. Methodology

The current study has used descriptive analysis and a non-parametric SD approach to assess individual research performance at the departmental level. The current study has used the stochastic rather than the deterministic approach. We have used STATA 15 and Microsoft Excel 2019 for data handling and descriptive statistics. However, we used GAUSS 10 software to apply the SD approach and its first, second, and third orders of SD. Following are the details of the SD approach.

3.1.2. Parametric Approach

In the parametric approach, we do descriptive and graphical analysis to examine the selected departments' research performance.Firstly, we document the variables' values: indicators, publications, and impact for the sampled departments. And then, we draw the graphs based on these variables' values. Secondly, we present the descriptive statistics of the core five indicators of research performance along with necessary graphical analysis.

3.1.3. Non-Parametric Approach: Stochastic Dominance

After discussing parametric methods, we use the non-parametric stochastic dominance (SD) KS type test (Barrett & Donald, 2003) to evaluate the performance comparison between the two departments.The stochastic dominance (SD) approach is commonly employed to assess whether one series exhibits stochastic dominance over another at a particular stochastic order. This study aims to investigate the potential dominance of the mathematics department over the computer sciences department by employing the first three SD rules. The three rules under consideration are commonly referred to as first-orderSD (FSD), second-order SD (SSD), and third-order SD (TSD). We take into consideration two publishing series to better comprehend the SD rules,denoted as X and Y, both having a stochastic outcome of publication, denoted as "P".

Let$\{X_i\}$, where i = 1, 2, ..., N represent i.i.d sample of publication series that follow a dominant distribution with the cumulative frequency distribution F_X(b). By making the assumption that theCDFs typically fall within the interval [0, x], where x is a positive value, and are continuous functions within the range [0, x], we establish the following criteria to explain whether the function $D_X^s(b)$integrates F_X(P) to a stochastic dominance order s = i.

$$D_X^1(\text{b}) = F_X(b) \qquad \text{For FSD} \quad (1)$$

$$D_X^2(\text{b}) = \int_0^x F_X(\delta)d\delta = \int_0^x D_X^1(\delta)d\delta \qquad \text{For SSD} \quad (2)$$

$$D_X^3(\text{b}) = \iint_0^x F_X(\gamma)d\delta d\gamma = \int_0^x D_X^2(\gamma)d\gamma \qquad \text{For TSD} \quad (3)$$

Let us consider a set of random variables $\{Y_i\}$, where i=1,2,...,N, that i.i.d samples from a non-dominated distribution with a CDF of $F_Y(b)$. Subsequently, we proceed to establish the distribution of $D_Y^s(b)$ for the function $F_Y(b)$, employing a similar approach as previously employed to define D_X^s(b).. Hence, the test encompasses the null and alternative hypotheses that aim to examine the stochastic dominance order of portfolio "X" over "Y".

$H_0^s : D_X^s(\text{b}) \leq D_Y^s(\text{b})$ for all b (b is a series of publications)
$H_1^s : D_Y^s(\text{b}) > D_X^s(\text{b})$ for some b

To examine the null hypothesis, the following KS test statistic is used.

$$K_s = \left(\frac{N^2}{2N}\right)^{1/2} \frac{sup}{x}[D_X^{-s}(b) - D_Y^{-s}(b)] \quad (4)$$

The orders of standard deviation considered in this study are at least second order (s = 2) or higher (s > 2). The utilization of the simulation methodology enables us to ascertain the p-values about the underlying null hypothesis and make an estimation of the suprema of test statistics, denoted as K_s(Barrett & Donald, 2003).

4. Results and Discussion

Table 1 documents the research performance of two departments of Islamabad's five universities through descriptive statistics of core indicators of research performance. The departments are named: Mathematics and Computer Sciences. The Table consists of three columns. Column 1 lists those core indicators that mainly contribute to professors' research performance. Columns 2 and 3 give the values of the core indicators for the mathematics and computer sciences departments, respectively. The table presents a concise overview of the productivity and citation impact results about the professors affiliated with two distinct academic departments. The purpose of the productivity indicators, as shown in the upper section of Table 1, is to provide a comprehensive overview of the publication output, specifically focusing on the types of documents and the order of authors. The impact indicators in the lower section of the table consist of two types: The primary data considered in this study include total citations, the proportion of self-citations, and citations per publication.

