Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776483

A CORRELATION ANALYSIS OF CLASSICAL H-TYPE INDICES AND ALTMETRIC H-TYPE INDICES FOR TOP TEN HIGHLY CITED AUTHORS IN ARTIFICIAL INTELLIGENCE

Rupesh Kumar A.
Assistant Professor, Department of Studies and Research in Library and Information Science, Tumkur University, Tumakuru, Karnataka, e-mail: a.rupeshkumar@gmail.com
ORCiD: 0000-0001-8787-6416

Abstract
The present study investigated the correlation between classical h-index (h) and Altmetric h-index (halt), classical g-index (g) and Altmetric g-index (galt) values of top ten highly cited authors in the field of artificial intelligence. Altmetric counterparts of h-index and g-index were calculated by using the same method as proposed for the original indices, but based on altmetric attention scores (AAS) in place of citation counts. While publication data and citation counts were collected from Web of Science database, altmetric attention scores were collected by using Webometric Analyst software. The indices were manually computed. Kendall's tau correlation test was conducted on h and halt, and on g and galt values. Statistically significant very high degree of positive correlation was found between h and halt (tau = 0.9189, p = 0.0004). Correlation between g and galt showed even higher degree of positivity with statistical significance (tau = 0.9775, p = 0.0001). The current literature on metric studies primarily focuses on analysing the correlation between citation counts and altmetric attention scores, confining to article-level metrics. This study attempted to measure the correlation between two well-known author-level metrics: h-index, g-index and their altmetric counterparts. It can be expanded by modelling the relationship between classical indices and their altmetric indices to understand the influence of Altmetrics on classical metrics.

Keywords: Artificial intelligence, Citation metrics, H-index, Altmetric h-index, G-index, Altmetric g-index, Kendall's Correlation

1. Introduction

In modern day publishing, research impact has become an increasingly important metric for measuring the significance of scientific and scholarly work. Research impact refers to the influence that research has on various stakeholders, including the academic community, policy makers, industry professionals, and the general public. The impact of research is typically measured by a variety of metrics, such as citation counts, social media engagement, media coverage, and policy or industry uptake. The significance of research impact lies in its ability to provide a quantitative measure of the quality and relevance of research, as well as its ability to inform decision-making processes and contribute to the advancement of knowledge. As a result, researchers and institutions are increasingly emphasizing the importance of research impact when assessing the value of scientific and scholarly work.

There are several metrics that are commonly used to measure the research impact of authors and journals. Citation counts, Journal Impact Factor (JIF) [1], CiteScore [2], h-index [3], g-index [4], i10 index [5] are noteworthy among them.

2. H-index

One of the most widely used metrics for measuring research impact is the h-index. The h-index was introduced by physicist Jorge Hirsch in 2005 as a way to measure both the productivity and impact of a researcher's work [3]. The h-index is based on the number of publications a researcher has published and the number of citations each publication has received. The h-index is often used to evaluate the impact of individual researchers, departments, or institutions, and it has become a standard metric in many academic fields. The significance of the h-index lies in its ability to provide a quantitative measure of the impact of research that can be used for funding, hiring, and promotion decisions. Despite the availability of other research metrics, the h-index remains one of the most widely used measures of research impact. This is partly due to its simplicity and ease of calculation, but also because it provides a more comprehensive view of a researcher's impact by considering both the number of publications and the number of citations they have received.

3. Altmetrics

In addition to the h-index, there are several other research metrics that are commonly used to measure research impact, such as the impact factor, citation counts, and Altmetrics. Altmetrics, which include social media mentions, online downloads, and other alternative measures of attention and impact, are becoming increasingly popular as a way to capture the broader reach and impact of research beyond traditional academic channels [6].

Altmetrics, also known as alternative metrics, have emerged as a complementary approach to traditional research metrics for measuring the impact and attention of scholarly work. Altmetrics encompass a wide range of non-traditional measures such as social media mentions, news articles, policy documents, and online downloads. Unlike traditional metrics such as citation counts or impact factors, which focus on academic citations and journals, altmetrics aim to capture a more diverse and real-time view of research impact by tracking the dissemination and engagement of research across various platforms. As scholarly communication continues to evolve in the digital age, altmetrics have become increasingly important for assessing the broader societal impact and relevance of research beyond traditional academic channels.

The present study attempts to measure the correlation between classical h-index (h) and altmetric h-index (halt); traditional g-index (g) and altmetric g-index (galt) taking into account the citation counts and altmetric attention scores (AAS) of top ten highly cited authors in the field of artificial intelligence.

4. Methods

4.1. Publication and Citation data

Web of Science Core collection was searched using the subject category “Computer Science, Artificial Intelligence”. The results were sorted in the descending order of citations. Top ten highly cited authors were identified. These authors' publication details and citation counts were retrieved from Web of Science for further analysis. Articles and reviews under the WoS Category “Computer Science, Artificial Intelligence” were considered and other types of publications were excluded from analysis. The results for each author were exported as comma separated values (csv) file format.

