



**HAL**  
open science

# Measuring Network-Driven Citations: An Adjusted Citation Count Metric

Emilie Bonhoure, Ransome Epie Bawack

► **To cite this version:**

Emilie Bonhoure, Ransome Epie Bawack. Measuring Network-Driven Citations: An Adjusted Citation Count Metric. 2023. <hal-03961131>

**HAL Id: hal-03961131**

**<https://hal.science/hal-03961131v1>**

Preprint submitted on 28 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Measuring Network-Driven Citations: An Adjusted Citation Count Metric

Emilie Bonhoure<sup>α,\*</sup> and Ransome Epie Bawack<sup>β</sup>

## Abstract

Citation count is extensively used within research systems to measure research performance and relevance. However, high numbers of citations could result from many different factors. This paper proposes an adjusted citation count metric that could help identify citations coming from researchers' networks. Using data from the Web of Science core collection, this research suggests and illustrates an approach to compute the proportion of citation counts from within and out of author networks, defined as the set of their co-authors. The results reveal the trends and effects of author networks on citation counts. The paper explains how these findings could be used to assess research performance. Its utilisation will depend mainly on the objectives of the assessor and the stage of the researcher's career. An algorithm is also proposed that could be used to automate the computation of the proposed metric. It could serve as a step towards refining and integrating this metric into existing bibliometric analysis packages built on the R software platform.

**Keywords:** citation, research impact, researcher network, bibliometric, research quality, performance

---

<sup>α</sup> Kedge Business School, 680 cours de la Libération, 33405 Talence, France. Orcid number: <https://orcid.org/0000-0002-6065-8395>.

\* Corresponding author: [emilie.bonhoure@kedgebs.com](mailto:emilie.bonhoure@kedgebs.com).

<sup>β</sup> Audencia Business School, 8 route de la Jonelière, 44000 Nantes, France. Orcid number: <https://orcid.org/0000-0002-5441-604X>.

## Introduction

Citation count is extensively used within research systems today to measure research impact, quality, and relevance (Aksnes et al., 2019; Heßler & Ziegler, 2022). For example, it has been used to assess researchers, research groups, departments, and institutions (Abbasi & Jamali, 2019; Berche et al., 2016; Mohammed et al., 2020). However, the accuracy of citation metrics for such assessments has been a longstanding debate in the field of information sciences (Caon et al., 2020; Downes & Lancaster, 2020; J. P. A. Ioannidis, 2015). The reason for this conundrum is the inability of citation metrics to fully measure research plausibility, originality, scientific value, and societal value, which are commonly perceived as characteristics of quality research (Aksnes et al., 2019). Thus, several researchers have investigated the contexts and purposes for which citation would suit research assessments (Halevi et al., 2017; Thelwall, 2018).

Today, citation plays a crucial role in assessing researchers for hiring, promotion, grants, and tenure since citation count is easily determined, making it a desirable metric to assess researchers (Moher et al., 2018; Zou & Peterson, 2016). Yet, this simple metric does not allow assessors to disentangle the individual effect of its components, making it difficult to assess what citation count truly measures. Indeed, this measure can be shaped by many factors related to one's research (Amjad et al., 2022; Hirv, 2022; Sjögarde & Didegah, 2022), ranging from research impact and visibility to the extent of the researcher's network (Aksnes et al., 2019; Tian et al., 2021; Waltman & Costas, 2014). For instance, it is difficult to straightforwardly perceive the number of citations resulting from the authors' network of co-authors or social network<sup>1</sup>. This phenomenon has captured the attention of the information science community and some researchers have revealed several strategic approaches used by researchers (J. P. A. Ioannidis, 2015; Tajedini et al., 2018). They also provide evidence that the professional social network of researchers can be used to predict their citation-based research performance, in particular their co-authors (Abbasi et al., 2011; Costas et al., 2010; J. P. A. Ioannidis, 2015). Yet, no approach or tool is available to assist assessors (employers, institutions) assess the impact of co-authorship networks on the performance of scholars.

This research aims to propose an approach towards calculating citation count that discriminates citations resulting from the professional social networks of researchers. We call this new metric *adjusted citation metric*, as opposed to the raw citation metric that does include them. We argue that estimating raw and adjusted citation counts along with the difference between both would help future assessors disentangle the effect of overall research impact from the impact of co-authorship networks. Citation count that does not include those from the researcher's network (*adjusted citation metric*) could be more substantial evidence of research scientific and societal impact and visibility. Meanwhile, it could also be strong evidence of the extent of one researcher's professional network, collaboration, and impact in a specific research discipline or community. The value of this unit of measurement depends on the goals of the assessor. Thus, this research contributes to the body of knowledge that seeks to improve the reliability of citation count in assessing research (Aksnes et al., 2019; Tajedini et al., 2018). It also provides a tool that can help assessors to assess researchers and research impact objectively and strategically. The following section summarises the extant literature on using citation metrics to assess research quality.

---

<sup>1</sup> A researcher's social network is considered all the potential colleagues he could have worked with (co-authors, doctoral students, members from the same faculty, etc). In the rest of the article, we restrict one researcher's network to "the set of all his co-authors" (see Appendix 1).

## Literature review

For over three decades, citation metrics have been used to evaluate the quality of researchers, research teams, and academic institutions (Aksnes et al., 2019; J. Ioannidis et al., 2016; Molléri et al., 2018). Several researchers advocate that citation count is a simple, transparent, precise, objective, useful, and affordable metric for research quality and performance assessment (Bornmann & Marx, 2014; Charlton & Andras, 2007). Given the availability of journal and publication databases, they are easy to calculate (Bornmann & Marx, 2014) and can be used to assess individual papers or authors (Waltman & Costas, 2014).

However, concerns have been raised in the academic literature about their use (Nightingale & Marshall, 2012). Some researchers argue that this metric can only be regarded as a reasonable proxy for true quality or impact (Wooldridge & King, 2019). It may be determined by other factors than the impact of the evaluated research (Aksnes et al., 2019; Amjad et al., 2022; Molléri et al., 2018; Waltman & Costas, 2014). In particular, scholars may be biased when adding citations to their work: they may not cite all the papers that influenced their work and tend to favour “confirmatory” studies (Bornmann, 2017). They can also make extensive use of self-citations (Kacem et al., 2020; Nightingale & Marshall, 2012; Szomszor et al., 2020).

Moreover, when using citation metrics as a measure of research performance, the type of citations and the number of citations in one paper are rarely considered (Rüdiger et al., 2021). Whether a paper is cited as a simple example of the findings of the citing paper does not imply the same research impact as if it is cited as a seminal paper (Liu et al., 2014; Wan & Liu, 2014a; Zhao & Strotmann, 2016). The citation metric does not distinguish either whether the paper is cited once or several times in one single citing paper (Wan & Liu, 2014b). Another drawback of citation metrics is related to their poor comparability. In particular, they do not account for differences (in terms of citation and publication patterns) across fields (Bornmann & Leydesdorff, 2013; Bornmann & Marx, 2014; J. Ioannidis et al., 2016). This concern has led several authors to propose field-normalised citation counts (Bornmann & Daniel, 2009; Bornmann & Leydesdorff, 2013; Waltman & van Eck, 2013). Similarly, these metrics do not consider the age of the cited publication (J. Ioannidis et al., 2016).

