Comparative Citation Analysis of Highly CitedResearchers in Web of Science and Google Scholar

Rouhallah Khademi
Assistant Prof., Knowledge and Information Science Department
Semnan University, Semnan, Iran
r.khademi@semnan.ac.ir

Mansoureh Serati Shirazi
Assistant Prof., Islamic World Science & Technology
Monitoring and Citation Institute (ISC)
Research Department of Scientometrics
Shiraz, Iran
Serati@isc.ac

ABSTRACT: *This study examines the correlation between citation counts in Google Scholar (GS) and Web of Science (WoS) for highly cited researchers across various disciplines. Although citation counts have almost been debated as a measure of research quality, they remain essential for evaluating academic impact. analyzing 2023 Highly Cited Researchers list from Clarivate™, citation trends were tracked over ten years using WoS and Google Scholar via "Publish or Perish" software. Pearson's correlation coefficient indicated a significant correlation between citations in both databases. Additionally, an independent-sample t-test revealed no significant difference in mean citations between GS and WoS. Exponential regression showed nearly identical growth rates for citations in both databases. The findings suggest that while Google Scholar captures more diverse citations, it correlates well with Web of Science, supporting its use as a valuable, free alternative for citation analysis.*

Keywords: Comparative Citation Analysis, Web of Science, Google Scholar

Received: 2 August 2024, Revised 12 September 2024, Accepted 27 September 2024

DOI: https://doi.org/10.6025/stm/2024/5/240-245

1. Introduction

Although many scholars have debated the use of citations to assess research quality, the general use of citations for evaluating research is based on the assumption that citation counts are an objective measure that credits and recognizes the value, impact, quality, or significance of an author's work.(Meho & Yang, 2006) Citation resources are essential tools for academics to search for, track, and conduct thorough citation analyses across diverse disciplines. Their significance lies in their ability to facilitate comprehensive citation analysis within an academic environment.(Adriaanse & Rensleigh, 2011). One of the tools to access citation datais citation indexes.

Eugene Garfield enabled the widespread use of citation analysis in academia by developing three citation indices: Science, Humanities, and Social Science Citation Indices. These were later combined and turned into an electronic version known as the Web of Science (Bakkalbasi et al. 2006). In 2004, Scopus and Google Scholar (GS) were introduced. Citation resources presented possible competition for the Web of Science (WOS), which enjoyed the monopoly for more than 40 years (Felter 2005: 43). Although there are many citation sources, these three databases have been used and studied as the most important citation databases. However, the question is, considering the difference in approaches and sources covered by these databases, is there any correlation and connection between the citations received by each person in the Google Scholar and Web of Science databases?

2. Literature Review

Many studies have investigated and compared these citation sources, the most important of which are in table 1. As the data in this table shows, this topic has interested researchers and journals since the first years of Google Scholar and Scopus's emergence until now.

The studies collectively reveal that Web of Science, Scopus, and Google Scholar each offer unique strengths and limitations in citation tracking and analysis. Google Scholar generally captures a broader range of citations due to its extensive scope, including sources outside traditional academic databases. However, Scopus and Web of Science provide more reliable, consistent, and curated citation data. These differences result in varying h-index values and citation counts across platforms, with each database showing specific advantages depending on the field of study and research needs. Using multiple citation databases offers a more comprehensive understanding of scholarly impact and research trends.