Table 1. Scientific Performance of Two Departments of Five Universities of Islamabad, Pakistan

Indicator	Mathematics	Computer Sciences
Article	5351	5121
Proceeding Papers	1143	2008
Book	1	4
Book Chapter	31	93
Reviews	90	232
Editorial	10	21
Total Publications	**6626**	**7524**
Total documents consisting of articles, proceedings papers and reviews	6584	7361
Duration of Sample	24	24
The mean of Publications per year	276.1	313.5
Impact		
Total citations	142277	165359
Mean of citations per publication	21.4	21.98
The ratio of self-citations to total citations.	7.8%	9.93%
Mean of h-index	10.37	5.21

Note: Table 1 shows the descriptive statistics of the scientific research performance of two departments of the five universities of Islamabad, Pakistan. The research indicators include the core research indicators, total publications during the sample and their impact.

Additionally, the widely utilized h index is also taken into account. These various types of indicators offer distinct forms of information pertaining to scientific performance. The aggregate count of citations and the average number of citations per publication exhibit minimal deviation from the raw data for the given analysis.

The inclusion of the h index is necessary due to its widespread adoption within the scientific community. Due to the disparate methodologies employed by different impact indicators, there is a potential for conflicting outcomes in their assessment of research performance. We propose using advanced indicators, such as the h-index, as a recommendation. These indicators alone enable a comprehensive and equitable assessment of performance. It is crucial to consider the unique characteristics of each Department and the specific requirements of the evaluation process when interpreting the results, even when utilizing advanced indicators.

Figure 1 displays each Department's overall number of publications divided by document type. For further details, please refer to Table 1. It should be noted that the classification of publications by document types, as determined by Thomson Reuters, often deviates from the classification used by journals(Zhang, Wang, & Zhao, 2017). According toCostas, Van Leeuwen, and Van Raan (2010), database producers typically categorize original research findings as "Articles" and extensive literature surveys as "Reviews" upon publication. According to the data presented in Figure 1, publications classified as "Article" are the predominant document type for both departments. Proceedings papers are important in various academic departments, including computer sciences and mathematics. The mathematics department has published significantly more articles (n=7524) than its counterpart department (n=6626).

In conjunction with considerations of authorship and document classification, the timing of the issuance of publications also constitutes a noteworthy aspect of the researchers' assessment. Is there an equal or unequal distribution of publications? Is there a discernible trend in productivity over the years, indicating whether it increases or decreases? As depicted in Figure 2, the variability in article publication among the researchers of the departments examined in this study is evident (also refer to Table 1).