4.2. Altmetric Attention Score (AAS)

For each publication, AAS was collected using the Webometric Analyst software[7]. Digital Object Identifier (DOI) available in each publication data was used for finding the AAS. Publications not having DOI were excluded from analysis.

4.3. Computation of indices

From the data collected as explained above, classical h-index, classical g-index, altmetric h-index, and altmetric g-index were computed.

4.4. Correlation Test

In order to understand the strength and direction of relationship between two sets of variables, namely, classical h-index & altmetric h-index, and classical g-index and altmetric g-index, Kendall's rank correlation test (tau)[8] was conducted. Microsoft Excel was used for data processing and R statistical analysis software was used for correlation test.

5. Results and Discussion

5.1. Distribution of citations

Table 1 shows the distribution of publications and citation counts of top ten highly cited authors in the field of artificial intelligence.

Table 1. Top ten highly cited authors in Artificial Intelligence

Rank	Author's Name	Total Number of citations	Number of publications with at least one citation	Number of publications with zero citation	Total number of publications	Average citations per publication
1	Van Gool, Luc	34263	109	4	113	303
2	Yan, Shuicheng	14330	166	1	167	86
3	Li, Xuelong	22821	316	1	317	72
4	Tao, Dacheng	22237	314	4	318	70
5	Cao, Jinde	21708	346	1	347	63
6	Zhang, Mengjie	6817	120	4	124	55
7	Abraham, Ajith	6728	201	7	208	32
8	Pedrycz, Witold	20165	610	19	629	32
9	Jiao, Licheng	10138	305	13	318	32
10	Hancock, Edwin	5190	200	39	239	22
	Total	**164397**	**2687**	**93**	**2780**	**59**

Table 1 shows the distribution of publications and citation counts. Authors have been ranked in the descending order of average citations per publication. Van Gool tops the list with the average of 303 citations per publication, while Hancock stands at rank 10 with average citations of 22. In total, the top ten authors have garnered 164397 citations for 2780 publications, averaging to 59 citations per publication. There were 93 publications with no citations.

5.2. H-type indices

Two h-type indices, namely, h-index and g-index were computed on the basis of publication and citation data collected under the present study. The data is presented in Table 2 in the ranked order of h-index.

Table 2. Classical H-index and G-index of Top ten highly cited authors in Artificial Intelligence

Rank	Author's Name	H-index	G-index
1	Cao, Jinde	84	126
2	Tao, Dacheng	80	138
3	Li, Xuelong	79	137
4	Pedrycz, Witold	69	106
5	Yan, Shuicheng	53	118
6	Jiao, Licheng	48	84
7	Van Gool, Luc	46	113
8	Abraham, Ajith	40	75
9	Zhang, Mengjie	39	80
10	Hancock, Edwin	38	63

In terms of h-index, Cao, Jinde topped the list with h-index of 84, followed by Tao, Dacheng with h-index of 80. However, Tao, Dacheng had the highest g-index of 138, closely followed by Li, Xuelong with g-index of 137. H-index ranged between 34 and 84, while the range of g-index was between 63 and 138. It is worth noting that Van Gool, who had the highest number of citations in the previous table, has a relatively lower h-index of 46 and g-index of 113 compared to some of the other authors. This suggests that while Van Gool may have many highly cited papers, the impact of their overall research output may be lower than some of the other authors. Overall, the data indicates that these authors have made significant contributions to their respective fields, with Cao having the highest h-index and Tao having the highest g-index. Figure 1 graphically presents a comparison between h-index and g-index.

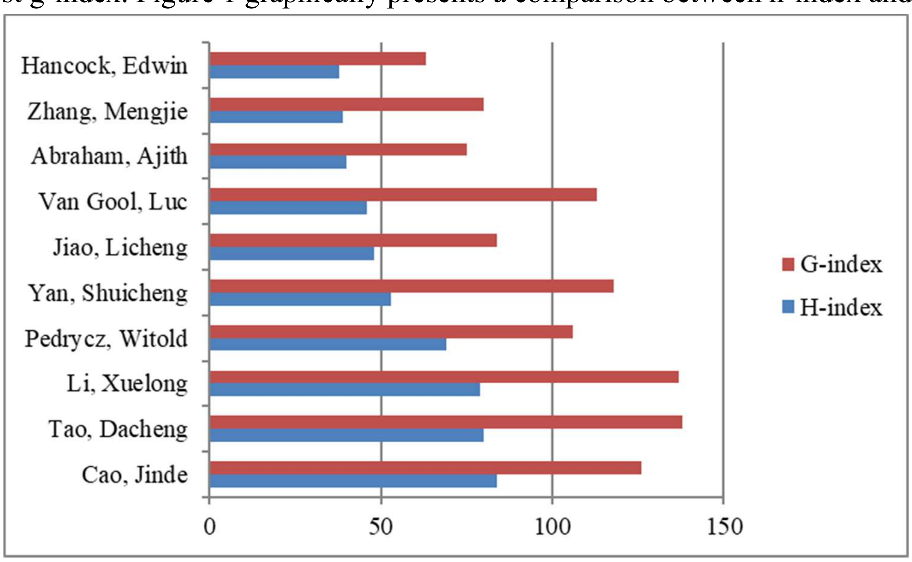

Fig. 1. Comparison between h-index and g-index

5.3. Altmetric h-type indices

The altmetric attention scores for each publication of the ten authors, having valid DOI were collected using the Webometric Analyst. Altmetric h-index and Altmetric g-index values were computed by using this data. Publications not having DOI were excluded. Table 3 shows the data on Altmetric h-type indices.

