IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 July 2023, accepted 7 August 2023, date of publication 10 August 2023, date of current version 17 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3304013

RESEARCH ARTICLE

Exploring the Significance of Publication-Age-Based Parameters for Evaluating Researcher Impact

GHULAM MUSTAFA[1], ABID RAUF[1], AHMAD SAMI AL-SHAMAYLEH[2], BILAL AHMED[3], (Graduate Student Member, IEEE), WAGDI ALRAWAGFEH[4], MUHAMMAD TANVIR AFZAL[5], AND ADNAN AKHUNZADA[4], (Senior Member, IEEE)

[1]Department of Computer Science, University of Engineering and Technology, Taxila 47080, Pakistan
[2]Department of Network and Cybersecurity, Faculty of Information Technology, Al-Ahliyya Amman University, Amman 19328, Jordan
[3]College of Information and Computer, Taiyuan University of Technology, Taiyuan 030000, China
[4]College of Computing and IT, University of Doha for Science and Technology, Doha, Qatar
[5]Department of Computing, Shifa Tameer-e-Millat University, Islamabad 44000, Pakistan

Corresponding author: Wagdi Alrawagfeh (wagdi.alrawagfeh@udst.edu.qa)

ABSTRACT In the modern era, bibliometric parameters have become crucial components in evaluating academic productivity. In the last three decades, the scientific community has suggested a variety of bibliometric parameters to rank researchers, such as publication count, citation count, author count, h-index, and its variants. There is an ongoing debate within the scientific community regarding which index provides the most accurate ranking of authors. Currently, the state-of-the-art evaluation of these indices is based on hypothetical scenarios and the limited and unfair distribution of researcher records in the datasets. In addition, these indices are often evaluated using dissimilar datasets, making it challenging to accurately compare their contributions and significance. To comprehensively analyze the performance of each index, it should be evaluated on extensive and balanced datasets. This study focuses on evaluating the effectiveness of some h-index extensions, such as the platinum h, M quotient, AW, AR, V, Ha, Hc, and AWCR indices. For the experiments, this study employed a comprehensive and balanced dataset from the mathematics field to evaluate the performance of these indices. Analysis of these indices was conducted in three stages. First, the correlations among these indices were calculated. A weak correlation between indices suggests that the rankings of authors obtained from these indices are not the same, whereas a strong correlation suggests the opposite. Second, the analysis examines the occurrence of awardees in all ranked lists, using prestigious award winners from four mathematics societies as a benchmark. Based on the analysis, the study reveals that the Ar index performs the best among the evaluated indices, bringing the highest number of awardees to the top 10% of the ranked list, with a maximum of 80%. However, none of the indices can bring all the awardees to the top rankings, indicating that there is no single index that can provide a complete and accurate evaluation of the author's ranking. Additionally, the study identified that the maximum awardees at the top of all ranked lists belong to the American Mathematical Society (AMS), suggesting that AMS may depend on these indices to determine its award recipients. Furthermore, the AWCR index is more suited to the AMS society by its contribution to returning the highest number of awardees belonging to the AMS society.

INDEX TERMS Bibliometric parameters, H index variants, publication age-based index, ranking, research evaluation, mathematics domain, data extraction.

I. INTRODUCTION

In the modern era, researchers are ranked based on a variety of bibliometric parameters, both qualitative and quantitative,

The associate editor coordinating the review of this manuscript and approving it for publication was Jolanta Mizera-Pietraszko.

to determine their level of success [1], [2], [3]. In addition to finding researchers' success, rankings also serve multiple other scenarios such as the selection of award winners, selection of candidates for tenured positions, students finding their supervisor for Ph.D. supervision, and finding appropriate reviewers for journals and conferences. Furthermore, researchers in any field are ranked based on the impact of their research work on the scientific community [4]. Therefore, ranking authors help to address questions related to issues such as identifying who should receive scholarships, whose research work is more impactful, who should be granted tenure, and who should be promoted [5].

In the past, conventional bibliometric indices such as publication count [6], citation count [7], and co-author count [8] were used to rank researchers. However, Hirsch [9] argued that using these measures individually cannot accurately determine the impact of researchers. Consequently, he proposed the h-index [9], which combines the properties of publication and citation counts to rank researchers. Since then, according to Bihari et al. [10] more than 70, the h-index variant and extensions have been proposed, and they have been widely used by the research community to rank authors. Some popular variants include the a-index [11], ar-index [12], m-quotient [13], contemporary h-index [13], f-index [14], Wu-index [7], and q2-index [15]. Furthermore, extensions include hm-Index [16], hc-index [17], fraction counts on paper [16], and age-based h-index [18], etc. While numerous parameters have been proposed in the literature to rank researchers, there is still a lack of consensus among the scientific community regarding the optimal parameter. This is partly because mostly proposed parameters are evaluated in hypothetical or fictional case scenarios, rather than extensive empirical evaluation [19]. Furthermore, some studies evaluating these measures have used small and imbalanced datasets, which may not be representative of the entire scientific community [20]. Additionally, the lack of a comprehensive empirical evaluation of these parameters makes it difficult to determine which parameter is most effective in practice. Furthermore, it is important to evaluate different ranking techniques on the same dataset to make meaningful comparisons and determine which parameters perform extraordinarily for ranking authors in a particular domain. This will also help to identify the strengths and weaknesses of each technique and guide the selection of the most appropriate technique for a particular purpose. For these reasons, the debate regarding the optimal ranking parameters for researchers continues.

This study aims to address the aforementioned deficiencies by evaluating various ranking parameters belonging to publication age-based categories such as the Platinum h index, AW index, M quotient index, AR index, V index, Ha index, Hc index, and AWCR index. We collected a comprehensive and balanced dataset of authors from the mathematics domain to evaluate these ranking parameters comprehensively. The primary objective of this study was to provide an optimal parameter for the scientific community to rank researchers. To achieve this objective, we address the following research questions:

- What type of correlation exists between publication age-based indices?
- Do the winners of international awards from prestigious mathematical societies also rank at the top positions of the lists obtained by publication age-based indices?
- Do the ranking criteria of publication age-based indices align more with any particular mathematically awarding society or are they consistent across all societies?

Main Contribution: In this study, we utilized a dataset from the mathematics domain consisting of data from 1050 authors. The dataset comprised 525 non-awardees of data collected from the dataset provided by Ghani et al. [21]. To balance the dataset, we added the data of the remaining 525 awardees from four prominent mathematics societies (AMS, IMU, LMS, and NASL). In the first phase of our experiment, we computed all publication age-based indices using this dataset, resulting in the corresponding ranking lists for each index. We then analyzed the correlation between the results of each index to determine their similarities and differences in the ranking patterns. Our findings indicate that most indices were strongly correlated or negatively correlated, suggesting that these indices follow either similar or unique patterns in ranking researchers.

The analysis of the researcher ranking lists showed that 49% of the awardees appeared in the top positions. Furthermore, the Ar index brings 80% of the awardees to the top 10% of the ranking list.

The structure of the remaining documents is as follows. The next section presents a summary of previously proposed techniques. The proposed methodology is presented in the following literature review. The experimental results are presented in the results section. The conclusion section provides a summary of the proposed study and the final section discusses future research directions.

II. LITERATURE REVIEW

Assessing author's scientific productivity is crucial in various aspects of research, including making informed decisions, nominating scientific awards, assigning research projects, promoting individuals, awarding tenures to skilled individuals, and granting contracts to experts. In the literature, more than 70 parameters have been proposed over the last two decades [10]. These parameters were used to rank the authors based on their research work. One of the earliest parameters for evaluating a scientist's research is publication count [15]. However, this parameter fails to capture the impact and quality analysis of an author's work because some researchers published articles in journals with low-impact factors to increase their publication count. To address this limitation, citation count was introduced, but it has some drawbacks, such as the time required for newly published papers to acquire citations, which can be a disadvantage for novice authors, even if they have a high publication count. Additionally, citation count may not accurately reflect the quality of a publication because researchers may cite papers by other authors to criticize them. Furthermore, researchers have also cited their work to increase their citations, called self-citation.

