

Article

A Non-Iterative Constrained Measure of Research Impact

Jakub Swacha 

Department of IT in Management, University of Szczecin, 71-004 Szczecin, Poland; jakub.swacha@usz.edu.pl

Abstract: The number of received citations and more complex bibliographic measures calculated based on them, such as the h-index, remain the most widely used indicators for measuring research impact in an objective and easy-to-compute way. However, using the number of received citations as a research impact measure has its shortcomings, some intrinsic (stemming from the doubt whether a citation is actually a confirmation of the cited paper's impact), some extrinsic (stemming from the ease of manipulating this measure by deliberately inserting multiple unmerited references). While the first can only be addressed by a careful interpretation of the measure with consideration of its limitations, the latter can be reduced to much extent by replacing simple citation counting with a more sophisticated procedure constraining the impact that a single publication may have on the others. One such solution is ArticleRank, which has, however, several disadvantages limiting its practical use. In this paper, we propose another solution to this problem, the Transitive Research Impact Score (TRIS), which is free of these disadvantages, and validate it on a sample dataset.

Keywords: scientometrics; citation analysis; research output measure; citation measure



Citation: Swacha, J. A Non-Iterative Constrained Measure of Research Impact. *Information* **2022**, *13*, 319.
<https://doi.org/10.3390/info13070319>

Academic Editor: Sarantos Kapidakis

Received: 1 May 2022
Accepted: 27 June 2022
Published: 29 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Comparing the research impact of scientific works, especially those written on different topics, is a very difficult task. There are a number of reasons, though, for which it has to be done, such as identifying seminal works in a given area, analyzing the growth and decline of certain research topics within a discipline, verifying outputs of research grants, or selecting the best candidates for tenure or scientific awards.

One way of measuring research impact is by expert review [1]. This is, however, a costly and time-consuming procedure that cannot be applied on a large scale. Moreover, it is flawed by the subjectivity of experts, who may favor works closer to their own research agenda or give unreliable opinions on works far from their own research agenda that they do not comprehend but would not admit for the fear of losing the trust in their overall competence and/or the expert position.

In scientometrics, the basic measure of research impact is the number of citations received by the paper or author under evaluation (see [2] and works cited therein). Thanks to the availability of large bibliographic databases, it is easy to obtain (provided the measurement is limited to a given bibliographic database) and objective (i.e., not directly affected by any expert's subjective opinion). Nonetheless, there are a few problems with this measure. First of all, there are different types of citations (see [3] and works cited therein); so, as aptly put by Ross, "counting citations seems like collecting coins in your wallet without distinguishing between quarters, nickels and dimes" [4]. Moreover, as shown by Bornmann and Daniel, publications are cited not only to acknowledge their influence, but also for other, including non-scientific, reasons [5].

Since the number of citations is used to evaluate the impact of researchers and journals, there is a growing problem of adding irrelevant and superfluous citations in a deliberate attempt to artificially inflate the measure value of the target works (and thus their authors and publication venues) [6]. Different pathologies of this kind have been observed [6], such as excessive self-citations, bi-directional mutual citations, chains, triangles, mesh, and even citation cartels, which perform manipulation on a large scale. Despite the ongoing

research effort in revealing citation manipulation, it can never be fully effective, as the same symptoms may result from completely legit reasons and fair intentions: one author's research on a single topic continued for a long term will inevitably lead to multiple self-citations, a scientific discourse between two researchers will end up with multiple bi-directional mutual citations, and a research community centered around a narrow problem area will result in dense citation networks of a larger degree.

The main reason making the citation number prone to manipulation is the combination of two facts: (1) adding one more reference to a scientific paper costs nothing, and (2) every citation, merited or not, counts as one. While we cannot do anything with the former, the latter can be countered through:

- making the measure transitive, so that publications receiving more attention themselves would impact the measures of its references more than those receiving little attention;
- constraining the total impact a single publication has on the measures of its references.

The first of the above can be obtained by making the impact score a publication passes on to its references proportional to the impact score it receives. The second of the above can be attained by following the simple rule: the more references one paper has, the less it adds to their respective scores. Note that this does not differentiate merited and unmerited citations; however, there is a valid reason for reducing the measured impact of publications having a long list of references, even if they are legit: after all, the average impact of an individual publication from such a long list could not be large.