Figure 1: Publications with different Document Types by Two Departments

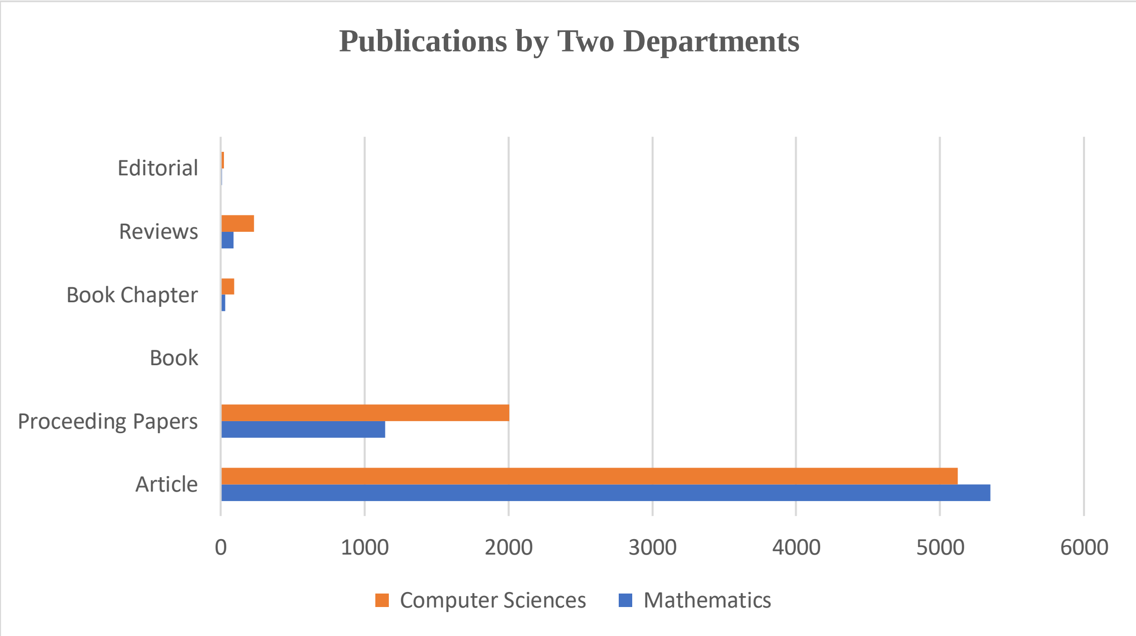
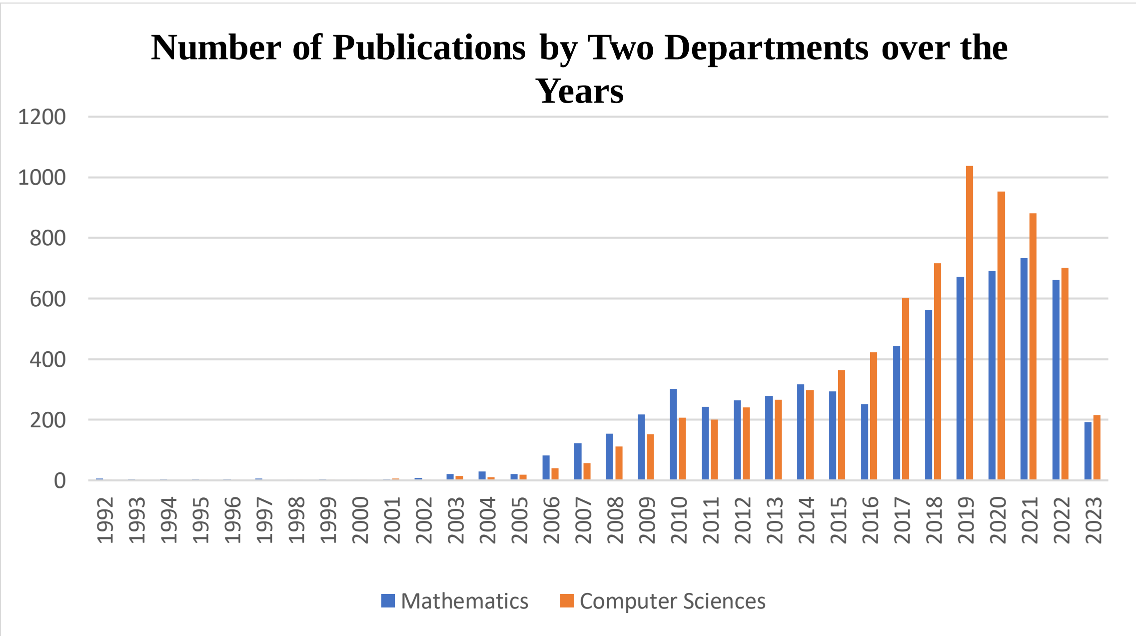


Figure 2:Yealy Publications by Two Departments



The Department of Computer Sciences attained peak levels of productivity approximately two decades into its academic trajectory, marked by the initiation of its publication endeavours. Subsequently, the Department has consistently produced a substantial volume of scholarly publications. The publications of both departments exhibited an upward trajectory since the commencement of the year, which reached a plateau between 1992 and 2017. Nevertheless, there is a noticeable disparity in publications between the mathematics department and its counterpart.