Publication Year	Author	Title	Source Title
2005	Pauly, Daniel; Stergiou, Konstantinos I.	Equivalence of results from two citation analyses: Thomson ISI's Citation Index and Google's Scholar service	Ethics in Science and Environmental Politics
2006	Bakkalbasi, Nisa; Bauer, Kathleen; Glover, Janis; Wang, Lei	Three options for citation tracking: Google Scholar, Scopus and Web of Science	Biomedical Digital Libraries
2006	Meho, Lokman I.; Yang, Kiduk	A New Era in Citation and Bibliometric Analyses: Web of Science, Scopus, and Google Scholar	Journal of the American Society for Information Science and Technology
2006	Yang, Kiduk; Meho, Lokman I.	Citation Analysis: A Comparison of Google Scholar, Scopus, and Web of Science	Proceedings of the American Society for Information Science and Technology
2008	Bar-Ilan, Judit	Which h-index? — A comparison of WoS, Scopus and Google Scholar	Scientometrics
2008	Kousha, Kayvan; Thelwall, Mike	Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines	Scientometrics
2010	Mingers, John; Lipitakis, Evangelia A. E. C. G.	Counting the citations: a comparison of Web of Science and Google Scholar in the field of business and management	Scientometrics
2010	Li, Jie; Burnham, Judy F.; Lemley, Trey; Britton, Robert M.	Citation Analysis: Comparison of Web of Science®, Scopus™, SciFinder®, and Google Scholar	Journal of Electronic Resources in Medical Libraries
2011	Adriaanse, Leslie S.; Rensleigh, Chris	Comparing Web of Science, Scopus and Google Scholar from an environmental sciences perspective	South African journal of libraries and information science
2013	Delgado-López-Cózar, Emilio; Repiso-Caballero, Rafael	The impact of scientific journals of communication: Comparing Google Scholar Metrics, Web of Science and Scopus	Comunicar

2013	Harzing, Anne-Wil	A preliminary test of Google Scholar as a source for citation data: a longitudinal study of Nobel prize winners	Scientometrics
2013	S. Adriaanse, Leslie; Rensleigh, Chris	Web of Science, Scopus and Google Scholar: A content comprehensiveness comparison	The Electronic Library
2014	De Winter, Joost C. F.; Zadpoor, Amir A.; Dodou, Dimitra	The expansion of Google Scholar versus Web of Science: a longitudinal study	Scientometrics
2016	Harzing, Anne-Wil; Alakangas, Satu	Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison	Scientometrics
2016	Moed, Henk F.; Bar-Ilan, Judit; Halevi, Gali	A new methodology for comparing Google Scholar and Scopus	Journal of Informetrics
2018	Martín-Martín, Alberto; Orduna-Malea, Enrique; Delgado López-Cózar, Emilio	Coverage of highly-cited documents in Google Scholar, Web of Science, and Scopus: a multidisciplinary comparison	Scientometrics
2018	Martín-Martín, Alberto; Orduna-Malea, Enrique; Thelwall, Mike; Delgado López-Cózar, Emilio	Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories	Journal of Informetrics
2019	Chapman, Karen; Ellinger, Alexander E.	An evaluation of Web of Science, Scopus and Google Scholar citations in operations management	The International Journal of Logistics Management
2021	Martín-Martín, Alberto; Thelwall, Mike; Orduna-Malea, Enrique; Delgado López-Cózar, Emilio	Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations	Scientometrics
2021	Levine-Clark, Michael; Gil, Esther L.	A new comparative citation analysis: Google Scholar, Microsoft Academic, Scopus, and Web of Science	Journal of Business & Finance Librarianship
2024	Gerasimov, Irina; Kc, Binita; Mehrabian, Armin; Acker, James; McGuire, Michael P.	Comparison of datasets citation coverage in Google Scholar, Web of Science, Scopus, Crossref, and DataCite	Scientometrics

Table 1. Short literature review of comparing citation databases

3. Methodology

The research population of this study includes the 2023 Highly Cited Researchers list reported by Clarivate™. This report has 21 categories, so we randomly selected an author from each category. Then, the citation trends of these samples were collected over ten years using the Web of Science database. Also, “Publish or Perish” software was used to collect data from Google Scholar. Correlation, Comparison of means, and trends were analyzed using SPSS and Excel software.

4. Results

Pearson’s correlation coefficient analysis indicated a significant correlation between citations received by highly cited researchers in Google Scholar and the Web of Science. The results of this test are shown in Table 2.