To address the limitations of publication and citation counts, Hirsch [9] introduced the h-index, which quickly gained popularity and became the dominant parameter for evaluating an author's research output. However, the h-index still has many weaknesses, such as being mainly based on long-term observations, resulting in older authors having larger h-indices than novice authors, and being field-specific. Moreover, an increase in the citation of h core articles does not impact the overall h index. Various variants and extensions have been introduced to overcome these limitations of h index [22]. Some well-known variants include the m-quotient [13], f-index [14], s-index [23] hm-Index [16], [24] etc. Many studies have been conducted to evaluate different indices on various datasets. For example, Kosmulski et al. [25] conducted a study to find the relationship between the h-index and the h(2) index using data from 19 professors belonging to the chemistry department of the University of Poland. The outcome of the study revealed a strong correlation between these two indices, indicating that their return results are the same.

Similarly, Van Raan [26] conducted a study in which they compared the h-index and its variants by employing data collected from a larger evaluation study of 147 research groups belonging to the chemistry department in the Netherlands. In this study, the author focused on group performance rather than individual performance. Moreover, they considered only three-year citations instead of whole-life citations to find the correlation between the h-index and its variants.

In 2016, Xiao et al. [27] evaluated 29 h-index variants. In this study, they calculated the correlation among these variants with the h-index and Wu index, which they considered as benchmarks. They identified from their analysis that the indices that are highly correlated to the h index are less correlated to the Wu index and vice versa. They also reported the highly correlated indices to show a slight improvement over the h-index or wu index.

Jin et al. [12] proposed a method for evaluating authors by combining different indices. The method utilizes a pair of indices, one as a quantitative measure and the other as a qualitative measure. The pairs considered indices such as the h-index, r-index, h-index, and ar-index. The outcome of this study revealed that the h-index and ar-index pairs performed well as indicators for research assessment.

In 2008, Schreiber [28], [29] conducted a study to compare the g-index with other indices, such as the h-index, r-index, and a-index, on a dataset consisting of 26 physicists belonging to the Physics Department at Chemnitz University of Technology. The study involved analyzing the citation records of these physicists, and the data sets used for analysis were collected from the Web of Science in January 2007. Based on these findings, Schreiber reported that the g-index is a more suitable measure for determining the overall impact of a scientist's publications than the h-index.

Moreover, in 2016, Raheel et al. [1] conducted a study in which they evaluated some h-index variants by employing the benchmark awardees dataset belonging to scientific societies of mathematics. They found that the complete-h index outperformed the other indices. Furthermore, Ameer and Afzal [30] in 2019 evaluated some of the quantitative parameters by employing a dataset belonging to neuroscience societies and identified that the R-index and hg-index were effective in ranking awardees in the top position. Ain et al [31] also conducted a study in 2019 to evaluate scientific quantitative parameters for researchers by employing a mathematics domain dataset. This study determined the correlation between the h-index and its variants for researchers and ranked the respective parameters based on award-winning researchers. However, these studies were limited in that they aimed to prove the association between the h-index or its variants with award-winning researchers who were awarded before proposing these parameters. Therefore, there is very little chance that these awardees are dependent on the h-index or its variants. It could only be a coincidence that they found some correlation between awardees and these parameters. To address this limitation, Usman et al. [32] conducted a study in which they evaluated the h-index and its variants by employing data from researchers in the civil engineering domain. They did not select researchers randomly; they selected awardees and non-awardees in the same period and included only those researchers who received awards after 2005. However, their dataset is sufficient to draw conclusive results regarding the parameters that are important for identifying award winners. Furthermore, in a recent study, Abdulrahman et al. [33] proposed a rule for the scientific community employing deep-learning models. For the experiments, they used different domain datasets, including civil engineering, mathematics, and neuroscience, as well as data from 500 researchers from each field. They reported that these rules achieved an accuracy of up to 70%.

The preceding paragraphs cover various studies that evaluated indices to assess or rank authors. The following observations were made from the discussion above. The indices discussed in the literature are typically assessed for accuracy in hypothetical scenarios. In addition, these indices are often applied to small, diverse, and imbalanced datasets. It is important to evaluate these indices using a comprehensive and balanced dataset specific to a certain domain. Furthermore, there is currently no established standard benchmark for assessing these indices; therefore, a benchmark dataset is necessary for evaluation purposes.

III. METHODOLOGY

After conducting an extensive analysis of the existing techniques proposed for author ranking, we formulated several research questions that we intend to address (given in the introduction section). The proposed methodology is illustrated in Figure 1. The subsequent sections provide detailed descriptions of each component of the methodology.

A. DOMAIN SELECTION

To evaluate the proposed technique, it was necessary to obtain data from a particular domain. To ensure diversity in the dataset, we chose the mathematical domain. This domain was selected because it is one of the most interconnected fields with other branches of science, such as computer science, physics, and chemistry. This demonstrates that the chosen domain is diversified in nature, and ranking authors according to this would be a significant accomplishment.

B. DATASET COLLECTION

For the experimentation, we gathered a comprehensive dataset to assess the effectiveness of the proposed methodology. The dataset comprises 1050 records, including 525 records of non-awardees and 525 awardees. To ensure diversity, we incorporated non-awardees' data from the dataset used by Ain et al. [31]. However, because this dataset had limited information on awardees until 2013, we updated our dataset by gathering data from various award societies' websites over the last two decades (AMS,[1]IMU,[2]LMS,[3]NASL[4]), collecting researcher names and years of awards. The Table 1 presents the counts of year-wise awardee data.

TABLE 1. Year wise awardees count.

Year	No of Awardees	Year	No of Awardees
1990	11	2007	16
1991	15	2008	18
1992	6	2009	24
1993	10	2010	22
1994	19	2011	18
1995	10	2012	22
1996	11	2013	15
1997	11	2014	33
1998	14	2015	12
1999	14	2016	12
2000	17	2017	16
2001	14	2018	23
2002	21	2019	12
2003	23	2020	13
2004	16	2021	20
2005	15	2022	20
2006	29	2023	13

We used the Publish or Perish [5] platform to extract the data of awardees by name. Adopting a "hold on" strategy, we gathered researchers' records before the year in which they received their award. Publish or Perish software deploys an advanced algorithm to extract both the primary data and metadata of authors from Google Scholars. We chose Google Scholar for data extraction because of its wide coverage of academic publications across various disciplines, global accessibility for researchers, and ability to retrieve both open-access and pay-walled publications. Previous studies comparing the coverage of Google Scholar with Web of Science (WoS) have found that Google Scholar's growth was 13% higher than that of WoS, and its citations have increased by an average of 1.5% per month over a year. Additionally, Google Scholar is constantly updated, with new data being uploaded weekly, ensuring that the information is always relevant and up-to-date [34]. To ensure fairness in our dataset, we collected an equal amount of non-award data for each year against the number of awardees in that year. For instance, in 1990, we collected 11 non-awardee data for the 11 awardees in that year, and we did the same for other years. The statistics of our dataset are shown in the Table 2.

TABLE 2. Data set statistics before preprocessing.

Features	Count
No of Authors	1050
No of Awardees	525
No of Non-Awardees	525
No of Total Citations	14370007
No of Total Publications	204896

C. DATA PREPOSSESSING

Before using data collected from sources such as Google Scholar for any analysis or evaluation, it is essential to cleanse the data thoroughly [35]. This is because the raw data may contain irrelevant or incorrect information, known as noise, which can negatively affect the validity of the results. The data-cleaning process involves identifying and removing this noise and can include steps such as verifying the accuracy of the data and eliminating duplicate entries. To improve the quality and relevance of our extensive collection of research materials, we conducted this study in two stages. First, we applied a filter to assess the consistency of each publication in the mathematics field, thus eliminating any irrelevant or non-mathematical content. Second, we conducted an author disambiguation process to identify and eliminate duplicate entries caused by authors publishing under different names. After verifying the aforementioned steps, the characteristics and properties of the final dataset for evaluation are listed in Table 3.

TABLE 3. Data set statistics after preprocessing.