A third type of concern is being raised over the potential deviant behaviours that the above points may lead to (Aksnes et al., 2019; J. P. A. Ioannidis, 2015). Citation analysis is “vulnerable to gaming” (Waltman & Costas, 2014) as researchers can modulate their citation behaviour to increase their metrics. In addition to self-citations, they could, for instance, agree with researchers from their network (like former co-authors) to cite each other’s paper as part of a win-win deal to increase their citation metrics. Researchers may also publish more low-impact papers instead of fewer high-quality ones to gather as many citations as possible (Molléri et al., 2018), thus replacing researchers’ “taste for science” by a greater “taste for publication” (Osterloh & Frey, 2015). One illustration of that is the positive correlation between the citation counts of a paper and the number of its co-authors. While it is widely shown that the former increases with the latter (Larivière et al., 2015), the reasons are unclear (Bornmann, 2017; Thelwall & Maflahi, 2020).

This network-related citation advantage could illustrate researchers’ tendency to cite their co-authors as part of a win-win deal to increase both their citation metrics. However, it could also illustrate the fact that citation metrics measure several components related to researchers’ performance, and not only their impact (Waltman & Costas, 2014). If citation metrics indeed primarily measure a researcher’s impact (quality, societal and informative value of his research) (Aksnes et al., 2019; J. Ioannidis et al., 2016; Molléri et al., 2018), they also include a component related to their professional network (Demaine, 2022; Tian et al., 2021; Zhang et al., 2021). From

employers' and institutions' perspective, citations from co-authors and members of a professional network is not necessarily a negative aspect because they reveal the researchers' collaboration network and sphere of expertise in a research discipline, starting from their co-authors (Milard, 2014). Indeed, it can be a way to assess the extent of one researcher's potential collaboration and intellectual exchanges (Milard, 2014, p. 2459). In this light, a measure of network effect on citation count would be a useful tool for research assessors. Such concerns have motivated researchers to propose several improvements to the citation count metric. For example, some propose normalising citation analysis used for evaluation purposes based on a control group of similar articles and suggest control group requirements for credible normalisation, especially when assessing the quality of research outputs across different disciplines (Kostoff, 2002). Considering one specific dimension of researchers' network (co-authorship), others propose to distinguish author self-citations from co-author self-citations (Costas et al., 2010). Others also suggest that total citation counts should be used to benchmark papers published in the same calendar year on a specific research theme or field: this metric would enable more informed performance evaluations of junior researchers (Dechow et al., 2019; Khan et al., 2014). Many other citation indicators have also been discussed (see Froghi et al., 2012; and Wildgaard et al., 2014 for a complete overview) such as the i10 index, the h-index (Froghi et al., 2012; Hirsch, 2005), the g-index (Bornmann et al., 2008; Egghe, 2006; Thompson et al., 2009), and the m-quotient (Bornmann et al., 2008; Thompson et al., 2009). Nevertheless, none of the existing measurement tools or approaches addresses the specific issue of separating network-based citations from total citations, which is the gap this research paper attempts to fill by proposing a citation count adjusted for network effect.

## **Methodology**

This paper adopts a three-step approach to propose an adjusted citation count metric that attempts to account for researchers' network effect. First, we discuss the proposed adjusted metric. Second, we explain how this metric can be computed. Third, we illustrate how this metric can be applied.

### **Explaining the adjusted citation metric**

We are proposing an adjusted citation count that would help measure researchers' research impact and network. Our proposed metric considers a target author's<sup>2</sup> citation count adjusted with a potential network effect. The target author's network is defined as the set of all his/her co-authors. As mentioned in the prior section, one part of the citations of an author could be from researchers who do not belong to the author's network. Another part of the citations of an author could be related to their professional network (set of co-authors). Such citations would reflect a more significant network effect. The adjusted citation metric eliminates the potential network effect by considering only the citations made by non-related authors, and thus only reveals the part of citations coming from authors out of the researchers' network. Conversely, should future assessors only consider the part of citations revealing the researcher's network, they should consider the difference between raw and adjusted citation metrics. Conceptually, the adjusted citation metric is the difference between the target author's citation count and the citations made in publications signed by the target author or one of their co-authors. For a given target author, it considers the citations made only by people outside their network.

---

<sup>2</sup> Appendix 1 provides more detailed definitions of the vocabulary used in this study.

This adjusted metric can be compared to the raw citation count through the adjusted citation ratio. It is the ratio of the adjusted citation count to the raw citation count. If this ratio is large (and particularly close to 1), it means that most citations of the target author come from researchers who are not part of their network. It implies that the target author's publications are likely cited due to the overall scientific impact rather than network effect. The reverse is at play when the ratio is closer to zero: it indicates that most of the author's citations result from a large network effect (co-authors cite the target author), meaning the target author belongs to a strong research community.

### **Computing the adjusted citation count metric**

The first step in computing the adjusted citation metric for any target author is obtaining all the author's publications and their citation count. The publications can be obtained from databases like Google Scholar, Web of Science, and Scopus. The second step involves identifying all members of their network (co-authors in the context of this research) and their respective publications. Third, identify and subtract all citations from the target author's citation count that originates from a co-authors publication of which they are not apart. The result of this subtraction is the adjusted citation count. Fourth, calculate the adjusted citation count ratio by dividing the adjusted citation count by the raw citation count (that perceived directly on the database, e.g., Google Scholar). An algorithm is proposed in the R programming language<sup>3</sup> to facilitate this computation. Implementing the proposed script requires loading the bibliometrix R package (Aria & Cuccurullo, 2017). The package provides a set of tools for bibliometric analysis. The script presents how to calculate the adjusted citation count and citation count ratio based on bibliometric data collected from the Web of Science (WoS) core collection.