In 2009, Li and Willett proposed ArticleRank [7], a metric alternative to the citation number, which, to some extent (we shall elaborate on that in the following section), addressed the two issues described above. Based on the PageRank algorithm, it performs the calculation in an iterative way, performing several passes over the dataset. Additionally, ArticleRank can only be compared among publications from one dataset. In this paper, we propose a novel measure also thought of as an alternative for the citation number, which is free from these drawbacks, while still being both transitive and constraining the impact of a single publication.

The new measure, called the Transitive Research Impact Score (TRIS), will be described in full detail in Section 3, after which we perform its validation using a sample dataset and then conclude with the main outcomes. Before that, in the following section, we present the related work on research impact measures.

2. Related Work

In 2005, Hirsch identified the following weaknesses of the citation number as a measure of an author's research impact: being hard to find (if one wants to consider all works of a given author rather than only those covered by some source); being prone to inflation by a small number of "big hits", which may not be representative of the individual if they are one of its many coauthors; and giving undue weight to highly cited review articles versus original research contributions [8]. As a solution, he proposed index h , defined as the number of papers with citation number $\geq h$. The h -index has currently become the primary measure of scientists' research impact, and is also used to measure the research impact of publication venues.

Several authors have proposed improvements to the h -index, which, in spite of their merits, have not yet managed to gain similar popularity. In 2006, Egghe introduced the g -index aimed to measure the global citation performance of a set of articles; having the set ranked descendingly by the number of received citations, the g -index is the largest number such that the top g articles together received at least g^2 citations [9]. The same year, Kosmulski introduced the $h(2)$ -index, defined as the highest natural number such that an author's $h(2)$ most-cited papers each received at least $h(2)^2$ citations [10]. Compared to the g -index, it thus gives less favor to authors publishing few papers with a high number of citations per paper. In 2007, the indices R and AR were proposed: the R -index is defined as the square root of the sum of citations received by the publications included in the h -index (to give a more precise measurement of one's impact), and its modification, the AR -index,

divides the number of citations of each publication by their age before summing them up, in an attempt to measure the current rather than historical impact of a researcher [11]. In 2009, the weighted h -index (or w -index) was proposed, in which only the first and corresponding authors receive the full impact score from a publication, whereas the score of the remaining authors decreases linearly with their increasing rank in the list of authors [12].

Both the h -index and its improvements are based on the number of citations of respective publications as their input, so they are, similarly to the citation number, prone to manipulation, as described in Section 1. They can also only be applied to measure the impact of a set of publications (e.g., written by one author or published at a single venue), not of an individual publication.

In contrast, the Φ measure proposed by Aragón can be applied to individual publications as well. Instead of counting citations, it counts the number of unique first authors of papers citing a given work [13]. While solving many deficiencies of using the citation number as a research impact measure, it introduces some of its own, such as considering only the first author and ignoring, in the calculation of the measure, multiple works from the same first author, no matter how important they could be for the field.

Another measure that can also be applied to individual publications is ArticleRank [7]. It is defined by the following formula (originally given in [7] (p. 4); here, we follow the variable naming of the only known publicly available implementation of ArticleRank in Neo4j Graph Data Science [14]):

$$ArticleRank_i(v) = (1 - d) + d \sum_{w \in N_{in}(v)} \frac{ArticleRank_{i-1}(w)}{|N_{out}(w)| + \overline{N_{out}}} \quad (1)$$

where:

- $ArticleRank_i(v)$ denotes the ArticleRank score of a paper v in iteration i ;
- d is a damping factor (by default set to 0.85, alike PageRank [15]);
- $N_{in}(v)$ is the set of papers having references to v in the citation network;
- $|N_{out}(w)|$ is the number of papers referenced by w ;
- $\overline{N_{out}}$ is the mean references number averaged over all papers in the network.

Note that the damping factor defines how fast the impact score of a paper spreads to the works it references, the inclusion of $|N_{out}(w)|$ in the denominator makes the impact score passed by a paper decreasing with the number of the works it references, and the inclusion of the mean value of N_{out} in the denominator reduces the impact score passed by papers having only a few references.

As can be observed in Formula (1), in a single pass, the impact score is only passed to the direct neighbors of the citing paper. For this reason, the calculation of ArticleRank is iterative, with the calculations being repeated multiple times over the same dataset. For instance, in the case of a small sample dataset containing merely 354 papers (including only 195 that had been cited at least once) and 752 citation linkages, it took 40 iterations before the measure values stabilized [7] (p. 6). This means it depends on the size and structure of the citation graph under measurement, and it cannot be known in advance, so the default value of 20 iterations used in Neo4j Graph Data Science [14] does not define a standard of reference, but is merely based on an assumption that 20 iterations are enough for the measure values to stabilize, which is often not enough (as exemplified by [7]).