Table 2: Performance of two Department Based on Publication

	Mathematic *versus* Computer Sciences	Computer Sciences *versus* Mathematics
SD Orders	**KS P-value**	
SD1	0.272	0.009
SD2	0.382	0.003
SD3	0.294	0.001

Note:SD of two pairs of publications series of mathematics and computer sciences departments of five universities of Islamabad, Pakistan. SD1, SD2, and SD3 are three p-values of stochastic orders first, second, and third respectively.

Following the documentation of results through summaries and descriptive analysis, we employed the non-parametric method known as stochastic dominance (SD) to evaluate the performance of both departments with respect to the publications parameter. In this study, the KS test byBarrett and Donald (2003)assesses the statistical significance. Specifically, the first-, second-, and third-order standard deviations (SD1, SD2, and SD3, respectively) are utilized. Table 2 displays the p-values obtained from the Kolmogorov-Smirnov (KS) test. The purpose of this test is to assess the null hypothesis that the target publication series exhibits stochastic dominance over the other series at the s^{th} order of the sample under consideration. The table displays three sets of results indicating the order of p-values for SD1, SD2, and SD3. The statistical significance of the p-values indicates that the null hypothesis, which posits that computer sciences stochastically dominate mathematics, is primarily supported. The findings of this study indicate that the Department of computer sciences exhibits stochastic dominance over other departments in its field. In a broader context, upon examining the p-values of the panel of publications series, it becomes evident that we can reject the null hypothesis that mathematics stochastically dominates computer sciences in all instances of SDs. This rejection is supported by the observation that the p-values exceed significance levels. In an alternative perspective, we posit the hypothesis that the computer sciences department exhibits stochastic dominance over mathematics, as evidenced by all observed values below any significance level.

Table 3 presents the descriptive statistics for the distributions of the indicators relative to the two departments under examination. By examining these descriptive statistics, one can acquire valuable insights about the measures of the dataset's central tendency, variability, symmetry, and shape. Such insights are instrumental in comprehending the characteristics and behaviour exhibited by the data.

In reference to the Computer Sciences Department,the average FSS (Index) is 1.75, while the middle value, or median, is 1.25. The upper limit of the observed values is 42, while the lower limit is 0. The data set exhibits moderate variability, evidenced by a standard deviation (SD) of 3.2 and a coefficient of variation of 1.23. The obtained skewness value of 3.29 indicates a positive skewness, implying that the distribution is skewed to the right. The kurtosis value 28.3 also suggests that the distribution has heavy tails and exhibits extreme values.

The average value for $First_A$is 1.25, and the median, is also 1. The upper limit of the observed values is 21, while the lower limit is 0. The data set's standard deviation (SD) is 1.15, and the coefficient of variation (CV) is 1.41. The skewness value of 3.25 suggests that the distribution is skewed to the right, indicating a longer tail on the right side of the distribution. On the other hand, the kurtosis value 20.78 indicates heavy tails and the presence of extreme values in the distribution.

The average value for $Last_A$is 2.54, and the median value is 1. The upper limit of the observed values is 84, while the lower limit is 0. The standard deviation (SD) is calculated to be 1.61, while the coefficient of variation (CV) is determined to be 5.9, suggesting a substantial level of variability. The obtained skewness value of 5.12 indicates a highly skewed distribution to the right. Additionally, the calculated kurtosis of 75.15 suggests that the distribution has heavy tails and exhibits extreme values.The average value for $HCA_{1\%}$ is 0.15, similarly, the average value for $HCA_{5\%}$is 0.67,

Regarding the Mathematics Department,the average FSS (Index) is 2.2; the middle value is represented by a median 1. The upper limit of the observed values is 89, while the lower limit is 0. The standard deviation (SD) is calculated to be 4.27, while the coefficient of variation is determined to be 1.71, suggesting a moderate level of variability. The obtained skewness value of 6.4 indicates a highly skewed distribution to the right. At the same time, the calculated kurtosis of 80.12

suggests the presence of heavy tails and extreme values in the distribution.