Table 3. Altmetric H-index and G-index of Top ten highly cited authors in Artificial Intelligence

Author's Name	Total number of publications	Number of publications without DOI	Number of publications with Altmetric Score	% of publications with Altmetric Score	Altmetric H-index	Altmetric G-index
Cao, Jinde	347	11	23	6.63	3	3
Tao, Dacheng	318	1	106	33.33	7	14
Li, Xuelong	317	1	89	28.08	6	7
Pedrycz, Witold	629	35	47	7.47	4	6
Yan, Shuicheng	167	2	60	35.93	6	9
Jiao, Licheng	318	23	31	9.75	3	4
Van Gool, Luc	113	19	51	45.13	9	14
Abraham, Ajith	208	53	32	15.38	4	8
Zhang, Mengjie	124	7	31	25.00	4	5
Hancock, Edwin	239	87	31	12.97	6	7
Total	**2780**	**239**	**501**	**18.02**		

It is evident from the data presented in the above table that only 18.02% of publications had the Altmetric attention score. Tao had the highest number of publications (106) with Altmetric attention score while Cao had the lowest (23). Van Gool had the highest Altmetric h-index of 9 whereas Tao and Van Gool both had the highest Altmetric g-index of 14. 239 publications did not have DOI and therefore, were excluded from the computation.

5.4. Correlation

The purpose of the present study was to assess the strength and direction of relationship between classical h-type indices and altmetric h-type indices. Kendall's rank correlation (tau) test was conducting using R statistical analysis software. Table 4 shows the results of the test.

Table 4. Kendall's rank correlation between classical and Altmetric h-type indices

	Altmetric h-index		Altmetric g-index	
Classical h-index	*tau =0.9189*	*p =0.0004*	---	
Classical g-index	---		*tau =0.9775*	*p = 0.0001*

It is evident from the above table that correlation between classical h-index and Altmetric h-index is very highly positive (tau = 0.9189, p-value = 0.0004) with statistical significance. Classical g-index and Altmetric g-index too have statistically significant very high degree of positive correlation (tau = 0.9775, p-value = 0.0001). This clearly shows that there is high degree of positive correlation between classical h-type indices and their altmetric counterparts.

6. Conclusion

Current body of literature on metric studies primarily focus on investigating the relationship between citation counts and altmetric attention scores, mostly confining to article-level metrics. The present study has extended the scope by exploring the relationship between classical and altmetric h-type indices, focusing on author level metrics. Artificial intelligence has been demonstrated as an illustration. Similar studies can be carried out in various fields. Further, it can be expanded by modelling the relationship between classical indices and their altmetric indices to understand the influence of Altmetrics on classical metrics.

References

[1] Clarivate, "The Clarivate Analytics Impact Factor," Web of Science Group, 2022. https://clarivate.com/webofsciencegroup/essays/impact-factor/ (accessed Aug. 26, 2023).

[2] H. Zijlstra and R. McCullough, "CiteScore: a new metric to help you track journal performance and make decisions," Elsevier, Dec. 07, 2017. https://www.elsevier.com/connect/editors-update/citescore-a-new-metric-to-help-you-choose-the-right-journal (accessed Feb. 06, 2022).

[3] J. E. Hirsch, "An index to quantify an individual's scientific research output," Proc. Natl. Acad. Sci., vol. 102, no. 46, pp. 16569–16572, Nov. 2005, doi: 10.1073/pnas.0507655102.

[4] L. Egghe, "Theory and practise of the g-index," Scientometrics, vol. 69, no. 1, pp. 131–152, Oct. 2006, doi: 10.1007/s11192-006-0144-7.

[5] J. Connor, "Google Scholar Citations Open To All," Google Scholar Blog, Nov. 16, 2011. https://scholar.googleblog.com/2011/11/google-scholar-citations-open-to-all.html (accessed Jan. 04, 2022).

[6] Altmetric, "What are altmetrics?," Altmetric, 2022. https://www.altmetric.com/about-us/what-are-altmetrics/ (accessed Jan. 05, 2022).

[7] M. Thelwall, Web Indicators for Research Evaluation: A Practical Guide. San Rafael, CA: Morgan & Claypool, 2017. doi: 10.1007/978-3-031-02304-0.

[8] M. G. Kendall, "A New Measure of Rank Correlation," Biometrika, vol. 30, no. 1/2, pp. 81–93, 1938, doi: 10.2307/2332226.