Additionally, the value of the damping factor is not actually fixed, as, according to [14], it should be adjusted to help counteract known issues with applying the ArticleRank algorithm: spider traps (when there are no references going out from a group of papers), rank sinks (when a network forms an infinite cycle), and dead-ends (when individual papers have no references).

Results obtained using a different number of iterations or damping factor values cannot be compared, but even if these parameters are set to the same values, ArticleRank values cannot still be compared across various datasets due to the inclusion of the mean

number of references ($\overline{N_{out}}$) in Formula (1). We consider this comparability limitation the main drawback of ArticleRank.

The next section introduces a new measure, called the Transitive Research Impact Score, providing advantages over the citation number similar to ArticleRank but free of the disadvantages of the latter.

3. Transitive Research Impact Score

The Transitive Research Impact Score (TRIS) is conceived as a measure of research impact alternative to the citation number. It aims to (1) link the measured impact a publication has on its references with the impact it has received itself, as well as to (2) constrain a publication's total impact so that it would not grow proportionally to the number of its references, as is the case with the citation number.

Moreover, TRIS is non-iterative. In order to explain how this is accomplished, one should observe that, as a rule, a citation is a one-directional link from a paper published later to another paper published earlier. Therefore, although to meet Design Aim 1, the value of TRIS of the citing paper has to be known in advance to calculate the value of TRIS of the cited paper, if we start calculating the measure from the latest papers in a dataset, continuing in the order of decreasing time of publication, then it could be calculated for every paper in the dataset.

In the real world, unfortunately, the rule mentioned above is not always preserved, as, sometimes, papers published earlier cite ones published later. There are several possible reasons for this, e.g.:

- the cited paper was known to the author of the citing paper already at the moment of writing the latter (e.g., as a preprint or through private correspondence), who also knew where it was scheduled for publication and thus could provide a future reference;
- the cited paper was actually published before the citing paper, but either the publisher of the former specified the publication year of $X-1$ instead of X for some reason (e.g., a delayed publication, which is a problem in many journals) or the publisher of the latter specified the publication year of $X+1$ instead of X for some reason (e.g., marketing, which applies to many book publishers who can thus advertise a book as new for a whole year longer).

Situations like the above described could lead to citation cycles (in the simplest case, Publication A citing Publication B, and Publication B citing Publication A as well), which would make the proposed non-iterative way of calculating the measure infeasible. For the sake of avoiding this, we propose a simple solution by dividing every paper's impact passed to its references into two parts, of which one is fixed, and the other variable, dependent on the impact the paper received itself. Both parts contribute to the impact score of the referenced works only if they were published before the referencing paper; otherwise, only the fixed part is considered. This way, no cycles may happen, and the only drawback is reducing the relative impact of future references. Yet, such situations are rare; therefore, on average, the effect on the calculated measure is negligible. Moreover, it could be argued that the relative impact of references to future work should indeed be smaller than that to past work (especially since, in practice, many future references merely announce a forthcoming publication on the same topic rather than give credit to it), which would make this trait an advantage rather than a drawback. This is even more of an advantage if we consider that it reduces the impact of self-citations on one's own recent or future works.

The implementation problem to solve is same-year references. Most of them are ordinary references to already-published work (e.g., a paper published in August cites one published in January); however, they may also include future references, so cycles are possible. If we could sort the papers depending on the exact date of their publication, this problem could be solved as in the general case; however, in practice, such precise information is usually available only for a small part of publications. Therefore, to solve this problem, publications are processed and arranged using the following three sorting keys:

- year of publication, descendingly;
- number of citations from the same year, ascendingly;
- total number of citations, ascendingly.

By using the number of citations from the same year as the second key, the publications that could only cite other same-year publications are processed before those that could probably be also cited by other same-year publications, and among the latter, those less probable to be cited by other same-year publications are processed before those more probable. To resolve such drawbacks, the total number of citations is used as the third sorting key. If all three sorting keys are equal, the order of the original dataset is kept; such a situation was found to be very rare, so it does not give reasons for complicating the sorting by including even more keys.