The average value for $First_A$is 1.15, and the middle value is 0. The upper limit of the observed values is 22, while the lower limit is 0. The standard deviation (SD) is 1.2, while the CV is 1.31. The observed skewness value of 3.25 suggests that the distribution is skewed to the right, while the calculated kurtosis of 25.34 indicates the presence of heavy tails and extreme values.

The average value for $Last_A$ is 2.56, and the middle value is represented by the median, which is 1. The upper limit of the observed values is 42, while the lower limit is 0. The standard deviation (SD) is calculated to be 2.31, while the CV is determined to be 2.71. These values suggest the presence of moderate variability within the data. The obtained skewness value of 4.12 indicates that the distribution under consideration exhibits a right-skewed pattern

Table 3: Descirptive Statistics of Performance Indicators of two Departments

Departments	**Indexes**	**Mean**	**Median**	**Max**	**Min**	**SD**	**Variant coeff**	**Skewness**	**Kurtosis**
Computer Sciences	FSS	1.75	1.25	42	0	3.2	1.23	3.29	28.3
	$First_A$	1.25	1	21	0	1.15	1.41	3.25	20.78
	$Last_A$	2.54	1	84	0	4.26	1.61	5.9	75.15
	$HCA_{1\%}$	0.15	0	6	0	0.34	2.61	5.12	60.12
	$HCA_{5\%}$	0.67	0	10	0	1.23	1.59	2.15	10.12
Mathematics	FSS	2.2	1	89	0	4.27	1.71	6.4	80.12
	$First_A$	1.15	0	22	0	1.2	1.31	3.25	25.34
	$Last_A$	2.56	1	42	0	3.4	2.31	2.71	15.16
	$HCA_{1\%}$	0.2	0	8	0	0.58	1.41	4.12	24.54
	$HCA_{5\%}$	1.18	0	30	0	1	2.13	4.61	55.83

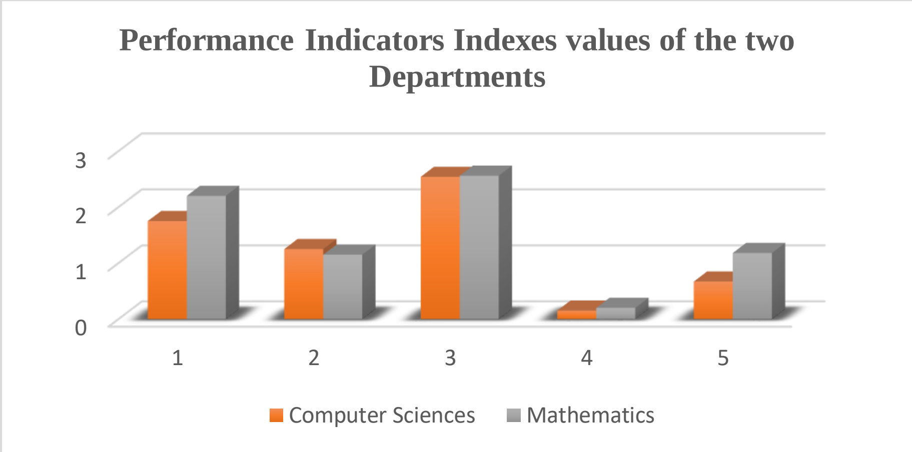
Note: The table shows the descriptive statistics of five core performance indicators of two selected departments

When examining Figure 3, a visual assessment can be conducted to compare the average performance indicators of the two departments. In figure 1 denotes FSS, denotes for $First_A$3 denotes for $Last_A$ 4 denotes $HCA_{1\%}$and 5 denotes for $HCA_{5\%}$for In Computer Sciences, it is evident that the third performance indicator exhibits the highest mean value, followed by the first indicator. In contrast, the fourth indicator demonstrates the lowest mean value. The second and fifth indicators are situated in an intermediate position.