Features	Count
No of Authors	1050
No of Awardees	525
No of Non-Awardees	525
No of Total Citations	143769007
No of Total Publications	204816

D. BENCHMARK DATA SET

To conduct a comprehensive evaluation of the publication age-based indices, we obtained lists of awards from

[1] https://www.ams.org/home/page
[2] https://www.mathunion.org/
[3] https://www.lms.ac.uk/
[4] https://dnva.no/norwegian-academy-science-and-letters
[5] https://harzing.com/resources/publish-or-perish

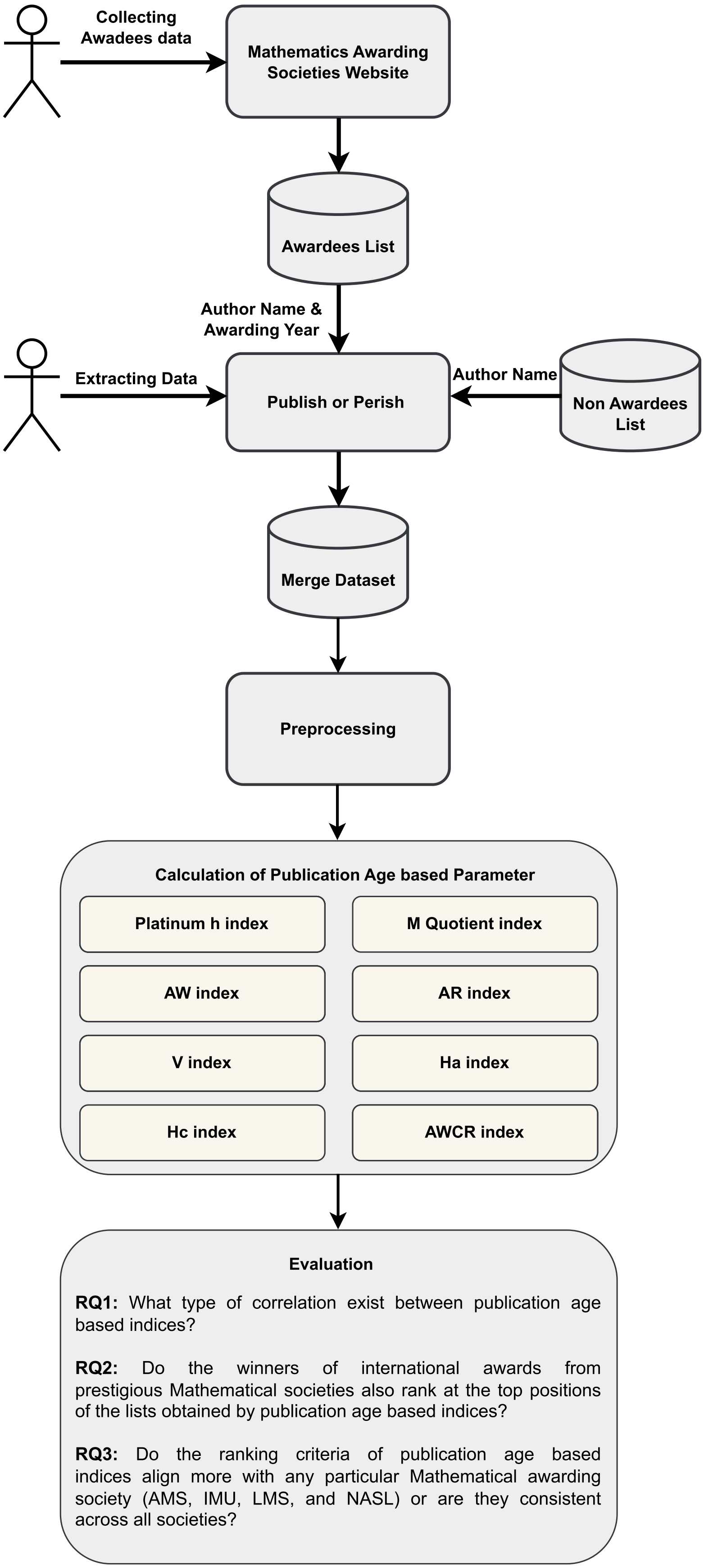

FIGURE 1. Proposed methodology diagram.

various mathematical societies that present awards. These lists include 30 internationally prestigious awards that are widely recognized in the mathematical community and are considered significant accomplishments for mathematicians and researchers. Several renowned mathematical societies and organizations present these 30 international prestigious awards, including the London Mathematics Society (LMS), International Mathematical Union (IMU), Norwegian Academy of Science (NASL), and American Mathematical Society (AMS). These organizations are committed to promoting and advancing mathematics and supporting mathematicians' research and academic endeavors worldwide. Researchers such as Raheel et al. [1], Ain et al. [31], and Ghani et al. [21] have also considered awardees from these societies as benchmarks. The reason for choosing these societies' awards as a benchmark is that there is no better benchmark available for evaluating these indices. The awards presented by these societies are among the most prestigious and competitive in the mathematical community and are based on a rigorous evaluation of the quality and originality of the researcher's work. The distribution of awardees in different societies is presented in Table 4.

TABLE 4. Number of awardees relative to society.

Socities Names	No of Awardees
AMS	257
IMU	59
LMS	188
NASL	21

E. CALCULATION OF INDICES

After collecting and pre-processing the data, we calculated the following eight indices. For implementation of indices we have used the python utility. These indices are variations of the h index and belong to the publication age-based category. The formula of the equation are given below:

- **M-quotient**
  The formula of M-quotient are given below:
  $$M - Quotient = \frac{h - index}{y} \tag{1}$$
  In above equation, y represents no of the year since the first publication and h-index represent the h-index of the author.
- **Hc – index**
  The formula of Hc – index are given below:
  $$hc - index = \alpha * \frac{C(i)}{(Y(now) - Y(i) + 1)} \tag{2}$$
  In above equation, Y(now) reflects the current year, Y(i) is the year of the publication, and C(i) refers to the citations of paper i.
  $$hc - index = \frac{C(i))}{1}, \frac{C(i))}{2} \ldots\ldots\ldots \frac{C(i))}{n} \tag{3}$$
- **Aw-index**
  The Aw-index is defined as the sum of the average number of citations per year of all articles. The formula of Aw-index are given below:
  $$Aw - index = \sqrt{\sum_{j=1}^{m} \frac{Cit_j}{a_j}} \tag{4}$$
  In above equation, Cit_j represent the total citation in one year, a_j represent individual year against paper.
- **Ar-index**
  The Ar-index is defined as the sum of the average number of citations per year of articles included in the h-core s. The formula of Ar-index are given below:
  $$Ar - index = \sqrt{\sum_{j=1}^{h} \frac{Cit_j}{a_j}} \tag{5}$$
  In above equation, Cit_j represent the total citation in one year, a_j represent individual year against paper and h represent the h index value.
- **Age Weight citation ratio (AWCR)**
  A measure of the average number of citations for an entire body of work, adjusted for the age of each paper
  $$AWCR = \sum_{i=1}^{n} \frac{Cit_i}{CY - y_i} \tag{6}$$
  where CY represent current year, y_i represent publication year and Cit_i represent the citation of publication.
- **V-index**
  The formula of v-index are given below:
  $$v - index = \frac{h}{p(y_{this} - y_o)} \tag{7}$$
  In above equation, h is the h-index, ythis is the current year and y0 is the year of first publication.
- **Platinum H-index**
  The formula of v-index are given below:
  $$platiniumh = \frac{h}{CL} * \frac{Cit_{all}}{Pub_{count}} \tag{8}$$
  In above equation, h is the h-index, CL is the career length (Current Year - Year of first Publication of Author), Cit_{all} is the total citation count and pub_{count} is the publication count
- **Ha-index**
  The ha-index against a dataset is the largest number of papers in the dataset that have obtained at least ha citations per year on average.
  $$max(h_a \leq \sum_{i=1}^{n} \frac{Cit_i}{y_i}) \tag{9}$$

After calculating the indices, we generated separate ranking lists for each index, resulting in eight distinct ranking lists for the authors. These ranking lists were then further analyzed to address the three research questions.

IV. EVALUATION

After calculating the ranked lists, this section addresses the evaluation of our three research questions. Our first question relates to the correlation between the indices.

Computation of Correlation between Indices: The primary goal of a correlation between indices is to examine the relationship and degree of similarity between them. To accomplish this, we used the Spearman correlation method (Corder and Foreman, 2009) to calculate the correlation.

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \tag{10}$$

In Equation the r_s represent Spearmans Correlation Coefficients, d represents difference between two rank of each observation and n represents number of observation. To answer RQ2, we determined the proportion of awardees present in the top 10%, 20%, and up to 100% of each ranking list to determine the most effective extension.