Consequently, the procedure for calculating TRIS can be specified as follows:

1. For each paper p in the dataset, we set its initial received impact score to 0, both in its fixed ($fi(p)$) and variable ($vi(p)$) parts; we also zero its citation count and same-year citation count.
2. For each paper p in the dataset:
 - 2.1. We obtain its impact score transition divisor $td(p)$, which is double the number of paper p 's references $cr(p)$ if $cr(p)$ is larger than 9, or the sum of 9 and $cr(p)$ otherwise.
 - 2.2. For each paper r referenced by paper p :
 - 2.2.1. We increase the referenced paper's fixed score $fi(r)$ by the reciprocal of $td(p)$;
 - 2.2.2. We increase the referenced paper's citation count by 1;
 - 2.2.3. If the publication year of paper p is equal to the publication year of paper r :
 - 2.2.3.1. We increase the referenced paper's same-year citation count by 1.
3. We sort the papers using three keys: (year of publication (in descending order), number of citations from the same year (in ascending order), total number of citations (in ascending order)).
4. For each paper p in the dataset, following the sorted order:
 - 4.1. For each paper r referenced by paper p :
 - 4.1.1. We increase the referenced paper's variable score $vi(r)$ by the impact score received so far by paper p (i.e., the sum of $fi(p)$ and $vi(p)$) divided by $td(p)$.
5. For each paper p in the dataset, $TRIS(p) = fi(p) + vi(p)$.

The number 9 (Point 2 of the procedure) is used to curb the impact score passed from papers with just a few references. Thanks to this, a paper having only one reference will pass only 10% of the impact score it itself received. This way, unlike ArticleRank [7], TRIS does not depend on the mean number of references in the considered dataset, making it comparable across different datasets. See Table 1 for the comparison of this and other aspects among the three measures.

Table 1. Comparison of the key traits of Citation Number, ArticleRank, and TRIS.

Trait	Citation Number	ArticleRank	TRIS
Comparable across datasets	Yes	No	Yes
Requires multiple iterations to calculate	No	Yes	No
Passes received impact to its references	No	Yes	Yes
Passing of the impact is limited in range	N/A	By the number of iterations	No
Limits the total impact of one paper	No	Yes	Yes
Gives value 0 to non-cited papers	Yes	No	Yes
Reduces the impact of future references	No	No	Yes

Also note the distinction between increasing the fixed part of the impact score (for which the citing paper's impact score is irrelevant, Point 2.2.1 of the procedure) and the variable one (which depends on the citing paper's impact score, Point 4.1.1 of the procedure). Considering that, TRIS can be defined using the following formula:

$$TRIS(p) = \sum_{c \in C(p)} \frac{1 + \delta(p, c) \cdot TRIS(c)}{cr(c) + \max(9, cr(c))}, \quad (2)$$

where:

- $TRIS(p)$ denotes the TRIS score of a paper p ;
- $C(p)$ denotes the set of papers having references to p ;
- $\delta(p, c)$ denotes the estimated time difference between papers p and c ; it equals 1 if paper p has been published before paper c , and 0 otherwise (which means that, in such a case, the value of $TRIS(c)$ is irrelevant); the estimation is based on the value of three sorting keys (year of publication (in descending order), number of citations from the same year (in ascending order), total number of citations (in ascending order));
- $cr(c)$ denotes the number of references of paper c .

4. Validation

In order to validate TRIS, we have implemented it in Python programming language; the full source code used to calculate TRIS is available at <https://github.com/jakubsw/tris> (accessed on 29 June 2022). The presented implementation expects the input citation graph data in JSON format generated by the 3dSciLi web tool [16].

The preliminary validation was aimed at checking whether TRIS gives expectable values in basic situations. For this purpose, three rudimentary citation graphs were prepared. They are listed in Table 2, together with measured values of the citation number, ArticleRank, and TRIS.

Table 2. Comparison of measurements on artificial graphs.

Citation Graph	Node	Citation Number	ArticleRank	TRIS
A → B → C (B precedes A in time, C precedes B in time)	A	0	0.15	0.00
	B	1	0.24	0.10
	C	1	0.29	0.20
A → B ↑ ↓ D ← C (A, B, C, D from the same year)	A	1	0.30	0.2111
	B	1	0.30	0.2000
	C	1	0.30	0.2100
	D	1	0.30	0.2110
A ₁ ... A ₁₀ → C B ₁₁ ... B ₁₃ → D ₁ → ⋮ ⋮ ⋮ ↓ B ₂₁ ... B ₂₃ → D ₃ → E ⋮ ⋮ ⋮ ↑ B ₅₁ ... B ₅₃ → D ₅ →	A	0	0.15	0.00
	B	0	0.15	0.00
	C	10	0.92	1.00
	D	3	0.38	0.30
	E	5	1.14	2.00

Legend: A, B, C, ... —considered publications; A → B—the citation of publication B by publication A.