In the field of Mathematics, it can be observed that the initial performance indicator exhibits the highest average value, with the subsequent indicator ranking second in terms of the mean value. The mean values of the second and fifth indicators are comparatively lower, while the fourth indicator exhibits the lowest mean value.

It is important to acknowledge that this interpretation is derived exclusively from the average values presented in the bar chart. The analysis fails to consider additional variables or the wide distribution of the dataset.Examining the bar chart, a visual assessment can be made regarding the average performance indicators of the two departments. In Computer Sciences, it is evident that the third performance indicator exhibits the highest mean value, followed by the first indicator, whereas the fourth indicator demonstrates the lowest mean value. The second and fifth indicators are situated at intermediate positions.

Figure 3: Performance Indicators Indexed values of the two Departments



In the field of Mathematics, it can be observed that the initial performance indicator exhibits the highest mean value, with the subsequent indicator ranking second in terms of magnitude. The mean values of the second and fifth indicators are comparatively lower, while the fourth indicator exhibits the lowest mean value.

Table 4 presents descriptive statistics for core performance indexes indicator, specifically citations and h-Index, based on the $HCA_{1\%}$ index.

In the Computer Sciences Department context, it is imperative to adhere to proper citation practises. The values in the dataset exhibit a range spanning from 573 to 1729. The data exhibits a rising trajectory in the number of citations, suggesting a future expansion in the research impact of the Department. The h-index values in this range span from 11 to 21. The h-index is a metric that quantifies the maximum number of papers within a dataset with equal or greater citations. The data indicates a moderate distribution of h-index values, suggesting detectable research productivity and influence level.

In the context of the Mathematics Department, the cited values exhibit a range spanning from 1277 to 21260. The data demonstrates a notable disparity in citation counts, suggesting a diverse spectrum of research influence among various publications. The h-index values in this study span a range of 19 to 84. The data indicates that the h-index values in Mathematics are relatively higher than those in the Computer Sciences Department, suggesting a potentially greater research impact and productivity in Mathematics.

Table 4: Descriptive based on $HCA_{1\%}$ Performance Indicators

Computer Sciences	
Citations	**h-Index**
573	14
666	15
676	13
703	11
782	16
1715	11
1729	21
Mathematics	
Citations	**h-Index**
1277	19
1865	24
4696	35
7656	47
8287	48
8862	54
21260	84

Figure 4 and 5 are drawn based on Table 4. Figure 4, we analyze the performance based on the citation score of sampled departments of the top 1% of researchers. Similarly, Figure 5 also considered the data of the top 1% of researchers but here we show the performance based on h-index. It is seen that the citation score of the mathematics department is high compared to its counterpart. Similarly, it would be obvious that the h-index is also high for the mathematics department.