A. RESULT AND DISCUSSION

In this section, we discuss the detailed overview of the results obtained by applying the proposed methodology. **RQ1: What type of correlation exists between publication age-based indices?**

To determine the relationship between the different indices, we computed the Spearman correlation between their ranking lists. The resulting correlation coefficient can range from -1 to $+1$, where positive values indicate a similar upward trend, whereas negative values imply a dissimilar trend, such as an upward trend for one parameter and a downward trend for another. When the correlation coefficients moved in opposite directions, the correlation was considered negative. A high correlation coefficient value indicates a strong relationship between the ranking lists, indicating that they behave similarly.

The Table 5 demonstrates the obtained correlation coefficients among the ranked lists, which can be interpreted based on their absolute values. These values belong to distinct ranges that convey specific meanings, as defined by various researchers in their studies [21], [30]. The ranges with their specific meaning are given below:

- Very Weak (0.0 to 0.19)
- Weak (0.20 to 0.39)
- Moderate Correlation (0.40 to 0.59)
- Strong Correlation (0.60 to 0.79)
- Very Strong correlation (0.80 to 1.0)
- Negative Correlation (value<0)

Similar to Table 5, Figure 2 visually represents the correlation matrix between the indices. In this figure, dark blue represents a very strong correlation between the indices, whereas blue represents a strong correlation. Moreover, light blue represents a moderate correlation, and white or bluish-white represents a negative correlation and so on.

For most indices, the relationship is either very strong or negative. Only one moderate relationship exists between the v index and platinum h index. From all of them, the performances of the Aw, Ar, AWCR, hc, and ha indices are consistent. It has either a very strong or negative correlation. The very strong correlation between these indices indicates that they behave similarly or follow the same pattern. The overall results showed either a very strong correlation or a negative correlation. Furthermore, it is not possible to make a general conclusion based solely on the correlation values obtained, as indices that perform poorly in this analysis may perform well for other research questions.

RQ2: Do the winners of international awards from prestigious mathematical societies also rank at the top positions of the lists obtained by publication age-based indices?

All the publication age-based parameters portraying a good performance are on average, resulted in approximately 49, 46, 49, 50, 51, 51, 52, 52, 52, and 51 percent of the awardees being placed in the top 10,20,30,40,50,60,70,80,90 and 100 percent of the ranked list respectively. However, the Aw index is the only parameter that performs the lowest in all top results of the ranked list, such as 10% of the awardees being placed in the top 10% of the ranked list, and similar results are portrayed by other top results. The results are shown in Figure 3.

The behavior of each parameter is as follows:

- The AR index outperformed all other indices by retrieving 80% of the awardees in the top 10% of records. Its behavior is not static in all the top results; after the top 60%, the platinum-h index outperformed all the other indices.
- The behavior of the M Quotient and the AWCR index is similar by retrieving almost the same number of awardees in the top results; for example, in the top 10%, both indices retrieved 40 percent awardees. When selecting more top records, the behavior of the M Quotient index changes and increases the percentage of awardees in the top results.
- The Ha index also performed well by retrieving 70% of the awardees in the top 10 results.
- The Aw index does not perform well as compared to other indices by retrieving only 10% of awardees in the top 10% of records of the ranked list.

In conclusion, the best-performing extension was the AR index, platinum h index, and Ha index, while the lowest-performing extension was the Aw index.

RQ3: Do the ranking criteria of publication age-based indices align more with any particular mathematically awarding society or are they consistent across all societies?

Figures 4, 5, and 6 illustrate the relationship between all four awarding societies and extensions of the h-index. To address RQ3, we calculated the percentage of occurrences of the same authors from the awarding list in the Top 10%, 50%, and 100% of the ranked list obtained by the publication age-based indices.

TABLE 5. Represent the correlation values.

Indices Name	Platinium h index	M qoutient index	AW index	Ar index	V index	AWCR	Hc index	Ha index
Platinium h index	1	0.696827	-0.615722	-0.464145	0.460789	-0.743494	-0.779804	-0.745341
M qoutient index	0.696827	1	-0.96918	-0.953414	0.896333	-0.929347	-0.825789	-0.864242
AW index	-0.615722	-0.967918	1	0.978813	-0.849728	0.950168	0.862603	0.887065
Ar index	-0.464145	-0.953414	0.978813	1	-0.88804	0.855378	0.718327	0.768122
V index	0.460789	0.896333	-0.809728	-0.88804	1	-0.668118	0.511756	-0.57728
AWCR	-0.743494	-0.929347	0.950168	0.855378	-0.668118	1	0.974704	0.985173
Hc index	-0.779804	-0.825789	0.862603	0.718327	-0.511756	0.974704	1	0.991612
Ha index	-0.745341	-0.864242	0.887065	0.768122	-0.57728	0.985173	0.991612	1

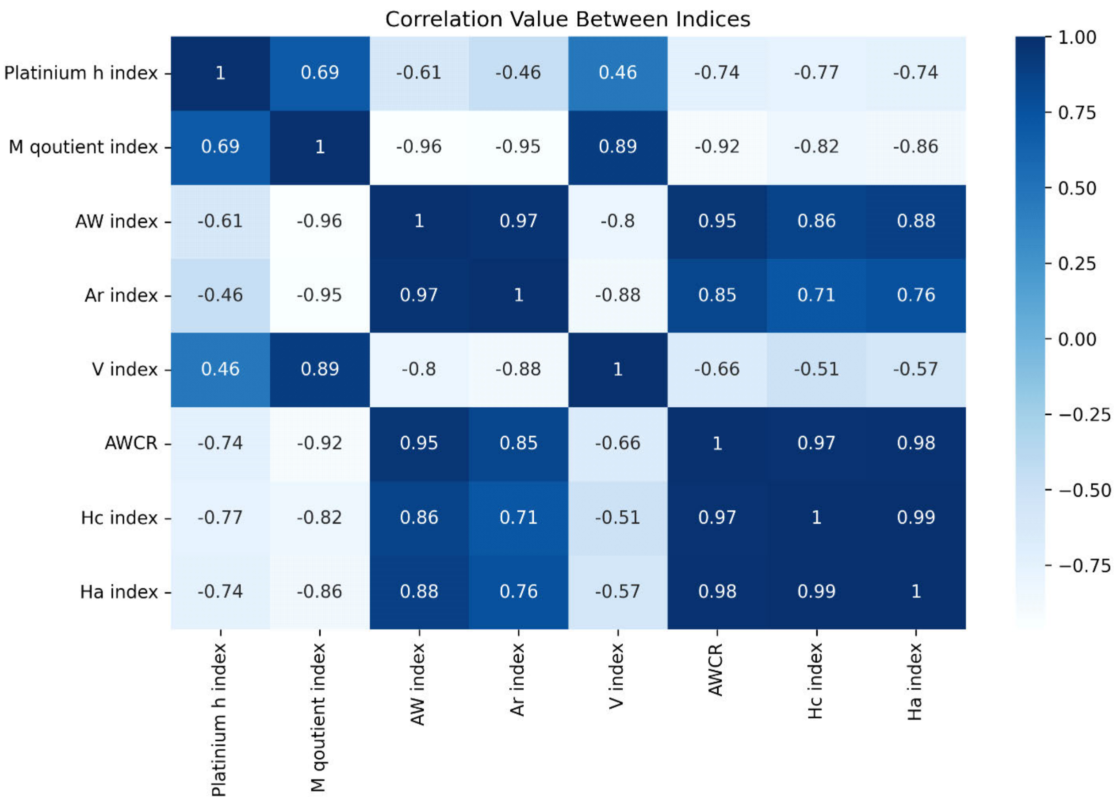

FIGURE 2. Correlation matrix.

By analyzing the dependence of awarding societies on these indices, we made the following observations.

1) AMS

- In the top 10% of the results (See Figure 4), the AWCR index outperformed all other indices by retrieving 75% of the awardees in the ranked list. Moreover, the Ha index showing average performance succeeded in retrieving 71% of the awardees. Furthermore, the Ar index showed low performance by retrieving 11% awardees, and the AW index returned nothing. Furthermore, the k norm index and w norm index retrieved 50 percent of awardees.
- In the top 50% of the results (See Figure 5), the AWCR Index outperformed all other indices by retrieving almost 61% of the awardees. Moreover, the Ha index shows average performance retrieved up to 53% of awardees. Furthermore, the M Quotient index showed poor performance, retrieving 25%.
- In the top 100% of the results(See Figure 6), the AWCR index remained good by retrieving 48% of awardees. The Ha and v indices showed equal performance by retrieving 46% of the awardees. The platinum h-index showed the lowest performance return of up to 34%.