The first of the citation graphs in Table 2 is the simplest one; it illustrates passing the impact node B received from Node A to its reference, Node C by both ArticleRank and TRIS, in contrast to the citation number, in the case of which the values for B and C are equal. Compared to ArticleRank, TRIS passes a noticeably larger share of the received impact; note though that it is relatively so high due to both A and B having just one reference, which would be very rare in citation graphs based on real-world data.

The second of the citation graphs in Table 2 presents a loop, which, as explained in the previous section, and unlike the other measures, cannot be handled as such by TRIS’ calculation algorithm, and is therefore broken (at Element A, making Element B the beginning of the chain of references). As we can observe, the effect of TRIS’ limitation on the measurement is very small (the precision of decimal numbers was increased here to precisely show the differences between the respective nodes).

The third of the citation graphs in Table 2 compares the measurements between Node C, receiving 10 citations from nodes that were not cited themselves, and Node E, receiving only 5 citations but from nodes that were cited 3 times each. While Citation Number measures the impact of C to be twice as high as of Node E, both ArticleRank and TRIS suggest a reverse ranking, with the latter giving much more weight to the impact of the citing works. This observation confirms TRIS to be an effective tool for diminishing the impact of a large number of papers, which have received little to no impact themselves.

The goal of the main validation was to check whether TRIS provides discriminating ability not worse than Citation Number. Discriminating ability is the key feature of any research impact measure, as it denotes the measure’s ability to identify a difference between any two compared papers. Poor discriminating ability hampers the use of a measure for its primary purpose, i.e., to construct rankings; an extreme example of this would be a measure returning the same fixed number for each of the considered papers: all of them would then have to be assigned the same rank.

The main validation has been performed on an exemplary real-world dataset containing the first 2000 results of a query on “citation metric” obtained from the Scopus bibliographic data platform [17] with default search settings (i.e., title/abstract/keywords). The obtained dataset contained 1973 works considered unique and 2962 citations within the dataset. For a glimpse of its complexity, Figure 1 shows the central part of its citation graph visualized with the 3dSciLi web tool [16]; publications are depicted as circles whose size is proportional to the number of received citations, whereas color denotes the age of publication (from blue—the oldest, to red—the latest) and links between circles depict citations. We consider the size and complexity of the exemplary dataset as sufficient for the purposes of the validation.



Figure 1. Three-dimensional citation graph of the exemplary dataset.

In Figure 2, we compare the alternative measures for each paper of the exemplary dataset receiving at least one citation; for non-cited papers, both TRIS and the citation number equal 0 (and ArticleRank equals 0.15), so there was no point in including them in the chart. Note that both axes are on the logarithmic scale, albeit of different bases: 2 in the case of the citation number, and 10 in the case of TRIS and ArticleRank. Also note that the points denoting ArticleRank were painted after the points denoting TRIS, so they may conceal some of them.