Table 2. Result of correlation test

Correlations			
		gs	wos
gs	Pearson Correlation	1	.689**
	Sig. (2-tailed)		.000
	N	210	210
wos	Pearson Correlation	.689**	1
	Sig. (2-tailed)	.000	
	N	210	210

**. Correlation is significant at the 0.01 level (2-tailed).

To assess the significance of the mean difference, we conducted an independent-sample t-test. The results indicated no significant difference in the mean of citations between both databases.

Table 3. Testing of the significance of the difference of the means of citation received on GS and WoS

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
y2014	Equal variances assumed	7.680	.008	2.264	40	.029	939.52381	415.06749	100.64112	1778.40649
	Equal variances not assumed			2.264	21.937	.034	939.52381	415.06749	78.58402	1800.46360
y2015	Equal variances assumed	7.176	.011	2.376	40	.022	1092.14286	459.61603	163.22420	2021.06151
	Equal variances not assumed			2.376	22.554	.026	1092.14286	459.61603	140.31276	2043.97295
y2016	Equal variances assumed	6.799	.013	2.568	40	.014	1221.09524	475.42186	260.23182	2181.95866
	Equal variances not assumed			2.568	23.863	.017	1221.09524	475.42186	239.57490	2202.61558
y2017	Equal variances assumed	5.810	.021	2.688	40	.010	1322.47619	491.97251	328.16265	2316.78973
	Equal variances not assumed			2.688	25.689	.012	1322.47619	491.97251	310.61593	2334.33645
y2018	Equal variances assumed	6.706	.013	2.899	40	.006	1434.52381	494.78212	434.53184	2434.51578
	Equal variances not assumed			2.899	27.413	.007	1434.52381	494.78212	420.02970	2449.01792
y2019	Equal variances assumed	5.895	.020	2.848	40	.007	1550.38095	544.39704	450.11350	2650.64841
	Equal variances not assumed			2.848	30.874	.008	1550.38095	544.39704	439.89171	2660.87020
y2020	Equal variances assumed	4.211	.047	2.654	40	.011	1644.23810	619.55207	392.07665	2896.39954
	Equal variances not assumed			2.654	35.489	.012	1644.23810	619.55207	387.09931	2901.37688
y2021	Equal variances assumed	2.851	.099	2.409	40	.021	1888.33333	783.97795	303.85480	3472.81187
	Equal variances not assumed			2.409	37.112	.021	1888.33333	783.97795	300.00550	3476.66116
y2022	Equal variances assumed	2.750	.105	2.434	40	.019	2052.09524	842.96453	348.40037	3755.79011
	Equal variances not assumed			2.434	37.553	.020	2052.09524	842.96453	344.93575	3759.25473
y2023	Equal variances assumed	3.558	.067	2.477	40	.018	2222.57143	897.40368	408.85093	4036.29193
	Equal variances not assumed			2.477	36.823	.018	2222.57143	897.40368	403.96395	4041.17891

Also, we have investigated the growth rate of receiving citations in both databases using exponential regression. The results indicated that the growth rate of receiving citations in both databases is nearly identical.This comparison is shown in figure1.