2) IMU

- In the top 10% result (See Figure 4), the Aw index brought the maximum number of awardees from IMU society and exhibited a performance level of 100%. The platinum index showed an average performance

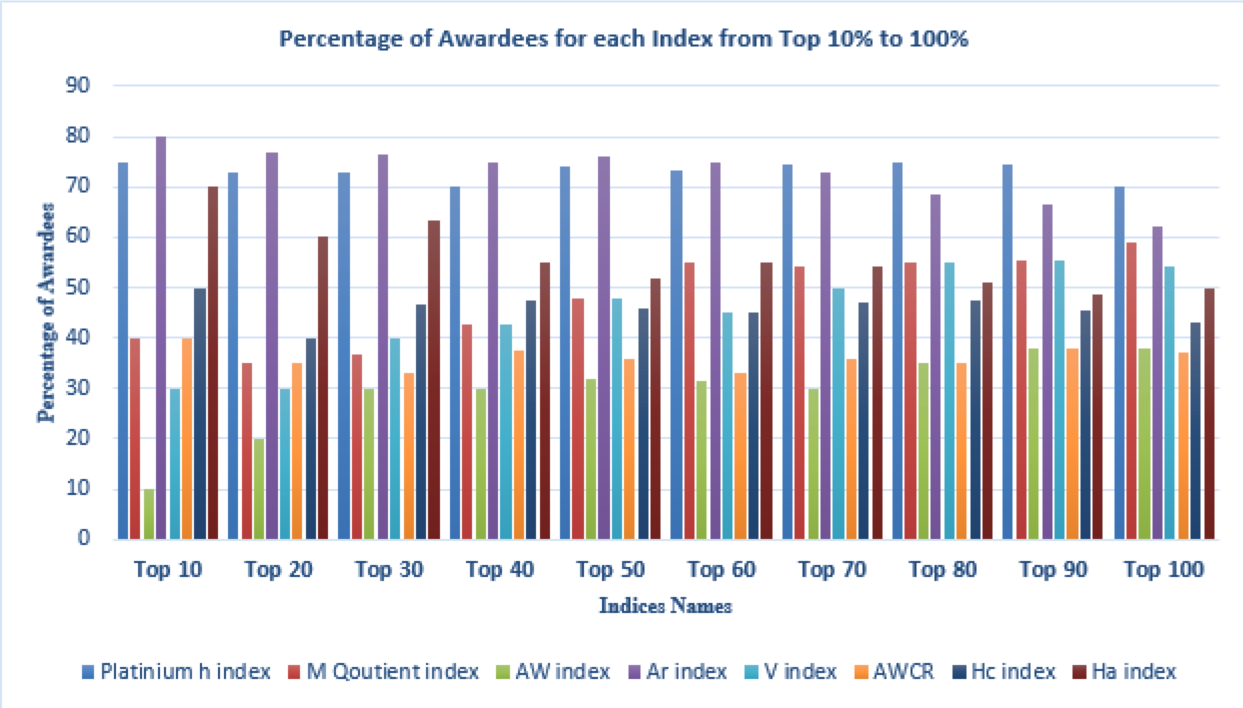

FIGURE 3. Top 10% to 100% occurrence of awardees.

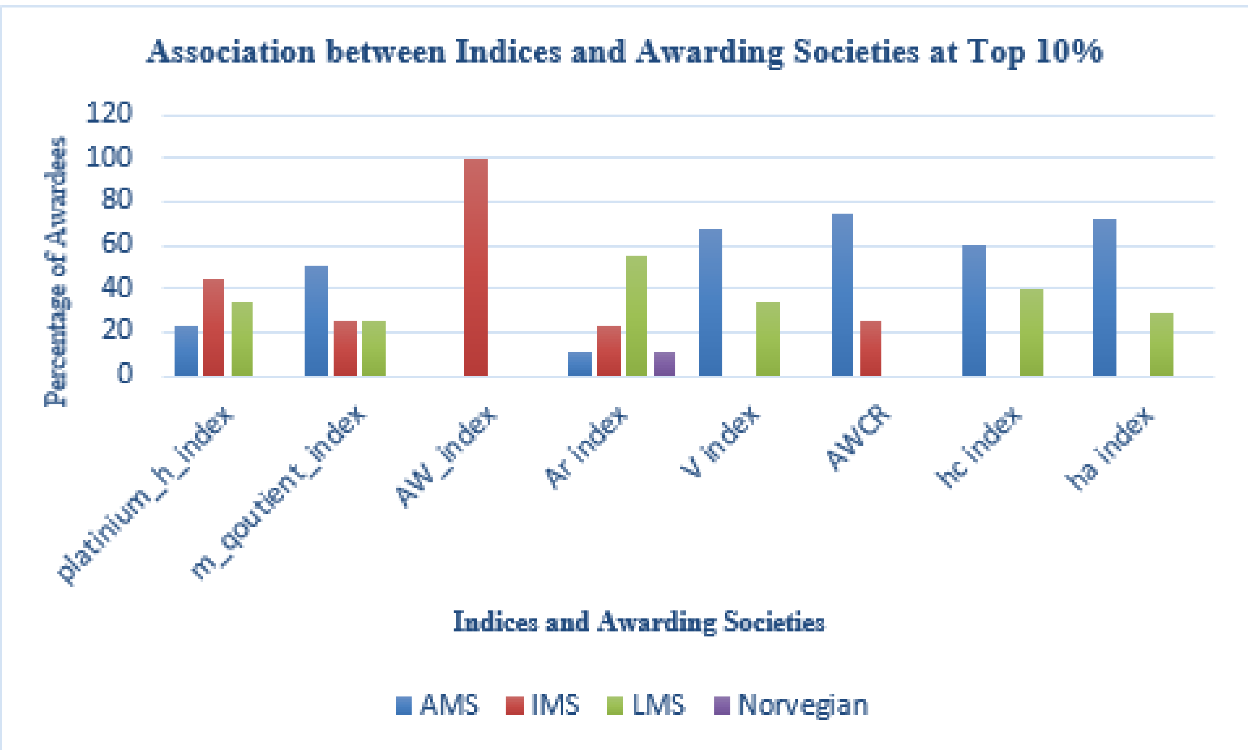

FIGURE 4. Variations of indices on awardees societies at top 10%.

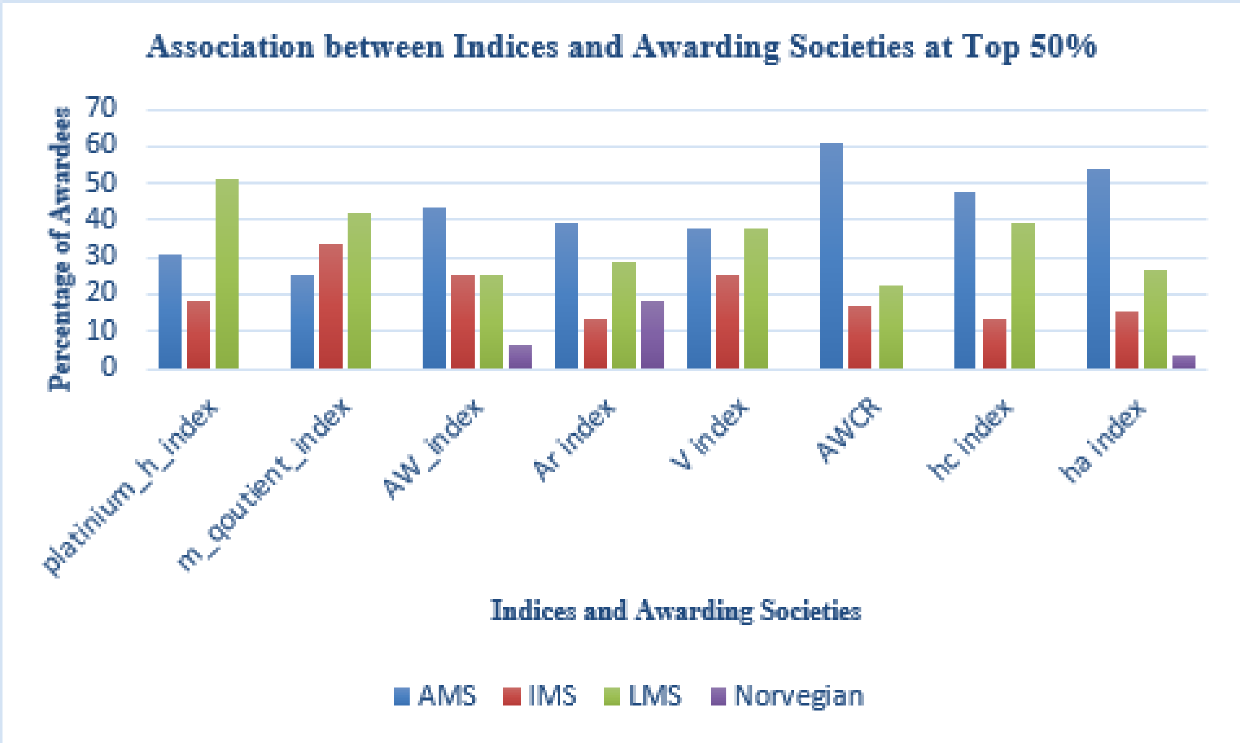

FIGURE 5. Variations of indices on awardees societies at top 50%.