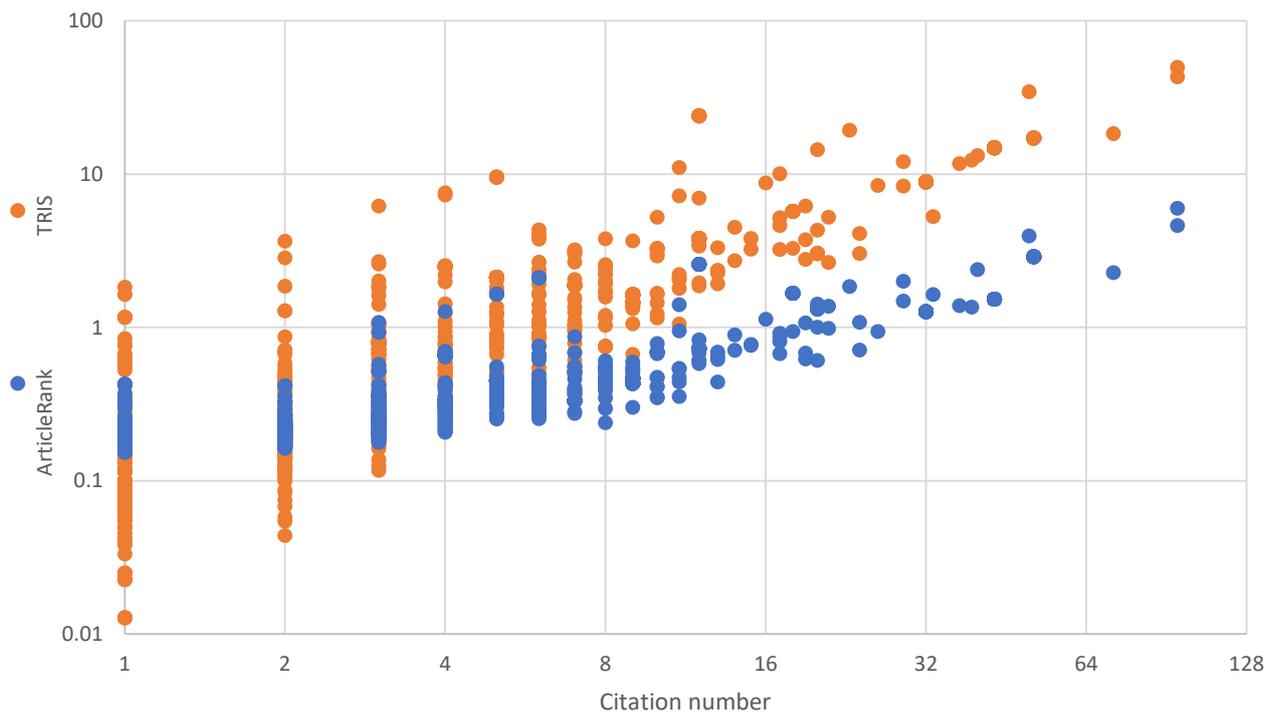


Figure 2. Comparison of measures obtained for all papers of the exemplary dataset.

As can be observed in Figure 2, as expected the three measures are visibly correlated, as the growth of the citation number is, in general, parallel with the growth of TRIS (this is also confirmed by the average TRIS values growing with the number of citations, see below) and ArticleRank.

However, TRIS measurements form a noticeably different pattern compared to the citation number. The most evident is the difference for publications having a low citation number, based on which multiple publications are placed at the same positions with regard to the horizontal axis, whereas their TRIS value is spread over a considerable range; for instance:

- 259 publications having a citation number of 1 were assigned 74 distinct TRIS values (ranging from 0.013 to 1.837; 0.148 on average);
- 130 publications having a citation number of 2 were assigned 104 distinct TRIS values (ranging from 0.044 to 3.652; 0.325 on average);
- 26 publications having a citation number of 5 were all assigned distinct TRIS values (ranging from 0.459 to 9.565; 1.414 on average);
- 9 publications having a citation number of 10 were all assigned distinct TRIS values (ranging from 1.155 to 5.250; 2.408 on average);
- 4 publications having a citation number of 20 were all assigned distinct TRIS values (ranging from 3.035 to 14.480; 6.220 on average).

We thus conclude that TRIS has much superior discriminating ability than the citation number, as not only almost all papers having at least a moderate number of citations are assigned a unique TRIS value, but even the lowly cited papers share their TRIS value usually with just a few other papers rather than tens or even hundreds of them.

The discriminating ability of TRIS is also superior to ArticleRank, as clearly indicated in Figure 2 by much larger value ranges, especially evident for papers having a citation number of less than 4.

The presented ranges of TRIS values assigned to papers having the same citation number also confirm that the intended mechanism of passing impact score acquired by a paper to its references works, and results in papers having fewer citations yet from high-impact sources receiving much higher TRIS values than papers cited by more yet low-impact papers. We can also confirm that the mechanism capping the impact one paper may have on other works as well, as the TRIS values of papers having just a few citations yet from high-impact papers, although high relative to those receiving also few citations yet from low-impact papers, are still much smaller than the TRIS values of the high-impact papers citing them.

Regarding the simplification introduced to the procedure in order to avoid having to perform multiple passes over the data, among the 2962 references in the exemplary dataset, there was not a single obvious future reference (i.e., having the cited publication year higher than the citing publication year) and 146 same-year references (4.9%). We have also established the maximum number of same-year references of one publication, which was 5 in the exemplary dataset (for 66% of papers having same-year references, it was 1). This means that, in realistic cases, thanks to the sorting procedure based on three keys, the effect of the introduced simplification on the score of high-impact publications (which are of primary interest, as most analyses focus on top-ranked papers) is negligible.