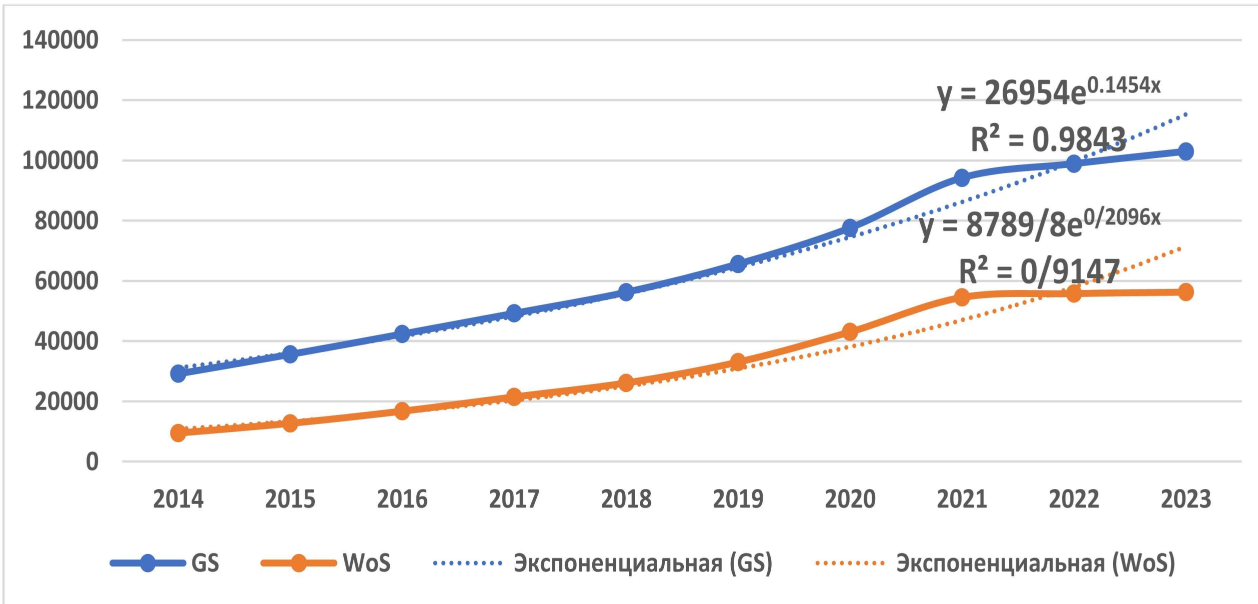

Figure 1. Growth rates of citations in GS and WoS

5. Conclusion

The findings from this study indicate that highly cited researchers tend to receive more citations in Google Scholar thanin Web of Science. This can be attributed to Google Scholar's ability to index a broader range of sources, including non-traditional and non-English publications, significantly contributing to higher citation counts. Despite the higher volume of citations in Google Scholar, the significant correlation between Google Scholar and Web of Science citations underscores the reliability of Google Scholar as a valuable tool for citation analysis.

The analysis reveals that the mean citation counts in both databases do not differ significantly, further validating the consistency between these platforms. The similar growth rates in citation counts across both databases, demonstrated through exponential regression, suggest that citation trends are comparable regardless of the database used. This finding is critical for researchers, as it indicates that while Google Scholar can capture a more diverse citations, it still aligns closely with the citation patterns observed in Web of Science.

These results highlight Google Scholar's potential as an accessible and free resource for comprehensive citation analysis. Given its broader coverage and the correlation with more established databases like the Web of Science, Google Scholar can serve as a complementary tool, especially for researchers in diverse and multidisciplinary fields. However, researchers should remain aware of the differences in indexing practices and the potential for varying citation quality between databases.

In conclusion, the study reinforces the importance of utilizing multiple citation databases for a more holistic understanding of scholarly impact and research trends. Google Scholar, with its extensive scope, coupled with the reliability of the Web of Science, provides a robust framework for citation analysis. Future research should continue to explore the dynamics of these databases to enhance further the accuracy and comprehensiveness of citation analyses in the academic community.

References

[1] Adriaanse, L. S., & Rensleigh, C. (2011). Comparing Web of Science, Scopus and Google Scholar from an environmental sciences perspective. *South African Journal of Libraries and Information Science*, *77*(2), 169–178.