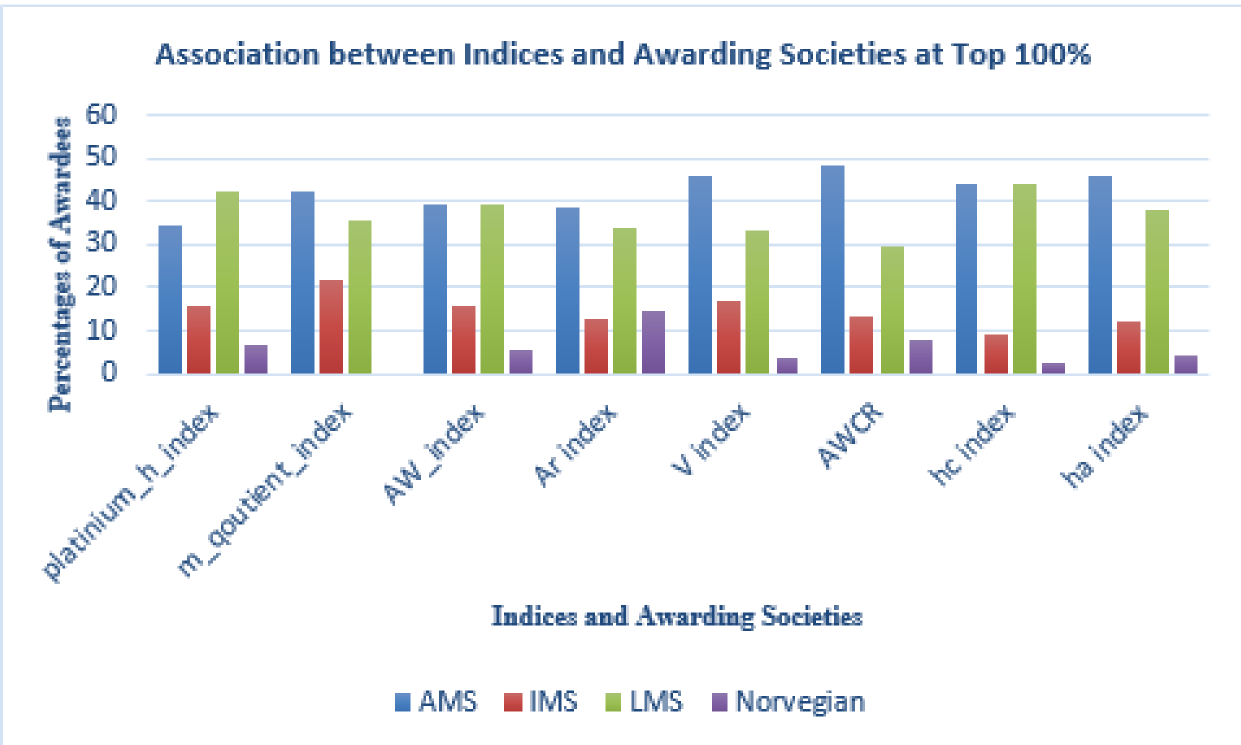

FIGURE 6. Variations of indices on awardees societies at top 100%.

of up to 44%. The Ar index shows poor performance of up to 22%, while the remaining indices return no awardees.

- In the top 50% of the results (See Figure 5), the m quotient index outperformed all others by 33%. The AW and V indices performed the same by retrieving 25% of awardees. The Hc and Ar indices show the lowest performance by retrieving 13% of awardees.
- In the 100% results (See Figure 6), the Aw index and AWCR outperformed all other indices by retrieving 16% of awardees. The Hc index showed the lowest performance by retrieving 9% of awardees.

3) LMS

- In the top 10% of the results (See Figure 4), the Ar index outperformed all other indices by retrieving 55% of awardees. The platinum index and the v index perform equally by retrieving 33% of the awardees. The AW and AWCR indices yielded no results.
- In the top 50% of the results (See Figure 5), the platinum index returns 51%, whereas the m quotient returns 41%. AWCR shows the lowest performance by retrieving 22% of the awardees.
- In the top 100% of results (See Figure 6), platinum and the Hc index retrieve 43% of awardees. The AW index shows the average performance by retrieving 39% of awardees. AWCR shows the lowest performance, retrieving 29% of the awardees.

4) NASL

- In the top 10% of the results (See Figure 4), the Ar index retrieved only 11% of the awardees, while the remaining indices returned no results.
- In the top 50% of the results (See Figure 5), the Ar index retrieved awardees of up to 18%, whereas the AW index retrieved up to 6%. The Ha index demonstrates a performance below 4%, whereas the remaining index returns no awardees.
- In the top 100% of the results (See Figure 6), the Ar index outperformed all other indices by up to 14%. The AWCR index retrieved up to 8%, whereas the platinum-h index retrieved 6%. The Hc index shows the lowest performance by retrieving 2% awardees and the m quotient index returns nothing.

After conducting a comprehensive analysis, we determined that certain indices are more appropriate for certain societies than others. For instance, the AR index outperforms other indices in all ranked lists, making it better suited for NASL society. Additionally, the platinum h-index is more appropriate for the LMS society, the AW index is better suited for the IMU, and the AWCR index is more suitable for the AMS society. Additionally, from the analysis, we have also identified that the highest number of awardees at the top of all ranked lists is from the American Mathematical Society (AMS), suggesting that AMS may depend on these indices to determine its award recipients. This analysis offers valuable insights into the strengths and weaknesses of various societies and their approaches to selecting and recognizing exceptional individuals in their respective fields. By identifying societies that are most effective at utilizing specific indices, researchers and analysts can develop more targeted and efficient strategies for predicting and honoring awardees. Furthermore, this analysis can aid in identifying areas where additional attention and resources may be necessary to ensure that deserving individuals receive recognition for their contributions, thus informing future improvements to the awarding and recognition process.

V. CONCLUSION

Evaluating scholars in the scientific community holds significant importance due to its numerous benefits. Researchers have proposed several parameters to identify influential scholars within specific domains. This research aims to provide a comprehensive evaluation of publication age-based parameters for assessing scholars. To evaluate these parameters, we utilized a dataset consisting of awardees and non-awardees in the field of mathematics, spanning two decades from 1990 to 2023. The collection of this dataset posed a major challenge, as we had to visit the societies' websites and extract awardees' data from the past two decades. Once the dataset was collected, the subsequent step involved evaluating our methodology using the collected data. In the initial stage of our experiments, we computed the correlation between different indices. The purpose was to examine the relationship among these indices. We observed that the majority of the indices demonstrated strong positive correlations, suggesting a consistent pattern in their rankings. However, a few indices displayed negative correlations, indicating a divergence in their rankings compared to the other indices. Next, we proceeded to analyze the occurrence of award winners within the ranked lists of each index, focusing on different percentile ranges such as the top 10%, 20%, 60%, 80%, and 100%. This analysis aimed to uncover valuable insights and trends regarding the performance of awardees in relation to each index. The results obtained from this analysis provided us with valuable information about the association between index rankings and the likelihood of receiving awards. Additionally, we conducted an analysis to explore the relationship between indices and awarding societies. As mentioned in the manuscript, we extracted data on awardees from four renowned and prestigious mathematics societies. The purpose of this analysis was to gain insights into whether these societies rely on specific indices when awarding researchers. By identifying the awarding societies that effectively utilize certain indices, researchers and analysts can develop more targeted and efficient strategies for predicting and acknowledging awardees. Furthermore, this analysis can contribute to improving the awarding and recognition process by identifying areas that require additional attention and resources. This ensures that the most deserving individuals are appropriately recognized and honored for their contributions. The findings of this study revealed that the Ar index outperformed the other evaluated indices, as it brought the highest number of awardees to the top 10% of the ranked list, reaching a maximum of 80%. However, it is important to note that none of the indices were able to include all the awardees in the top rankings, indicating that no single index can offer a complete and accurate evaluation of author rankings. Furthermore, the study identified that the American Mathematical Society (AMS) had the highest number of awardees at the top of all the ranked lists. This suggests that AMS may rely on these indices when selecting their award recipients. Additionally, the AWCR index was found to be particularly well-suited for the AMS society, as it returned the highest number of awardees affiliated with the AMS.