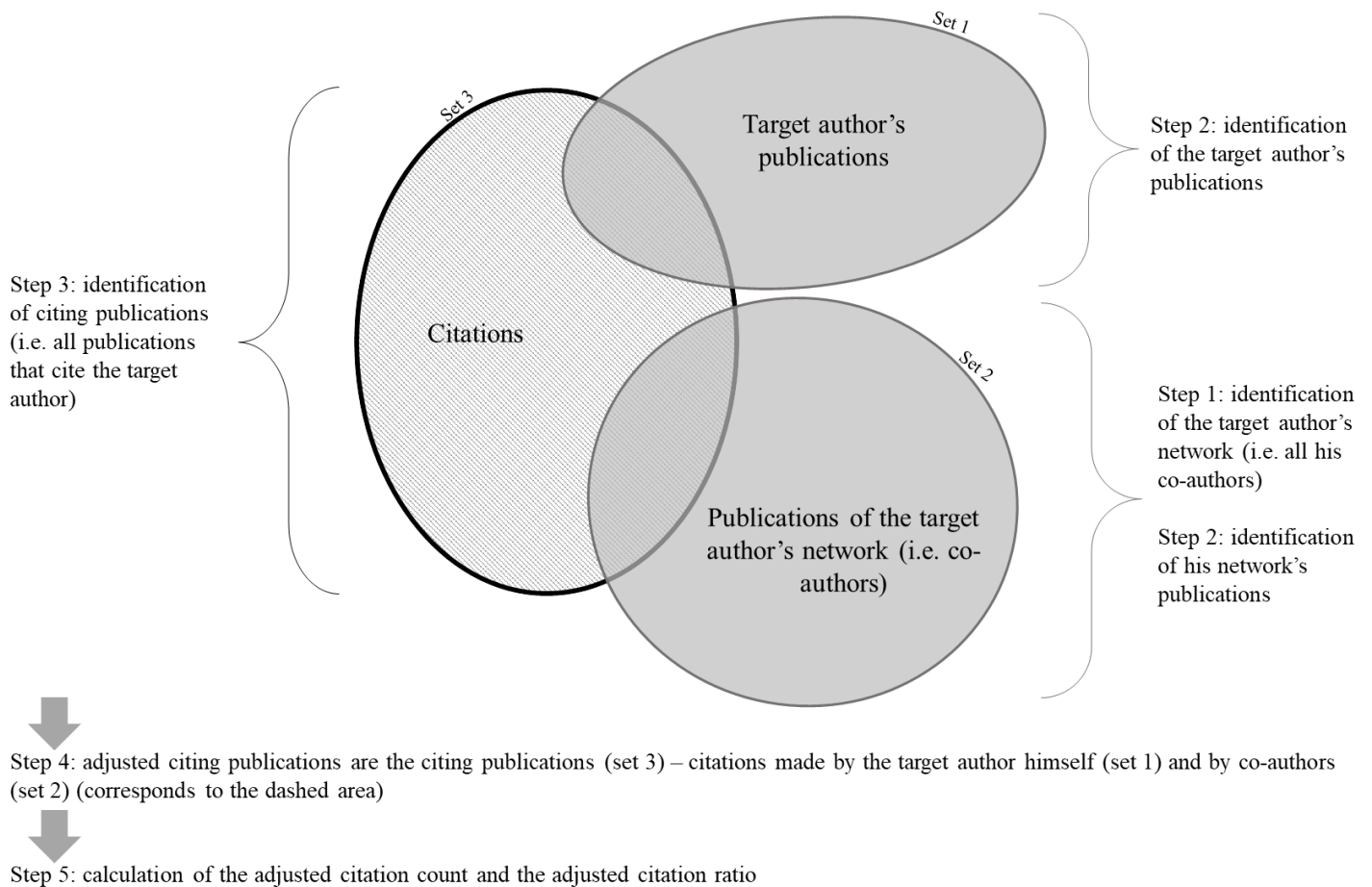
Using this script requires the user seeking to compute the proposed metrics as follows<sup>4</sup>. First, the investigator has to identify all co-authors of the target author. The target author's publications records should be downloaded from WoS and uploaded into R. Our algorithm will automatically extract the list of co-authors. Second, extract the publication records of all co-authors identified in step 1 and upload it into R. Third, extract the list of publications that cite the target author from WoS and load it into R. Fourth, the algorithm compares the lists loaded in step 2 and step 3. It produces an adjusted set of citing publications: citing publications (extracted and loaded in step 3) minus the set of co-authors' publications (extracted and loaded in step 2). In other words, the adjusted set of citing papers consists of the papers that cite the target author, which the target author or their co-authors did not write. Fifth, the algorithm computes two new measures. It first computes the adjusted citation count, which is the number of papers in the set elaborated in step 4. It then computes the adjusted citation count ratio, which is the ratio of the adjusted citation count to the raw citation count.

Fig. 1 summarises the steps described above. It should be noted that, for the purpose of this proof of concept, we did not restrict the analysis to any publication or citation window. Instead, we considered one target author's publications and co-authors throughout their professional career.

---

<sup>3</sup> The script can be requested from the authors.

<sup>4</sup> The steps here correspond to the ones mentioned as comments in the R source code.



**Fig. 1** Algorithm to compute the adjusted citation count metric. Note: the intersection between set 1 and set 3 represents the self-citations; the intersection between set 2 and set 3 represents the publications of the target author' network that cite at least one publication co-signed by the target author.

### Application: calculation of the adjusted citation metric for five Google Scholar themes

As an illustration, we calculate the adjusted citation count metric and the adjusted citation count ratio for all the authors mentioned in five Google Scholar (GS) themes: “data analytics and visualisation”, “entrepreneurial economics”, “environmental marketing”, “nanobioengineering”, and “wine economics”. These themes were selected randomly and significantly distanced from the research fields of the authors to avoid interpretation biases or potential conflicts of interest. Table 1 describes the research themes investigated and provides some statistics on target authors' citations and network taken from GS.

**Table 1** Research themes from Google Scholar analysed

Field	Nb of authors	Max nb of citations	Min nb of citations	Average nb of co-authors
data analytics and visualisation	14	9 081	3	23
entrepreneurial economics	8	6 978	27	12
environmental marketing	16	13 866	1	11
nanobioengineering	13	73 547	15	26
wine economics	28	25 246	4	13

Note: The reported figures have been collected from GS on May 14th, 2021. The "Nb of authors" column specifies the number of target authors for each theme; the "Max (min respectively) nb of citations" reports the number of citations of the target author who had the largest (lowest resp.) number of citations. The "Average nb of co-authors" reports the average number of different co-authors per target author in each theme. These figures do not include target authors without citations.

---

Three target authors were excluded during the adjusted citation metric computation due to our inability to identify them uniquely in WoS.

After computing the adjusted citation count metric, we plot the adjusted citation count ratio as a function of the raw citation count to identify the citation behaviour patterns of target authors. However, given several limitations of the WoS platform, some manual adjustments are required to identify the target author's network (step 1 of the algorithm). An important caveat of WoS is that it does not identify authors uniquely, resulting in (i) many homonyms that are different authors and (ii) authors that should be identified using several "names". Another flaw is that one can only download publication records in batches of 500 records, making more difficult the work on target authors with many co-authors and thus many cumulative publications. The limitations of WoS are presented with more details in the "Limitations and future research agenda" section.

To deal with these caveats, we developed an alternative and manual way to proceed with step 1 of the algorithm. Co-authors were identified visually using GS. Moreover, we restricted the set of co-authors to the co-authors of the ten most cited co-authored papers of the target author on GS (hereafter the "main" co-authors). If for one target author, less than 10 papers are cited, we include the co-authors of the non-cited papers, in the order presented by GS (namely from the most recent to the oldest), until we reach the 10 co-authored paper thresholds. We then search for the list of publications of the target author and the main co-authors on WoS. We end up with a list of the publications of the target author and main co-authors. This list is the data we used to proceed with the rest of the algorithm and the automated computation of the adjusted citation count metric.