[2] Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries*, *3*(1), 7. https://doi.org/10.1186/1742-5581-3-7

[3] Bar-Ilan, J. (2008). Which h-index? — A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, *74*(2), 257–271. https://doi.org/10.1007/s11192-008-0216-y

[4] Chapman, K., & Ellinger, A. E. (2019). An evaluation of Web of Science, Scopus and Google Scholar citations in operations management. *The International Journal of Logistics Management*, *30*(4), 1039–1053. https://doi.org/10.1108/IJLM-04-2019-0110

[5] De Winter, J. C. F., Zadpoor, A. A., & Dodou, D. (2014). The expansion of Google Scholar versus Web of Science: A longitudinal study. *Scientometrics*, *98*(2), 1547–1565. https://doi.org/10.1007/s11192-013-1089-2

[6] Delgado-López-Cózar, E., & Repiso-Caballero, R. (2013). The impact of scientific journals of communication: Comparing Google Scholar Metrics, Web of Science and Scopus. *Comunicar*, *21*(41), 45–52. https://doi.org/10.3916/C41-2013-04

[7] Gerasimov, I., Kc, B., Mehrabian, A., Acker, J., & McGuire, M. P. (2024). Comparison of datasets citation coverage in Google Scholar, Web of Science, Scopus, Crossref, and DataCite. *Scientometrics*, *129*(7), 3681–3704. https://doi.org/10.1007/s11192-024-05073-5

[8] Harzing, A.-W. (2013). A preliminary test of Google Scholar as a source for citation data: A longitudinal study of Nobel prize winners. *Scientometrics*, *94*(3), 1057–1075. https://doi.org/10.1007/s11192-012-0777-7

[9] Harzing, A.-W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804. https://doi.org/10.1007/s11192-015-1798-9

[10] Kousha, K., & Thelwall, M. (2008). Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. *Scientometrics*, *74*(2), 273–294. https://doi.org/10.1007/s11192-008-0217-x

[11] Levine-Clark, M., & Gil, E. L. (2021). A new comparative citation analysis: Google Scholar, Microsoft Academic, Scopus, and Web of Science. *Journal of Business & Finance Librarianship*, *26*(1–2), 145–163. https://doi.org/10.1080/08963568.2021.1916724

[12] Li, J., Burnham, J. F., Lemley, T., & Britton, R. M. (2010). Citation Analysis: Comparison of Web of Science®, Scopus™, SciFinder®, and Google Scholar. *Journal of Electronic Resources in Medical Libraries*, *7*(3), 196–217. https://doi.org/10.1080/15424065.2010.505518

[13] Martín-Martín, A., Orduna-Malea, E., & Delgado López-Cózar, E. (2018). Coverage of highly-cited documents in Google Scholar, Web of Science, and Scopus: A multidisciplinary comparison. *Scientometrics*, *116*(3), 2175–2188. https://doi.org/10.1007/s11192-018-2820-9

[14] Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160–1177. https://doi.org/10.1016/j.joi.2018.09.002

[15] Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4

[16] Meho, L. I., & Yang, K. (2006). *A New Era in Citation and Bibliometric Analyses: Web of Science, Scopus, and Google Scholar* (arXiv:cs/0612132). arXiv. http://arxiv.org/abs/cs/0612132

[17] Mingers, J., & Lipitakis, E. A. E. C. G. (2010). Counting the citations: A comparison of Web of Science and Google Scholar in the field of business and management. *Scientometrics*, *85*(2), 613–625. https://doi.org/10.1007/s11192-010-0270-0

[18] Moed, H. F., Bar-Ilan, J., & Halevi, G. (2016). A new methodology for comparing Google Scholar and Scopus. *Journal of Informetrics*, *10*(2), 533–551. https://doi.org/10.1016/j.joi.2016.04.017

[19] Pauly, D., & Stergiou, K. I. (2005). Equivalence of results from two citation analyses: Thomson ISI's Citation Index and Google's Scholar service. *Ethics in Science and Environmental Politics*, *2005*(1), 33-35. https://www.int-res.com/articles/esep/2005/E65.pdf

[20] S. Adriaanse, L., & Rensleigh, C. (2013). Web of Science, Scopus and Google Scholar: A content comprehensiveness comparison. *The Electronic Library*, *31*(6), 727–744. https://doi.org/10.1108/EL-12-2011-0174

[21] Yang, K., & Meho, L. I. (2006). Citation Analysis: A Comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for Information Science and Technology*, *43*(1), 1–15. https://doi.org/10.1002/meet.14504301185