Figure 4: Performance of $HCA_{1\%}$ based on Citation Score

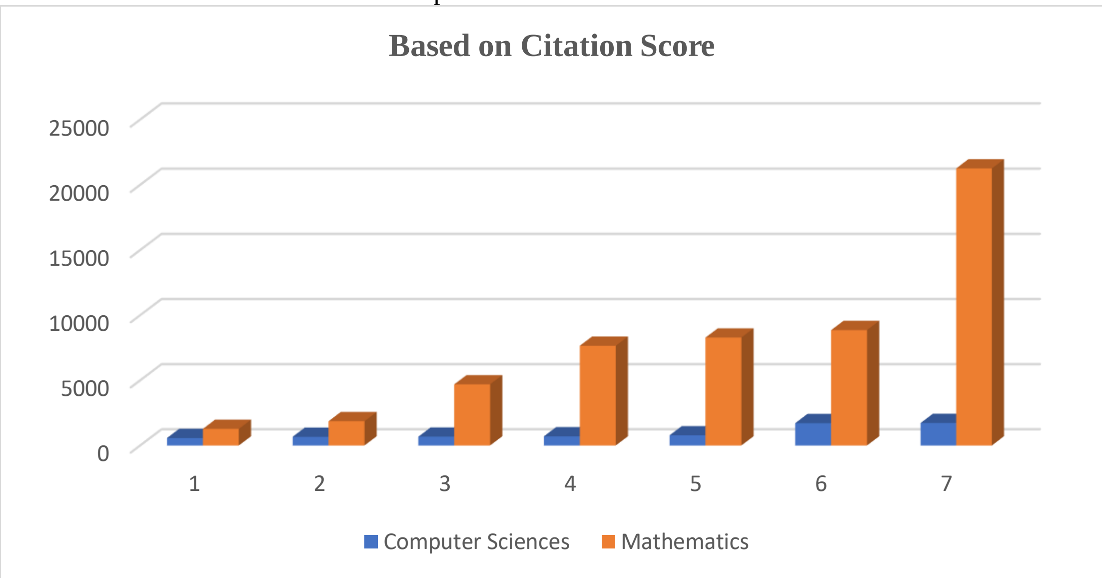
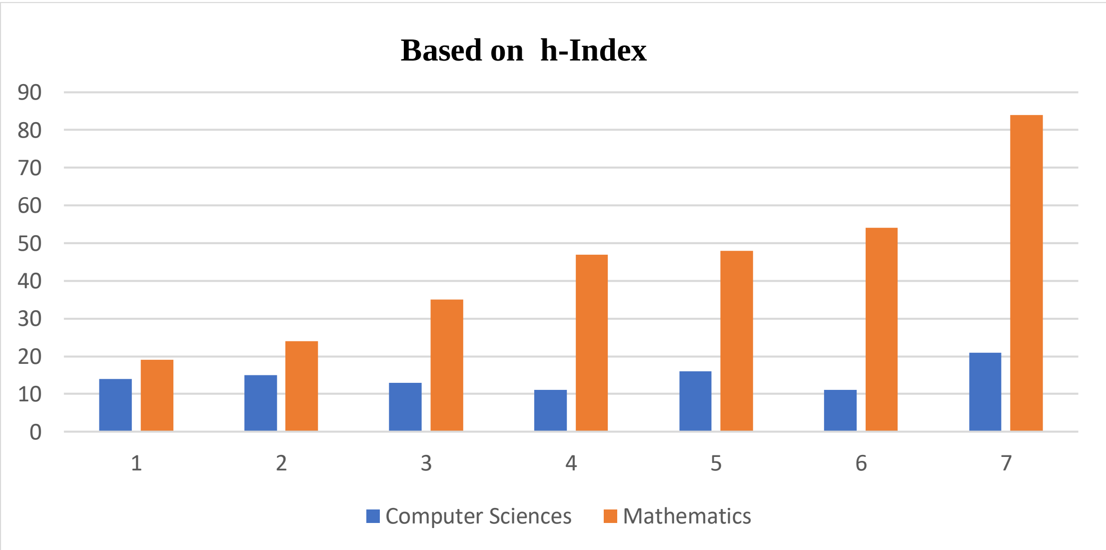


Figure 5: Performance of $HCA_{1\%}$ based on h-Index



5. Conclusion

In this study, we assess the performance of the mathematics and computer sciences professors of five universities in Islamabad, Pakistan, using parametric and non-parametric stochastic dominance approaches. In the parametric approach, we do a bibliometric analysis and examine the performance of both departments. In bibliometrics analysis, we consider the selected documents' publications: articles, proceeding papers, editorials, books, book chapters and reviews. We also consider citations of publications, h-index and self-citations to examine researchers' impact or productivity. For the formal method, we applied the stochastic dominance approach to examine the performance comparisons of two departments based on publications. The core findings are presented through formal tables and graphs.

Many nations are moving towards research policies that prioritize excellence. When assessing the performance of a researcher or Department in terms of research, it is advisable to complement bibliometric analyses with the examination of additional indicators. It is highly advisable to consider additional criteria when evaluating individual research performance. The criteria encompass various aspects, including teaching, guidance, engagement in group endeavors, and fostering collaboration. Additionally, certain quantitative factors cannot be assessed through bibliometrics alone, such as the number of patents, invitations to speak, international agreements, awards, and technology transfers.