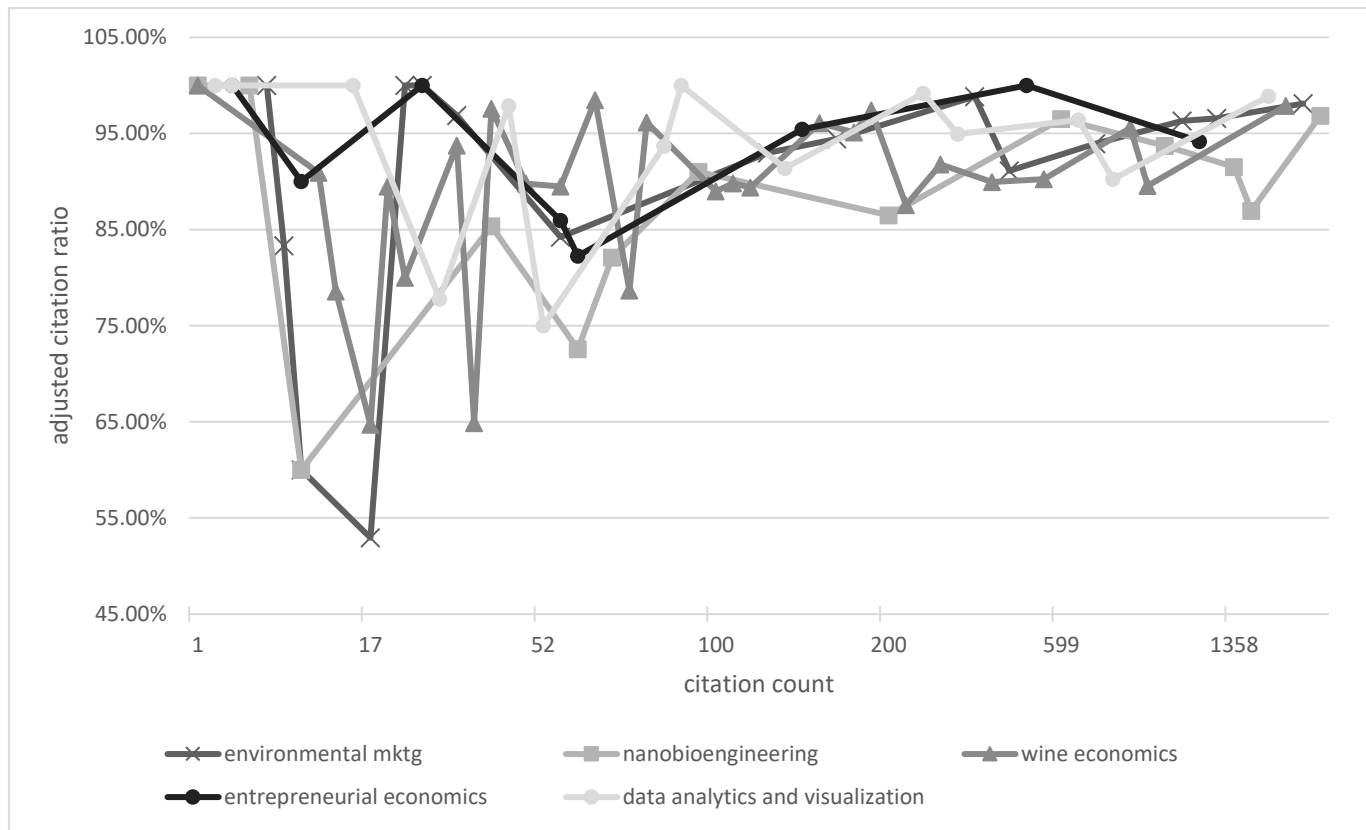
## **Results and discussion**

### **Results of the application**

As a proof of concept, this study recalculates the citation count of several authors after adjusting it for network effects. Fig. 2 presents the results for the authors of the five GS themes investigated. For more visibility, Appendix 2 (Fig. 3 to Fig. 7) shows the results on five different graphs<sup>5</sup>.

---

<sup>5</sup> For Fig.2 to Fig.7, it should be noted that each point on the graphs represents one single target author, since there are no target authors in the same field who have exactly the same raw citation count.



**Fig. 2** Adjusted citation ratio and (raw) citation count

The results for the five research themes investigated reveal the same general trend. This trend depends on the total number of citations for each author. For those who have more citations (minimum 50 to 100, depending on the themes), the adjusted citation ratio increases with the number of citations. It also ranges among high levels: from around 80-85% to 98%-99%. In other words, most citations of highly cited authors come from researchers outside their research network.

Among authors that are less cited (below the 50 to the 100-citation threshold, depending on the themes), findings are more heterogeneous. There are two phenomena at play. First, some of these researchers have an adjusted citation ratio equal to 100%. This score provides evidence of the total absence of a network effect. Either such researchers do not use their network to make their article circulate or they do not have a large enough network to do so. In both cases, their citations only illustrate their overall research impact, as none of their co-authors cites them. Conversely to the first phenomenon, the second highlights authors who have much lower adjusted citation count ratios. Some of these ratios fall below the 75%-level, indicating that more than a quarter of the target author's citations are based on network effects.

The general trend is more apparent when one focuses on the "environmental marketing" research theme. Below the 50-citation threshold, authors either have very high adjusted citation count ratios (three of them have a ratio equal to 100%) or much lower ones (one has a ratio equal to 53%, another one to 60%). From this threshold, the adjusted citation ratio slowly increases to reach 96% to 98% levels, without ever reaching 100%. Results on the "nanobioengineering" field also illustrate the general trend. Researchers who have fewer citations (below 100 citations) present two types of adjusted citation ratios. Either their ratios are very high, particularly equal to 100%

for the two of them who have one and five citations, or they have lower ratios (see the lower ratio for that theme that is 60%). Above the 100-citation threshold, the adjusted ratio starts increasing and reaches levels up to 97%.

## **Discussion**

This study presents an “adjusted” metric of citation counts that proposes to solve one of its major flaws: the difficulty in disentangling network effects from overall research impact within citation measures. An algorithm is developed to compute this new metric automatically. Using information extracted from WoS allows the user to determine all co-authors of one target author and calculate their adjusted citation count and ratio. Five research themes are used to prove the relevance of the proposed concept (metric) and to put forth an algorithm that can be used to automate its computation. The results show that the adjusted citation ratio is correlated but non-linear to the raw number of citations.

Despite the apparent exogeneity of citation count, these results provide evidence of the importance of disentangling network from overall research-impact effects when considering citation metrics. Citations are well-known measures of performance in academia (Aksnes et al., 2019; J. Ioannidis et al., 2016; Molléri et al., 2018). Generally, a high citation count is considered a signal of research impact. However, it can also indicate the extent of authors’ networks (particularly their co-authors). Thus, it can indicate how large the potential for one author’s research collaborations is to institutions such as one future employer. In such cases, citation count metrics are also a measure of researchers’ network and sphere of influence.

These results highlight different adjusted-citation patterns. Contrary to what could be expected, there are two distinct patterns for authors with fewer citations (below a 50-to-100 threshold). With the first (more surprising) one, citations are entirely due to one’s research impact (authors with an adjusted citation count ratio equal to 100%). Thus, from institutions’ and employers’ perspective, citation metrics should be used only to measure these authors’ research impact. In turn, it would be impossible to conclude on their collaboration potential since they could be either reluctant or unable to mobilize their network to make their article circulate. The second pattern for less cited authors is the opposite – a significant part of the citations of these authors’ publications come from their network. Some of them have adjusted citation count ratios equal to 50% or 60%, meaning that about half of their citations are made in publications from colleagues, former colleagues, overall co-authors. Therefore, citations of such researchers are due, for one part, to their research impact, and, for the other part, to their large network and their ability to mobilize it. In all cases (first or second pattern), institutions’ and employers’ strategy could depend on what they are looking for. If they want to recruit researchers with a high research impact, they could use the adjusted citation metric and ratio. If they want to recruit researchers with a high potential of collaborations and intellectual exchanges, they could compute the difference between the raw and our adjusted citation metrics.

The picture is different for the researchers who are more cited. For them, the relationship between citation counts and the adjusted citation ratio is more linear. The more an author is cited, the higher the adjusted citation count ratio. Most of their citations (around or above 90%) come from outside their network. This fact is in line with the citation metric as a research-impact measure. Nevertheless, a small portion of their citations keeps coming from their network, reinforcing the need to distinguish between both. Again, institutions’ and employers’ strategy could depend on their own preferences. If they want to recruit researchers with a high research impact, they should favour those with a high adjusted citation ratio. If they want to recruit researchers with a high (past or future)

number of collaborations, they should favour those with a low adjusted citation ratio (and thus a high raw-adjusted citation metric difference).

Another significant result relates to the heterogeneity across fields. The threshold above which the adjusted citation ratio starts increasing depends on the research area. In the “entrepreneurial economics” research theme, the adjusted citation count ratios do not go below 80%, even for less cited authors. Conversely, those ratios are much more heterogeneous in the “wine economics” field, even for more cited authors (with ratios ranging between 87% and 98%). One explanation for that could be related to the nature of each field: the general topic they belong to (whether they are part of biology or physics, management sciences, or human sciences) and how transdisciplinary they are. For instance, “nanobioengineering” is less transdisciplinary than “wine economics”, which contains studies on marketing, finance, economics, sustainable development, and agricultural issues. It may explain why this theme presents more heterogeneous results<sup>6</sup>.