Based on the current study, we have identified that the Ar index performs well in the field of mathematics. However, various studies in the literature present different findings. For instance, Madiha et al in 2019 suggested that the h index performs well in the domain of neuroscience, while another study conducted by Raheel et al in 2018 stated that the Wu index yields better results for civil engineering data, and so on. Thus far, researchers have not reached a consensus on a single parameter that can be universally applied. Now, there is an open challenge exist for research community to evaluate over more than seventy indices that have been proposed in the literature, belonging to different categories. This evaluation should be conducted on a vast and comprehensive dataset encompassing multiple domains of data to identify the most influential research assessment parameter which can be applied globally.

VI. FUTURE WORK

In addition to the category of indices analyzed in this study, other indices belonging to different categories have been proposed by researchers and ranking communities. To assess their performance, it is significant to evaluate these indices on a comprehensive dataset of other domains.

ACKNOWLEDGMENT

The authors would like to thank the Qatar National Library for their generous support in providing Open Access funding for this research. They also like to thank Al-Ahliyya Amman University and the University of Engineering and Technology Peshawar, Pakistan, for their invaluable assistance and support, which greatly facilitated the execution of this study. Their contributions were instrumental in the successful completion of this research work.

REFERENCES

[1] M. Raheel, S. Ayaz, and M. T. Afzal, "Evaluation of h-index, its variants and extensions based on publication age & citation intensity in civil engineering," *Scientometrics*, vol. 114, no. 3, pp. 1107–1127, Mar. 2018.

[2] S. Ayaz and M. T. Afzal, "Identification of conversion factor for completing-h index for the field of mathematics," *Scientometrics*, vol. 109, no. 3, pp. 1511–1524, Dec. 2016.

[3] G. Mustafa, A. Rauf, B. Ahmed, M. T. Afzal, A. Akhunzada, and S. Z. Alharthi, "Comprehensive evaluation of publication and citation metrics for quantifying scholarly influence," *IEEE Access*, vol. 11, pp. 65759–65774, 2023.

[4] G. Mustafa, M. Usman, M. T. Afzal, A. Shahid, and A. Koubaa, "A comprehensive evaluation of metadata-based features to classify research paper's topics," *IEEE Access*, vol. 9, pp. 133500–133509, 2021.

[5] G. Mustafa, M. Usman, L. Yu, M. T. Afzal, M. Sulaiman, and A. Shahid, "Multi-label classification of research articles using Word2Vec and identification of similarity threshold," *Sci. Rep.*, vol. 11, no. 1, p. 21900, Nov. 2021.

[6] K. Balog, L. Azzopardi, and M. de Rijke, "Formal models for expert finding in enterprise corpora," in *Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2006, pp. 43–50.

[7] D. H. L. Cameron, B. Aleman-Meza, S. Decker, and I. B. Arpinar, "SEMEF: A taxonomy-based discovery of experts, expertise and collaboration networks," Ph.D. dissertation, Dept. Comput. Sci., Univ. Georgia, Atlanta, GA, USA, 2007.

[8] H. Deng, J. Han, M. R. Lyu, and I. King, "Modeling and exploiting heterogeneous bibliographic networks for expertise ranking," in *Proc. 12th ACM/IEEE-CS Joint Conf. Digit. Libraries*, Jun. 2012, pp. 71–80.

[9] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 46, pp. 16569–16572, Nov. 2005.

[10] A. Bihari, S. Tripathi, and A. Deepak, "A review on h-index and its alternative indices," *J. Inf. Sci.*, vol. 49, no. 3, pp. 624–665, 2021.

[11] Q. L. Burrell, "On the h-index, the size of the Hirsch core and Jin's A-index," *J. Informetrics*, vol. 1, no. 2, pp. 170–177, Apr. 2007.

[12] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The R- and AR-indices: Complementing the h-index," *Chin. Sci. Bull.*, vol. 52, no. 6, pp. 855–863, Mar. 2007.

[13] N. R. Khan, C. J. Thompson, D. R. Taylor, K. S. Gabrick, A. F. Choudhri, F. R. Boop, and P. Klimo, "Part II: Should the *h*-index be modified? An analysis of the *m*-quotient, contemporary *h*-index, authorship value, and impact factor," *World Neurosurg.*, vol. 80, no. 6, pp. 766–774, Dec. 2013.

[14] D. Katsaros, L. Akritidis, and P. Bozanis, "Thefindex: Quantifying the impact of coterminal citations on scientists' ranking," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 60, no. 5, pp. 1051–1056, May 2009.

[15] F. J. Cabrerizo, S. Alonso, E. Herrera-Viedma, and F. Herrera, "Q2-index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core," *J. Informetrics*, vol. 4, no. 1, pp. 23–28, Jan. 2010.

[16] M. Schreiber, "A modification of the h-index: The hm-index accounts for multi-authored manuscripts," *J. Informetrics*, vol. 2, no. 3, pp. 211–216, Jul. 2008.

[17] A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos, "Generalized Hirsch h-index for disclosing latent facts in citation networks," *Scientometrics*, vol. 72, no. 2, pp. 253–280, Aug. 2007.

[18] L. Egghe and R. Rousseau, "An h-index weighted by citation impact," *Inf. Process. Manage.*, vol. 44, no. 2, pp. 770–780, Mar. 2008.

[19] L. Bornmann, R. Mutz, S. E. Hug, and H.-D. Daniel, "A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants," *J. Informetrics*, vol. 5, no. 3, pp. 346–359, Jul. 2011.

[20] S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "h-index: A review focused in its variants, computation and standardization for different scientific fields," *J. Informetrics*, vol. 3, no. 4, pp. 273–289, Oct. 2009.

[21] R. Ghani, F. Qayyum, M. T. Afzal, and H. Maurer, "Comprehensive evaluation of *h*-index and its extensions in the domain of mathematics," *Scientometrics*, vol. 118, pp. 809–822, Jan. 2019.

[22] J. Ding, C. Liu, and G. A. Kandonga, "Exploring the limitations of the h-index and h-type indexes in measuring the research performance of authors," *Scientometrics*, vol. 122, no. 3, pp. 1303–1322, Mar. 2020.

[23] Y. Liu and Z. Sun, "S-index: A new measure to evaluate the scientific impact of scholars," in *Proc. 3rd Int. Conf. Adv. Electron. Mater., Comput. Softw. Eng. (AEMCSE)*, Apr. 2020, pp. 529–535.

[24] R. Dillon, "The u-index: An alternative metric to measure academic impact," in *Proc. XII Int. Conf. Virtual Campus (JICV)*, Sep. 2022, pp. 1–3.