The enhancement or replacement of bibliometrics is necessary, especially in disciplines outside the natural and life sciences realm. The current databases available for the humanities and social sciences, including philosophy, languages, law, sociology, history, psychology, art, and mathematics, are inadequate in their coverage of these disciplines. Consequently, these disciplines face limitations in effectively utilizing bibliometrics.

References

Abramo, G., Costa, C., & D'Angelo, C. A. (2015). A multivariate stochastic model to assess research performance. *Scientometrics, 102*, 1755-1772.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2023). Correlating Article Citedness and Journal Impact: An Empirical Investigation by Field on a Large-Scale Dataset. *Scientometrics, 128*(3), 1877-1894.

Barrett, G. F., & Donald, S. G. (2003). Consistent Tests for Stochastic Dominance. *Econometrica, 71*(1), 71-104.

Bornmann, L., & Williams, R. (2017). Can the Journal Impact Factor be used as a Criterion for the Selection of Junior researchers? A Large-Scale Empirical Study based on ResearcherID Data. *Journal of Informetrics, 11*(3), 788-799.

Costas, R., Van Leeuwen, T. N., & Van Raan, A. F. (2010). Is Scientific Literature Subject to a 'Sell-By-Date'? A General Methodology to Analyze the 'Durability'of Scientific Documents. *Journal of the American Society for Information Science and technology, 61*(2), 329-339.

Gao, Y., Shi, S., Ma, W., Chen, J., Cai, Y., Ge, L., . . . Tian, J. (2019). Bibliometric Analysis of Global Research on PD-1 and PD-L1 in the Field of Cancer. *International Immunopharmacology, 72*(7), 374-384.

Hamann, J. (2016). The Visible Hand of Research Performance Assessment. *Higher Education, 72*, 761-779.

Ho, M.-T., Mantello, P., Nguyen, H.-K. T., & Vuong, Q.-H. (2021). Affective Computing Scholarship and the Rise of China: A View from 25 Years of Bibliometric Data. *Humanities and Social Sciences Communications, 8*(1), 1-14.

Kurtz, M. J., & Henneken, E. A. (2017). Measuring Metrics-a 40-Year Longitudinal Cross-Validation of Citations, Downloads, and Peer Review in Astrophysics. *Journal of the Association for Information Science and Technology, 68*(3), 695-708.

Leydesdorff, L., Bornmann, L., Mutz, R., & Opthof, T. (2011). Turning the Tables on Citation Analysis One More Time: Principles for Comparing Sets of Documents. *Journal of the American Society for Information Science and Technology, 62*(7), 1370-1381.

Onodera, N., & Yoshikane, F. (2015). Factors Affecting Citation Rates of Research Articles. *Journal of the Association for Information Science and Technology, 66*(4), 739-764.

Rafols, I., & Stirling, A. (2021). *Designing Indicators for Opening up Evaluation: Insights from Research Assessment*.

Sahel, J.-A. (2011). Quality Versus Quantity: Assessing Individual Research Performance. *Science translational medicine, 3*(84), 84cm13-84cm13.

Tang, X., Wan, X., & Zhang, X. (2014). *Cross-Language Context-Aware Citation Recommendation in Scientific Articles.* Paper presented at the Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval.

Torres-Salinas, D., Robinson-García, N., & Gorraiz, J. (2017). Filling the Citation Gap: Measuring the Multidimensional Impact of the Academic Book at Institutional Level with PlumX. *Scientometrics, 113*(3), 1371-1384.

Zhang, J., Wang, Y., & Zhao, Y. (2017). Investigation on the Statistical Methods in Research Studies of Library and Information Science. *The Electronic Library, 35*(6), 1070-1086.

Zhu, K., Lin, R., & Li, H. (2022). Study of Virtual Reality for Mild Cognitive Impairment: A Bibliometric Analysis Using CiteSpace. *International Journal of Nursing Sciences, 9*(1), 129-136.