### **Contributions**

Overall, the proof of concept validates the idea of the adjusted citation metric. By documenting different patterns depending on the fields and authors’ citation counts, this study supports the importance of disentangling authors’ network from research impact in assessing performance, especially measured by the number of citations. This evidence is the first contribution of this paper. This paper also contributes to the research stream that advocates citation count as a measure of research quality that is precise and useful for future policy (Charlton & Andras, 2007). It helps researchers understand citation behaviour (Bornmann, 2017; Bosquet & Combes, 2013; Didegah & Thelwall, 2013; Larivière et al., 2015; Urlings et al., 2021) and suggests that additional factors should be considered when exploring citation measures. Thus, it responds to calls for bibliometric indicators to be scrutinised and improved regularly to mitigate the use of incorrect or misleading scientometric instruments (J. P. A. Ioannidis, 2015; Kenna et al., 2017). Specifically, it proposes an adjusted citation metric that could be used to enhance the precision and usefulness of citation counts for defining policies regarding the assessment of researchers and research teams in institutions. It proposes a new approach to assessing both research impact and researchers’ network through citation analysis.

This new approach complements extant approaches (Vrat, 2020) by proposing a simple metric to assess research impact based on citation count. The approach involves adjusting citation counts by eliminating citations from the researchers’ co-author network, thus reducing potential bias in the metric. Symmetrically, this metric could be used to assess researchers’ (co-author) network, and thus the extent of their collaborations and intellectual exchanges among the research community. In this case, either the adjusted citation metric or the difference between raw and adjusted citation metrics could be considered. We thus offer a simple metric to measure researchers’ and institutions’ network and potential collaborations (Wiechetek & Pastuszak, 2022) The proposed adjusted citation count can then be normalised with a control group using the guidelines for credible citation analysis proposed by Kostoff (2002) to determine researcher or research team impact or network. The proposed adjusted citation count can also be used to develop benchmarks for comparison purposes in different research themes or domains (Khan et al., 2014).

---

<sup>6</sup> Researchers tagged as working on “wine economics” are likely to publish on other issues, thus with a wide range of co-authors. It may explain very different strategies across researchers.

A more methodological contribution is that this study proposes an algorithm that automatically computes the adjusted citation metric (adjusted citation count and ratio). Although it still needs improvement, the algorithm is a step towards efficiently calculating the proposed citation metrics. It could be incorporated into established bibliometric R packages to make it easily accessible to researchers or institutions performing bibliometric analysis. It also enhances the ability of institutions or research groups to disentangle research-impact from network effects quickly.

From a practitioner perspective, the algorithm is proposed freely to encourage researchers to investigate further and refine this research impact and network metric. It would help investigators save time on the long and tedious process of computing the metrics manually. The algorithm can be applied to any research theme or field, as demonstrated in this study using five different research themes. Also, other researchers can improve the algorithm and implement it in bibliometric analysis software packages. The findings made using this algorithm go far beyond just calculating indices – they have substantial implications for decision-makers who assess the performance of researchers, research impact and researchers' network. For example, instead of measuring researchers' performance through raw citation counts, institutions may use the adjusted metric to disentangle impact and network effects. Such a metric could also help them refine their strategies in terms of recruitment.

Those willing to recruit based only on their research impact and recognition in terms of research should look for the highest possible adjusted citation ratio. Those looking for researchers with a strong network and a large base of co-authors may prefer lower adjusted citation ratios when recruiting. Alternatively, institutions could use the adjusted citation count to assess the performance of their research department. The adjusted citation metric may also be used at the individual level. Whether they prefer to signal their impact or their network, researchers may display differently this new metric. They could also use it to assess the performance of other scholars. For instance, when conducting a literature review or a bibliometric study, one could favour publications or authors with the highest adjusted citation ratios to ensure that the results with more impact are considered, excluding for any potential network effect. Conversely, when conducting a network analysis (on the model of Maltseva & Batagelj, 2022's) based on citing publications, they could favour publications or authors with the lowest adjusted citation ratios.

Lastly, the adjusted metric could be used to estimate the true impact of a global theory outside its initial field. For instance, the agency theory (Jensen & Meckling, 1976) has been developed within the financial economics research area. Later on, it has been extended to management sciences and is now extensively used in many other social science disciplines. One way to assess the impact of this theory (and possibly to compare it with other theories) would be to look at an adjusted citation count of its seminal paper. This adjusted citation count would only consider the citations (of the seminal paper) made by publications in fields other than the originating field, which is financial economics. Such a process could also work for highly-cited and high-impact authors like the Nobel prize winners.

### **Limitations and future research agenda**

One limitation of this analysis is related to the manual treatment made to deal with the main WoS caveats. One of them is that the platform does not identify authors using a unique ID. This platform limitation results in two

challenges: first, authors can be identified differently in WoS depending on how their names are spelt<sup>7</sup>. Furthermore, the platform does not correct for homonyms, thus resulting in several people that could be identified using the same names<sup>8</sup>. This first caveat has one major consequence on the adjusted citation metric computation. It may seem logical to include all the possible “names” (with or without the first name, with initials, etc.) of the target author and corresponding co-authors to make sure one extracts all their publications. However, this would also include many irrelevant homonyms, thus, many irrelevant publication records in the dataset.

Another WoS caveat is related to the maximum number of publication records downloaded at once (500). In the case of target authors who have many co-authors who have cumulatively written thousands of papers, downloading the complete list of their publications would be too long. This challenge is more apparent when one includes all the name variants for each author.

The restriction of the co-authors to a list of “main” co-authors leads to potential biases in the adjusted citation metric. Since only main co-authors are considered, the set of all the co-authors of a target author may be larger. We do not look at the citations made in the publications of non-main co-authors. This decision leads to an upward bias in the calculated adjusted citation metric, representing a maximum threshold for each target author. Furthermore, this upward bias in the adjusted citation metric may be larger for authors that publish more (and have more co-authors).

Given that this illustration is a proof of concept, future research should consider all the co-authors of a target author to obtain more accurate adjusted citation counts. These issues would also be easily handled if the source database (WoS in this paper) identified authors uniquely and allowed massive downloads of bibliometric data. Using our algorithm, it would be easy to automatically extract the target author’s name, the complete list of his co-authors and their publications, along with the list of citations from WoS or any other bibliometric platform.

Another limitation is related to our definition of a target author’s network. This study only considers the target author’s co-authors. So far, the proposed adjusted citation metric is thus very similar to “co-author self-citations” sometimes mentioned in extant literature (Costas et al., 2010; J. P. A. Ioannidis, 2015), while its aim is to account for all the target author’s network. Many other researchers than co-authors could be part of this network: colleagues, former classmates, PhD students, (former) supervisors, etc. Even if these people did not co-sign any publication with the target author, some citations coming from them might be more network- than impact-related. Thus, further research should consider expanding the scope of the target author’s network and propose means of capturing such network effects in the proposed citation metrics.

Additionally, the adjusted citation metric considers one target author as part of his own network. It includes self-citations into network citations, even though they do not actually represent the network and collaboration potential of one target author (Kacem et al., 2020; Szomszor et al., 2020). Therefore, further research should distinguish between self-citations from network citations to refine the assessment of one researcher’s network potential (Costas et al., 2010). The new metric does not distinguish either “local” co-authors (namely from the same department or university as the target author) from “foreign” ones (namely from other department or university). Since international collaboration is usually valued, one could refine the network scope by limiting it to “foreign” co-authors. This way, institutions and potential employers would be able to look for researchers with the highest international network effect.

---

<sup>7</sup> For instance, the author John Smith can be denominated, in WoS, under three different “names”: “Smith John”, “Smith”, or “Smith J.”.