[25] M. Kosmulski, "A new Hirsch-type index saves time and works equally well as the original h-index," *ISSI Newslett.*, vol. 2, no. 3, pp. 4–6, 2006.
[26] A. F. J. van Raan, "Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups," *Scientometrics*, vol. 67, no. 3, pp. 491–502, Jun. 2006.
[27] S. Xiao, J. Yan, C. Li, B. Jin, X. Wang, X. Yang, S. M. Chu, and H. Zha, "On modeling and predicting individual paper citation count over time," in *Proc. IJCAI*, 2016, pp. 2676–2682.
[28] M. Schreiber, "Self-citation corrections for the Hirsch index," *Europhys. Lett.*, vol. 78, no. 3, p. 30002, May 2007.
[29] K. Mane and U. Shrawankar, "An improved indexing technique for tribal art retrieval system," in *Proc. IEEE Int. Students' Conf. Electr., Electron. Comput. Sci. (SCEECS)*, Feb. 2023, pp. 1–5.
[30] M. Ameer and M. T. Afzal, "Evaluation of h-index and its qualitative and quantitative variants in neuroscience," *Scientometrics*, vol. 121, no. 2, pp. 653–673, Nov. 2019.
[31] Q.-U. Ain, H. Riaz, and M. T. Afzal, "Evaluation of h-index and its citation intensity based variants in the field of mathematics," *Scientometrics*, vol. 119, no. 1, pp. 187–211, Apr. 2019.
[32] M. Usman, G. Mustafa, and M. T. Afzal, "Ranking of author assessment parameters using logistic regression," *Scientometrics*, vol. 126, no. 1, pp. 335–353, Jan. 2021.
[33] A. A. Alshdadi, M. Usman, M. O. Alassafi, M. T. Afzal, and R. AlGhamdi, "Formulation of rules for the scientific community using deep learning," *Scientometrics*, vol. 128, no. 3, pp. 1825–1852, Mar. 2023.
[34] A. Martín-Martín, M. Thelwall, E. Orduna-Malea, and E. D. López-Cózar, "Google scholar, Microsoft academic, scopus, dimensions, web of science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations," *Scientometrics*, vol. 126, no. 1, pp. 871–906, Jan. 2021.
[35] X. Jiang, X. Sun, and H. Zhuge, "Graph-based algorithms for ranking researchers: Not all swans are white!" *Scientometrics*, vol. 96, no. 3, pp. 743–759, Sep. 2013.

GHULAM MUSTAFA received the B.S. degree (Hons.) in software engineering from COMSATS University Islamabad (CUI), Abbottabad Campus, in 2017, and the M.S. degree (Hons.) in computer science from the Capital University of Science and Technology, Islamabad. He is currently pursuing the Ph.D. degree in computer science from the University of Engineering and Technology Taxila, Taxila, Pakistan. He was associated with academia and industry for the last six years. He is currently a Senior Lecturer with the Faculty of Computing, Shifa Tameer-e-Millat University, Islamabad. Previously, he served five years with the CS Department, Capital University of Science and Technology, as an Associate and a Junior Lecturer. Furthermore, he was with A&F Solution Software House, as a Web Frontend Designer and a Backend Developer. Moreover, with academia he is also an active freelancer for the last four years, doing his projects in different languages, such as Python, Java, and C++. In his academic career, he has taught different computer science labs, such as the Introduction to Programming Lab (C++), Object Oriented Programming Lab (C++), Advanced Computer Programming Lab (JAVA), Database System Lab, Data Structure Lab (C++), Computer Communication and Network Lab (CNN), and the Web Development Lab, while serving as a Junior Lecturer. Moreover, he also taught different computer science courses, such as the Introduction to Programming, Object Oriented Programming, Discrete Structure, Theory of Automata, and Software Engineering, while serving as a Senior and an Associate Lecture. He received two Gold Medals for the B.S. degree from CUI, Abbottabad Campus, and the Gold medal for the M.S. degree in computer science from the Capital University of Science and Technology.

ABID RAUF received the M.Sc. degree in statistics from Quaid-i-Azam University, Islamabad, Pakistan, in 2000, the M.S. degree in information security from Sichuan University, Chengdu, China, in 2007, and the Ph.D. degree from the National University of Sciences and Technology (NUST), Islamabad, in 2021. He is currently a Lecturer with the Department of Computer Science, University of Engineering and Technology Taxila, Taxila, Pakistan. He has expertise in the area of information security and privacy and statistical methods for machine learning. His research interests include cryptographic protocols and algorithms, security in the Internet of Things (IoT), security in wireless sensor networks and privacy, machine learning and privacy, and covert communications.

AHMAD SAMI AL-SHAMAYLEH received the master's degree in information systems from The University of Jordan, Jordan, in 2014, and the Ph.D. degree in artificial intelligence from the University of Malaya, Malaysia, in 2020. He is currently an Assistant Professor with the Faculty of Information Technology, Al-Ahliyya Amman University, Jordan. His research interests include artificial intelligence, human–computer interaction, the IoT, Arabic NLP, Arabic sign language recognition, language resources production, the design and evaluation of interactive applications for handicapped people, multimodality, and software engineering.

BILAL AHMED (Graduate Student Member, IEEE) received the master's degree in Pakistan. He is currently pursuing the Ph.D. degree with the Taiyuan University of Technology, China. He was associated with academia and industry for the last seven years. In his academic career, he taught different computer science courses, such as machine learning, data science, and information visualization, and also courses on various programming languages, such as Python and C++. His research interests include machine learning, deep learning, data science, scientometrics, and recommendation systems.

WAGDI ALRAWAGFEH received the Ph.D. degree in computer science (multi-agent systems) the from Memorial university of Newfoundland, Canada. He is currently working as an Assistant professor with the College of Computing & IT, University of Doha for Science and Technology, Qatar. He has more than ten years of experience working with environmental aspects directly related to education, programming, and information technology. Through his experiences working as a Teacher Assistant, a Research Assistant and Software Developer at Memorial University, Canada, and as a Teacher at Al Isra Private University, Jordan, he gained a variety of skills in imparting knowledge regarding both theoretical and practical topics of information technology. During his experience as a computer science teacher, he is successfully communicated the material to master's students and used several teaching techniques to overcome any challenges during the class. His continuous experience in the IT field has enhanced his aspiration to utilize and apply his IT knowledge to business solutions and computerizing business processes. As a result, this aspiration motives him to encourage the students to apply what they learn in the class to solve or automate the business processes in daily life.

MUHAMMAD TANVIR AFZAL received the M.Sc. degree (Hons.) in computer science from Quiad-i-Azam University, Islamabad, Pakistan, and the Ph.D. degree (Hons.) in computer science from the Graz University of Technology, Austria. He was associated with academia and industry at various levels for the last 20 years. He was with the Capital University of Science and Technology, Islamabad, as a Professor, an Associate Professor, and an Assistant Professor of CS. He is currently the Director and a Professor with the Shifa School of Computing and Director Campus at one of the largest campus of Shifa Tameer-e-Millat University, Islamabad. Previously, he was the Director and a Professor with the Namal Institute Mianwali. Furthermore, he was with NESCOM, COMSATS University Islamabad, and Jin Technologies, Islamabad. He has authored more than 125 research articles, including more than 50 published articles in impact factor leading journals in the field of data science, information retrieval and visualization, semantics, digital libraries, artificial intelligence, and scientometrics. He has authored two books and has edited two books in computer science. His cumulative impact factor is more than 123, with citations of more than 1300. He played pivotal role in making collaborations between MAJU-JUCS, MAJU-IICM, and TUG-UNIMAS. He conducted more than 100 curricular, co-curricular, and extra-curricular activities in the last five years, including seminars, workshops, and national competitions (ExcITeCup), and invited international and national speakers from Google, Oracle, IICM, IFIS, and SEGA Europe. Under his supervision, more than 70 postgraduate students (M.S. and Ph.D.) have defended their research theses successfully and a number of Ph.D. and M.S. students are pursuing their research with him. He served as a Master Trainer and the Program Director at a national level training for a public sector organization in Pakistan on human factors engineering and conducted the training of more than 150 hours for experts from the industry. He remained a Master Trainer for data science trainings in leading industrial organizations. He was a recipient of multiple international research fundings. He served as the Ph.D. Symposium Chair, the Session Chair, the Finance Chair, a Committee Member, and an Editor for several IEEE, ACM, Springer, and Elsevier international conferences and journals, respectively. He is an Editor-in-Chief for reputed impact factor journal, such as *Journal of Universal Computer Science*.

ADNAN AKHUNZADA (Senior Member, IEEE) is an accomplished Cybersecurity Specialist and a Consultant, boasting extensive industrial experience and expertise. With more than a decade of experience in the field, he is highly regarded as a Professional Member of ACM. He has made significant contributions to the cybersecurity landscape through impactful published research, successful commercial products, and holding U.S. patents. He has advised some of the largest ICT companies worldwide, effectively securing multi-million-dollar projects. As a recognized authority in the field, his expertise spans the design of innovative SIEM systems, IDS, IPS, threat intelligence platforms, secure protocols, AI for cybersecurity, secure future internet, adversarial, and privacy preserving machine learning. With a combination of proven industry success, technical prowess, and a dedication to advancing cybersecurity, he remains at the forefront of the field, continuously shaping its future.