5. Discussion

The deficiencies in measuring the research impact with the number of received citations gave rise to a number of alternative measures. There is a growing trend for consideration of metrics alternative to citations, known as *altmetrics* [18], counting actions performed on articles, such as views, mentions, or recommendations. Altmetrics, however, have their own issues, such as source data volatility [19], gender and regional biases [20], or being prone to manipulation by web bots [20], and, most of all, they try to measure the research impact on society in general rather than scientific community [21].

A more conservative approach is to calculate synthetic measures based on the citation number, which are free from some of its disadvantages, such as the *h*-index [8] and its modifications [9–12]. As they all involve some kind of data aggregation, these indexes can only be applied to measure the research impact of authors or journals, but not of individual publications. One proposal aimed at measuring also the impact of individual publications (the Φ measure [13]) uses just a small subset of available data (the first author of citing publications), vastly limiting its discriminating ability.

Another proposal in this vein is ArticleRank [7], based on the PageRank algorithm [15], whose effectiveness has been well proven by web search engines. Its key advantage is weighting the value of received citations depending on the research impact of the citing paper. ArticleRank, though, is limited in its comparability, as the obtained values depend on the properties of the dataset, in which the measurement is made, and the number of iterations, which, in turn, depends on the size and structure of a particular citation graph.

Publication citations are not, however, web links. Whereas it is common for several web pages devoted to similar topics to have links to each other, research papers generally cite only papers published earlier. This observation suggests that there is no need to perform multiple iterations, as in ArticleRank, to obtain a transitive measure, which would consider the citing papers' research impact in assessing the research impact of publications they reference; it is only needed to perform the calculation on papers in the dataset in the order of their respective publication time.

This concept has been implemented in the proposed TRIS measure. It retains the key virtues of ArticleRank (being transitive and constraining the total impact a single publication may have on the papers it cites), but is free of its shortcomings (requiring multiple iterations to calculate and not enabling measurement comparisons across different datasets). As presented in the previous section, TRIS also excels ArticleRank in discriminating ability, greatly reducing the number of papers receiving the same impact score.

The main limitation of TRIS regards the way it handles future references, which do not receive any part of the impact received by a citing paper. Although this could be considered even as an advantage, the actual problem is that for most real-world datasets, it is impossible to determine the exact publication date, hence identifying a same-year reference as a future one. To overcome this, TRIS estimates the publication order of same-year publications with the number of same-year citations and the total number of citations, which solves the problem in most practical situations, but not all of them. Nonetheless, we consider the risk of notably affecting the impact score due to this reason as negligible in real-world datasets.

TRIS, to a great extent, weakens the effect of manipulating the citation number by including multiple references in low-impact works. For this reason, we envisage its use as a measure complementary to the citation number, capable of verifying the quality of citations indicated by a high citation number. We do not see TRIS as a replacement for the citation number for several reasons, primarily the latter being natural, simpler, and widespread.

What can be seen as another limitation is that TRIS is a measure of the actual research impact, not an estimator of potential research impact. For example, a newly published breakthrough paper will have little of the former, but a lot of the latter. Estimating the latter is, however, a non-trivial problem, and there is a long ongoing debate on how to solve it. A good introduction to this vein of research is provided by Wang et al., who both review prior literature on the topic and also describe a solution exploiting simultaneously citations, authors, journals, and time information [22].

A limitation that TRIS shares with all measures based on bibliographic data is being prone to data quality issues typical for such data. These include not only common editing errors, but also problems with distinguishing authors having the same names, or a single author using various names or name forms. Nine types of such issues were identified by Azeroual and Lewoniewski [23].

6. Conclusions

In this paper, we introduced the Transitive Research Impact Score (TRIS), a measure of the research impact of scientific publications alternative to the number of citations. The measure is less prone to manipulation by unmerited citations, as it values citations from high-impact publications much higher than from low-impact publications, and it constrains the impact that a single publication may have on the others.

Compared to another measure having similar advantages, ArticleRank, it allows measurement comparisons across different datasets and has superior discriminating ability, allowing one to distinguish the impact scores of tens of papers having the same number of citations. It is also much easier to compute, as it does not require multiple passes over the data. This is possible thanks to a simplification of the measure calculation procedure, at the cost of reducing the impact score generated by future references. The results obtained on an exemplary dataset show that the effect of this simplification is negligible for typical cases.

Our future work will be to compare the results based on the citation number, TRIS, and possibly other measures obtained on various large datasets. This could reveal the extent to which TRIS is capable of identifying truly important research works more than the citation number, as well as indicating manipulation of the citation number by unmerited references.