<sup>8</sup> For instance, “Smith J.” could identify John Smith or Janet Smith.

A final limitation is that this article only proposes a proof of concept for our new metric. It formally introduces the adjusted citation metrics and proposes an approach to computing them. However, significant research is needed to develop this metric fully and assess its determinants. It would be necessary to empirically investigate how this metric is affected by the field of study, number of co-authors and research experience of the target author.

For instance, we do not restrict our analysis to any publication and citation windows, while they could be crucial (Chen et al., 2022). Thus, we do not distinguish between scholars depending on their level, while whether they are young or old scholars would surely affect their citation count. Normalizing by career level will thus be a necessary step to refine the adjusted citation metric. Generally, career, field, age, or other normalizations could be useful to refine the analysis of the research impact and network of one researcher, one department or one university, and compare one with another. Future assessors could use one of these normalized metrics depending on their specific needs.

## **Conclusion**

This paper explores the relationship between citations, research impact, and researcher network to propose an adjusted citation metric that would help distinguish citation counts resulting from researchers' network effects and those resulting from their overall research impact. Thus, it is at the intersection of information used to assess research performance and its impact on researchers, research teams, and institutions that evaluate research. It proposes a simple yet unique and innovative approach to using citation count as a basis for researcher impact and network assessment. The proposed metric is novel, replicable, and generalisable. Given the dynamic nature of the information science field, the authors do not claim that the proposed metric or tool is without bias. Instead, they advocate for the importance of this metric in helping institutions better assess researchers they are hiring, promoting, or funding as it is done using other metrics like h-index and i-index. Depending on the institution's objectives, a researcher who acquires more citations through network effects may be more desirable than one with no network effect and vice versa. Thus, the authors would like to challenge other researchers to take an interest in this metric and improve it.

## **Statements and declaration**

We have no known conflict of interest to disclose.

## References

- Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594–607. <https://doi.org/10.1016/j.joi.2011.05.007>
- Abbasi, A., & Jamali, H. R. (2019). University research diversification effect on its citation-based performance: A study of Australian universities. *17th International Conference on Scientometrics and Informetrics, ISSI 2019 - Proceedings, 1*, 489–500.
- Aksnes, D., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1), 215824401982957. <https://doi.org/10.1177/2158244019829575>
- Amjad, T., Shahid, N., Daud, A., & Khatoon, A. (2022). Citation burst prediction in a bibliometric network. *Scientometrics*, 127(5), 2773–2790. <https://doi.org/10.1007/s11192-022-04344-3>
- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975.
- Berche, B., Holovatch, Y., Kenna, R., & Mryglod, O. (2016). Academic research groups: Evaluation of their quality and quality of their evaluation. *Journal of Physics: Conference Series*, 681, 12004–12004. <https://doi.org/10.1088/1742-6596/681/1/012004>
- Bornmann, L. (2017). Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? An analysis based on data from F1000Prime and normalized citation scores. *Journal of the Association for Information Science and Technology*, 68(4), 1036–1047. <https://doi.org/10.1002/asi.23728>
- Bornmann, L., & Daniel, H.-D. (2009). Universality of citation distributions-A validation of Radicchi et al.'s relative indicator  $cf = c/c_0$  at the micro level using data from chemistry. *Journal of the American Society for Information Science and Technology*, 60(8), 1664–1670. <https://doi.org/10.1002/asi.21076>
- Bornmann, L., & Leydesdorff, L. (2013). The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000. *Journal of Informetrics*, 7(2), 286–291. <https://doi.org/10.1016/j.joi.2012.12.003>
- Bornmann, L., & Marx, W. (2014). Distributions instead of single numbers: Percentiles and beam plots for the assessment of single researchers. *Journal of the Association for Information Science and Technology*, 65(1), 206–208. <https://doi.org/10.1002/asi.22996>
- Bornmann, L., Mutz, R., & Daniel, H.-D. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, 59(5), 830–837. <https://doi.org/10.1002/asi.20806>
- Bosquet, C., & Combes, P.-P. (2013). Are academics who publish more also more cited? Individual determinants of publication and citation records. *Scientometrics*, 97(3), 831–857. <https://doi.org/10.1007/s11192-013-0996-6>
- Caon, M., Trapp, J., & Baldock, C. (2020). Citations are a good way to determine the quality of research. *Physical and Engineering Sciences in Medicine*, 43(4), 1145–1148. <https://doi.org/10.1007/s13246-020-00941-9>

Charlton, B. G., & Andras, P. (2007). Evaluating universities using simple scientometric research-output metrics: Total citation counts per university for a retrospective seven-year rolling sample. *Science and Public Policy*, 34(8), 555–563. <https://doi.org/10.3152/030234207X254413>

Chen, S., Song, Y., Shu, F., & Larivière, V. (2022). Interdisciplinarity and impact: The effects of the citation time window. *Scientometrics*, 127(5), 2621–2642. <https://doi.org/10.1007/s11192-022-04338-1>

Costas, R., van Leeuwen, T. N., & Bordons, M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics*, 82(3), 517–537. <https://doi.org/10.1007/s11192-010-0187-7>

Dechow, P. M., Sloan, R. G., & Zeng, J. (Jieyin). (2019). Is It a Home Run? Measuring Relative Citation Rates in Accounting Research. *Accounting Horizons*, 34(1), 67–91. <https://doi.org/10.2308/acch-52570>

Demaine, J. (2022). Fractionalization of research impact reveals global trends in university collaboration. *Scientometrics*, 127(5), 2235–2247. <https://doi.org/10.1007/s11192-021-04246-w>

Didegah, F., & Thelwall, M. (2013). Determinants of research citation impact in nanoscience and nanotechnology. *Journal of the American Society for Information Science and Technology*, 64(5), 1055–1064. <https://doi.org/10.1002/asi.22806>

Downes, B. J., & Lancaster, J. (2020). Celebrating women conducting research in freshwater ecology ... And how the citation game is damaging them. *Marine and Freshwater Research*, 71(2), 139–155. <https://doi.org/10.1071/MF18436>

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152. <https://doi.org/10.1007/s11192-006-0144-7>

Froghi, S., Ahmed, K., Finch, A., Fitzpatrick, J. M., Khan, M. S., & Dasgupta, P. (2012). Indicators for research performance evaluation: An overview. *BJU International*, 109(3), 321–324. <https://doi.org/10.1111/j.1464-410X.2011.10856.x>

Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. *Journal of Informetrics*, 11(3), 823–834. <https://doi.org/10.1016/j.joi.2017.06.005>

Heßler, N., & Ziegler, A. (2022). Evidence-based recommendations for increasing the citation frequency of original articles. *Scientometrics*, 127(6), 3367–3381. <https://doi.org/10.1007/s11192-022-04378-7>

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>

Hirv, T. (2022). The interplay of the size of the research system, ways of collaboration, level, and method of funding in determining bibliometric outputs. *Scientometrics*, 127(3), 1295–1316. <https://doi.org/10.1007/s11192-021-04232-2>

Ioannidis, J., Boyack, K., & Wouters, P. (2016). Citation metrics: A primer on how (not) to normalize. *PLOS Biology*, 14(9), e1002542. <https://doi.org/10.1371/journal.pbio.1002542>

Ioannidis, J. P. A. (2015). A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation. *Journal of Psychosomatic Research*, 78(1), 7–11. <https://doi.org/10.1016/j.jpsychores.2014.11.008>

Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)