Funding: This research received no external funding.

Data Availability Statement: Exemplary data used for validation were obtained from a generally accessible commercial source (Scopus); their copy is available from the author under conditions set by their original provider.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Taylor, J. The Assessment of Research Quality in UK Universities: Peer Review or Metrics? *Br. J. Manag.* **2011**, *22*, 202–217. [CrossRef]
2. Herrmannova, D.; Patton, R.M.; Knoth, P.; Stahl, C.G. Do citations and readership identify seminal publications? *Scientometrics* **2018**, *115*, 239–262. [CrossRef]
3. Swacha, J. A simple taxonomy for computer science paper relationships. *Studia Inform.* **2013**, *32*, 81–92.
4. Ross, H. Citation Bribery. *Chem. Eur. J.* **2019**, *25*, 13449–13451. [CrossRef] [PubMed]
5. Bornmann, L.; Daniel, H. What do citation counts measure? A review of studies on citing behavior. *J. Doc.* **2008**, *64*, 45–80. [CrossRef]
6. Chakraborty, J.; Pradhan, D.; Dutta, H.S.; Nandi, S.; Chakraborty, T. On Good and Bad Intentions behind Anomalous Citation Patterns among Journals in Computer Sciences. *arXiv* **2018**, arXiv:1807.10804.
7. Li, J.; Willett, P. ArticleRank: A PageRank-based alternative to numbers of citations for analysing citation networks. *Aslib Proc.* **2009**, *61*, 605–618. [CrossRef]
8. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [CrossRef] [PubMed]
9. Egghe, L. Theory and practise of the g-index. *Scientometrics* **2006**, *69*, 131–152. [CrossRef]
10. Kosmulski, M. A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI Newsl.* **2006**, *2*, 4–6.
11. Jin, B.; Liang, L.; Rousseau, R.; Egghe, L. The R- and AR-indices: Complementing the h-index. *Chin. Sci. Bull.* **2007**, *52*, 855–863. [CrossRef]
12. Zhang, C.-T. A proposal for calculating weighted citations based on author rank. *EMBO Rep.* **2009**, *10*, 416–417. [CrossRef] [PubMed]
13. Aragón, A. A measure for the impact of research. *Sci. Rep.* **2013**, *3*, 1649. [CrossRef] [PubMed]
14. Neo4j GDS: Article Rank. Available online: <https://neo4j.com/docs/graph-data-science/current/algorithms/article-rank/> (accessed on 29 June 2022).
15. Page, L.; Brin, S.; Motwani, R.; Winograd, T. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab Technical Report 1999-66. Available online: <http://dbpubs.stanford.edu:8090/pub/1999-66> (accessed on 22 May 2022).
16. Swacha, J. Three dimensions of science: A web tool for 3D visualization of scientific literature. In Proceedings of the 2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Champaign, IL, USA, 27–30 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 274–277. [CrossRef]
17. Scopus. Available online: <https://www.scopus.com> (accessed on 22 May 2022).
18. Williams, A.E. Altmetrics: An Overview and Evaluation. *Online Inf. Rev.* **2017**, *41*, 311–317. [CrossRef]
19. Fang, Z.; Dudek, J.; Costas, R. Facing the Volatility of Tweets in Altmetric Research. *J. Assoc. Inf. Sci. Technol.* **2022**. [CrossRef]

20. Chapman, C.A.; Hemingway, C.A.; Sarkar, D.; Gogarten, J.F.; Stenseth, N.C. Altmetric Scores in Conservation Science Have Gender and Regional Biases. *Conserv. Soc.* **2022**, *286*, 20192047. [[CrossRef](#)]
21. Thelwall, M. Measuring Societal Impacts of Research with Altmetrics? Common Problems and Mistakes. *J. Econ. Surv.* **2021**, *35*, 1302–1314. [[CrossRef](#)]
22. Wang, Y.; Tong, Y.; Zeng, M. Ranking Scientific Articles by Exploiting Citations, Authors, Journals, and Time Information. *Proc. AAAI Conf. Artif. Intell.* **2013**, *27*, 933–939. [[CrossRef](#)]
23. Azeroual, O.; Lewoniewski, W. How to Inspect and Measure Data Quality about Scientific Publications: Use Case of Wikipedia and CRIS Databases. *Algorithms* **2020**, *13*, 107. [[CrossRef](#)]