- Kacem, A., Flatt, J. W., & Mayr, P. (2020). Tracking self-citations in academic publishing. *Scientometrics*, *123*(2), 1157–1165. <https://doi.org/10.1007/s11192-020-03413-9>
- Kenna, R., Mryglod, O., & Berche, B. (2017). A scientists' view of scientometrics: Not everything that counts can be counted. *ArXiv Preprint ArXiv:1703.10407*.
- Khan, N. R., Thompson, C. J., Taylor, D. R., Venable, G. T., Matthew Wham, R., Madison Michael, L., & Klimo, P. (2014). An analysis of publication productivity for 1225 academic neurosurgeons and 99 departments in the United States: Clinical article. *Journal of Neurosurgery*, *120*(3), 746–755. <https://doi.org/10.3171/2013.11.JNS131708>
- Kostoff, R. N. (2002). Citation analysis of research performer quality. *Scientometrics*, *53*(1), 49–71. <https://doi.org/10.1023/A:1014831920172>
- Larivière, V., Gingras, Y., Sugimoto, C., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, *66*(7), 1323–1332. <https://doi.org/10.1002/asi.23266>
- Liu, J., Chen, H.-H., Ho, M. H.-C., & Li, Y.-C. (2014). Citations with different levels of relevancy: Tracing the main paths of legal opinions. *Journal of the Association for Information Science and Technology*, *65*(12), 2479–2488. <https://doi.org/10.1002/asi.23135>
- Maltseva, D., & Batagelj, V. (2022). Collaboration between authors in the field of social network analysis. *Scientometrics*, *127*(6), 3437–3470. <https://doi.org/10.1007/s11192-022-04364-z>
- Milard, B. (2014). The social circles behind scientific references: Relationships between citing and cited authors in chemistry publications. *Journal of the Association for Information Science and Technology*, *65*(12), Article 12. <https://doi.org/10.1002/asi.23149>
- Mohammed, S., Morgan, A., & Nyantakyi, E. (2020). On the influence of uncited publications on a researcher's h-index. *Scientometrics*, *122*(3), 1791–1799. <https://doi.org/10.1007/s11192-020-03356-1>
- Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology*, *16*(3), e2004089–e2004089. <https://doi.org/10.1371/journal.pbio.2004089>
- Molléri, J. S., Petersen, K., & Mendes, E. (2018). Towards understanding the relation between citations and research quality in software engineering studies. *Scientometrics*, *117*(3), 1453–1478. <https://doi.org/10.1007/s11192-018-2907-3>
- Nightingale, J., & Marshall, G. (2012). Citation analysis as a measure of article quality, journal influence and individual researcher performance. *Radiography*, *18*(2), 60–67. <https://doi.org/10.1016/j.radi.2011.10.044>
- Osterloh, M., & Frey, B. (2015). Ranking games. *Evaluation Review*, *39*(1), 102–129. <https://doi.org/10.1177/0193841X14524957>
- Rüdiger, M. S., Antons, D., & Salge, T.-O. (2021). The explanatory power of citations: A new approach to unpacking impact in science. *Scientometrics*, *126*(12), 9779–9809. <https://doi.org/10.1007/s11192-021-04103-w>
- Sjögarde, P., & Didegah, F. (2022). The association between topic growth and citation impact of research publications. *Scientometrics*, *127*(4), 1903–1921. <https://doi.org/10.1007/s11192-022-04293-x>

Szomszor, M., Pendlebury, D. A., & Adams, J. (2020). How much is too much? The difference between research influence and self-citation excess. *Scientometrics*, *123*(2), 1119–1147. <https://doi.org/10.1007/s11192-020-03417-5>

Tajedini, O., Ghazizade, A., & Sadatmoosavi, A. (2018). Identifying the Effects of Co-authorship Strategies on the Citation-based Performance of Scholars: A Social Networks Analysis. *Journal of Scientometric Research*, *7*(1), 19–28. <https://doi.org/10.5530/jscires.7.1.3>

Thelwall, M. (2018). Microsoft Academic automatic document searches: Accuracy for journal articles and suitability for citation analysis. *Journal of Informetrics*, *12*(1), 1–9. <https://doi.org/10.1016/j.joi.2017.11.001>

Thelwall, M., & Maflahi, N. (2020). Academic collaboration rates and citation associations vary substantially between countries and fields. *Journal of the Association for Information Science and Technology*, *71*(8), 968–978. <https://doi.org/10.1002/asi.24315>

Thompson, D., Callen, E., & Nahata, M. (2009). New indices in scholarship assessment. *American Journal of Pharmaceutical Education*, *73*(6), 111. <https://doi.org/10.5688/aj7306111>

Tian, S., Xu, X., & Li, P. (2021). Acknowledgement network and citation count: The moderating role of collaboration network. *Scientometrics*, *126*(9), 7837–7857. <https://doi.org/10.1007/s11192-021-04090-y>

Urlings, M., Duyx, B., Swaen, G., Bouter, L., & Zeegers, M. (2021). Citation bias and other determinants of citation in biomedical research: Findings from six citation networks. *Journal of Clinical Epidemiology*, *132*, 71–78. <https://doi.org/10.1016/j.jclinepi.2020.11.019>

Vrat, P. (2020). Assessment of research impact through citation analysis: A new approach. *Journal of Advances in Management Research*, *17*(1), 160–172. <https://doi.org/10.1108/JAMR-06-2019-0102>

Waltman, L., & Costas, R. (2014). F1000 recommendations as a potential new data source for research evaluation: A comparison with citations. *Journal of the Association for Information Science and Technology*, *65*(3), 433–445. <https://doi.org/10.1002/asi.23040>

Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, *7*(4), 833–849. <https://doi.org/10.1016/j.joi.2013.08.002>

Wan, X., & Liu, F. (2014a). Are all literature citations equally important? Automatic citation strength estimation and its applications. *Journal of the Association for Information Science and Technology*, *65*(9), 1929–1938. <https://doi.org/10.1002/asi.23083>

Wan, X., & Liu, F. (2014b). WL-index: Leveraging citation mention number to quantify an individual's scientific impact. *Journal of the Association for Information Science and Technology*, *65*(12), 2509–2517. <https://doi.org/10.1002/asi.23151>

Wiechetek, Ł., & Pastuszak, Z. (2022). Academic social networks metrics: An effective indicator for university performance? *Scientometrics*, *127*(3), 1381–1401. <https://doi.org/10.1007/s11192-021-04258-6>

Wildgaard, L., Schneider, J., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics*, *101*(1), 125–158. <https://doi.org/10.1007/s11192-014-1423-3>

Wooldridge, J., & King, M. B. (2019). Altmetric scores: An early indicator of research impact. *Journal of the Association for Information Science and Technology*, *70*(3), 271–282. <https://doi.org/10.1002/asi.24122>

Zhang, X., Xie, Q., & Song, M. (2021). Measuring the impact of novelty, bibliometric, and academic-network factors on citation count using a neural network. *Journal of Informetrics*, 15(2), 101140. <https://doi.org/10.1016/j.joi.2021.101140>

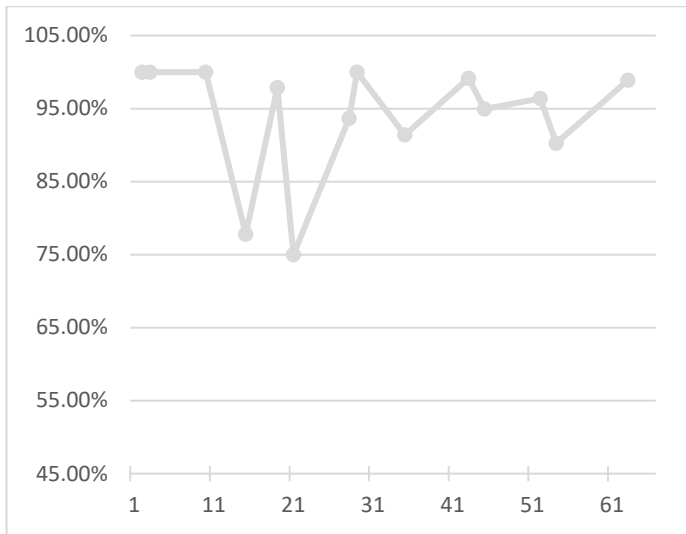
Zhao, D., & Strotmann, A. (2016). Dimensions and uncertainties of author citation rankings: Lessons learned from frequency-weighted in-text citation counting. *Journal of the Association for Information Science and Technology*, 67(3), 671–682. <https://doi.org/10.1002/asi.23418>

Zou, C., & Peterson, J. B. (2016). Quantifying the scientific output of new researchers using the zp-index. *Scientometrics*, 106(3), 901–916. <https://doi.org/10.1007/s11192-015-1807-z>

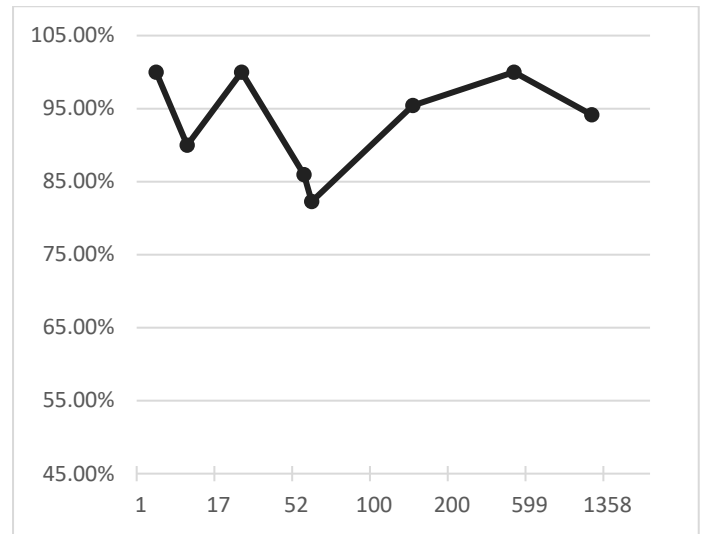
## Appendix 1: Vocabulary

Adjusted citation metric	Also named “adjusted citation count” or “adjusted citation count metric”. It refers to the number of distinct publications that cite one target author and are not co-signed by themselves or co-authors.
Adjusted citation ratio	Also called “adjusted citation count ratio”. It is the ratio of the adjusted citation metric over the raw citation metric. It represents the part of citations made by non-related authors.
Citations	They refer to the mentions made by publications of another publication.
Citation metric	Also called “citation count”, “raw citation metric/count”. It refers to the number of distinct publications that cite one target author.
Co-author	It refers to any researcher that co-signed a publication with the target author at least once.
Main co-author	It refers to the co-authors of the ten most cited (co-authored) publications of one target author on Google Scholar. (NB: publications that are single-authored by the target author are not considered).
Network	For one given target author, this is the set of all his co-authors.
Non-related author	It refers to a researcher who is not a target author’s co-author (i.e. who never co-signed a publication with the target author).
Target author	It refers to the scholar for whom one considers citation counts, co-authors, or non-related researchers. It is the author for whom we compute the adjusted citation metric and ratio.

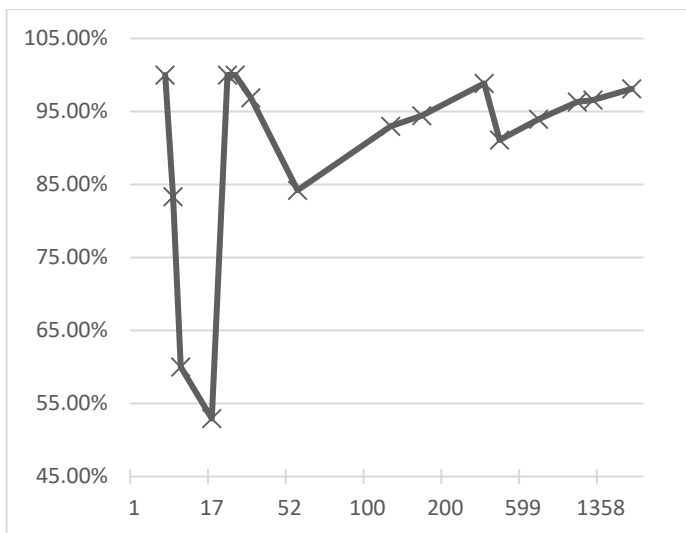
## Appendix 2: Results of the application (focus)



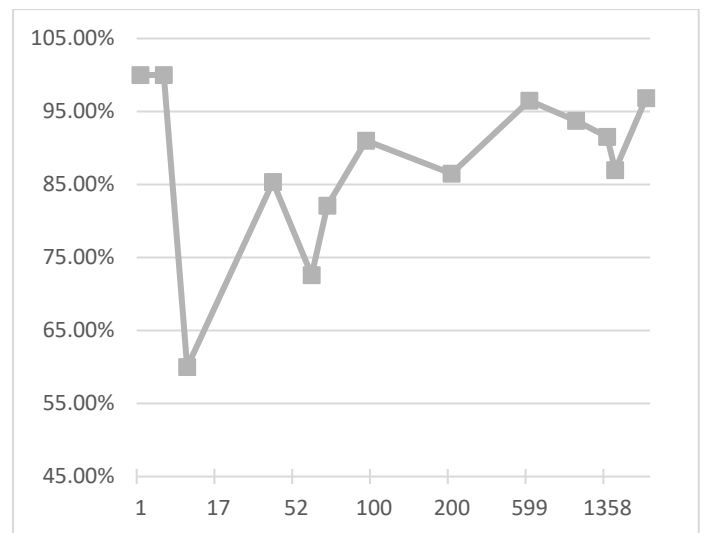
**Fig. 3** Adjusted citation ratio and citation count (data analytics and visualisation)



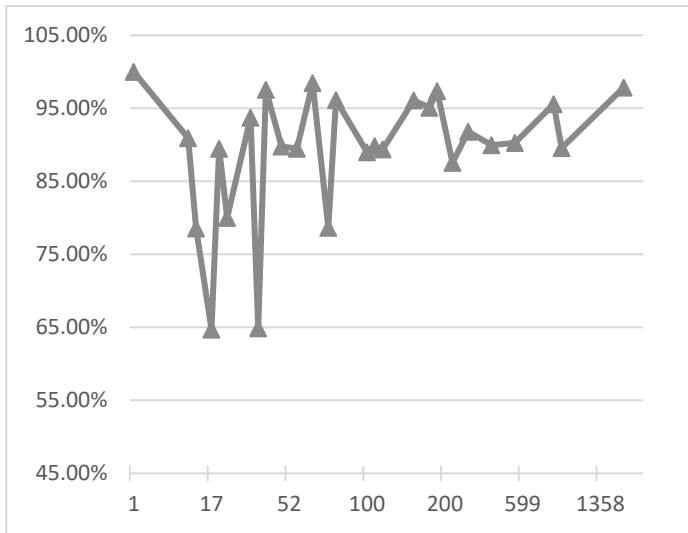
**Fig. 4** Adjusted citation ratio and citation count (entrepreneurial economics)



**Fig. 5** Adjusted citation ratio and citation count (environmental marketing)



**Fig. 6** Adjusted citation ratio and citation count (nanobioengineering)



**Fig. 7** Adjusted citation ratio and citation count (wine economics)

Note: on all these graphs, the x-axis is the citation count, and the y axis is the adjusted citation ratio.