Quantifying citation inflation: measurement methods and their implications for research evaluation

Gege Lin[*, **], Guoqiang Liang[***], Haiyan Hou[*], Zhigang Hu[*]

[*]*lingegenl@mail.dlut.edu.cn; houhaiyan@dlut.edu.cn; huzhigang@dlut.edu.cn*
0000-0003-0137-1471; 0000-0002-2790-9973;0000-0003-1835-4264;
WISE Lab, Dalian University of Technology, Dalian (China)

[**] *g.lin@cwts.leidenuniv.nl;*
0000-0003-0137-1471;
Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

[***]*lianggq@bjut.edu.cn*
0000-0002-9669-4048
School of Economics and Management, Beijing University of Technology, Beijing (China)

Abstract: With the exponential growth of scientific literature and changing citation behavior among scholars, citation inflation has become an increasingly important issue in research evaluation. In this study, we propose a new citation deflator based on reference growth rate (CDR) instead of using the publication rate as a citation deflator (CDP). Our comparative analysis of the two deflators shows that failing to deflate citation counts leads to significant underestimation of the scientific impact of prior generations. Moreover, CDP overemphasizes the value of early citations and underestimates the value of recent citations, resulting in over-amplification of early citations. Therefore, we suggest a systemic overhaul of citation counting methods and the creation of a citation deflator to mitigate inflationary temporal bias.

1. Introduction

Citation analysis is a crucial tool for assessing the influence of scientific research and can inform decision-making in various contexts, including academic hiring, funding allocation, and policy development (Dangzhi Zhao et al., 2020; Fortunato et al., 2018; Petersen et al., 2019; Thelwall, 2019). The citation of a paper can be used to gauge the impact and value of a research output in the academic community, as well as its influence on future research (Huang et al., 2022; Petersen et al., 2014). Furthermore, citation analysis can aid in identifying academic development trends and research hotspots in various fields. It can also assist scientists in assessing the impact and visibility of their research output, which is crucial for their academic career and professional advancement (D. Wang et al., 2013).

However, citation inflation has become an increasingly important issue in research evaluation due to the exponential growth of scientific literature and the changing citation behavior of scholars(Marks, 2022; Teixeira da Silva & Dobránszki, 2018). The scientific production is steadily increasing, with publications exhibiting an annual growth of 4% and the number of references per publication increasing at a rate of 1.8% per year (Petersen et al., 2019). Pan et al. (2018) show that the top 1% of Science publications from 2000 had more than 100 citations as of 2005, whereas the top 1% of publications from 1965 had only 50 citations as of 1970. The problem is actually quite simple - when citations are produced in different historical periods, their "real values" are inconsistent and thus cannot be simply added together. This is similar to adjusting real prices for monetary inflation. The consequences of this problem have been observed in recent studies analysing the citation dynamics of individual careers and publications throughout the citation life cycle (Petersen et al., 2014, 2019; Yin & Wang, 2017). As a result, this issue is likely to have an impact on quantitative evaluation in hiring and promotion decisions, as well as the study of the scientific enterprise itself, when methods rely on cumulative citation counts or longitudinal citation analysis (Fortunato et al., 2018).

Based on the growth rate of academic papers, Petersen et al. (2019) proposed a citation deflator, that does not rely solely on reference increase. They concluded that the "real value" of a citation is halved in relative terms every 16 years. However, it should be noted that the increase of scientific publications reflects the speed of obsolescence of scientific literature (Burton & Kebler, 1960) rather than the rate of citation value decline. Similar to the Price index, citation inflation should be measured based on changes in reference counts. Just like Sleeping Beauty literature (Van Raan, 2004), which refers to delayed recognition literature, they may be old but still have value. Although these articles may have been overlooked or received insufficient attention after publication for various reasons, they can be rediscovered and become important references as time goes by and the academic environment changes (Miura et al., 2021). In addition, the value of Sleeping Beauty literature may diminish over time as similar works emerge. While some Sleeping Beauty literature may represent significant contributions to the field, it is important to contextualize these works within the broader scientific landscape and consider their citation value in light of newer research.

Therefore, we propose a new citation deflator based on reference growth rate (CDR), which balances the effects of citation inflation and time on citation counts. By considering temporal bias and CDR, the citation deflator provides a more refined and accurate method for measuring the impact and influence of papers, and can contribute to improving the evaluation and decision-making of scientific research.

2. Material and methods

We analyzed all English-language publications (articles) in the Clarivate Analytics Web of Science (WoS) database that were published between 1980 and the end of 2022. The WoS data was collected and parsed by CWTS of Leiden University. For each item, we extracted its publication year, WoS category (WC), country, etc. We used the Leiden classification system to group individual publications into broad research domains according to direct citation links between articles, using the methodology introduced by Waltman & Van Eck (2012) and (Traag et al., 2019). We aggregated publications into 5 broad domains: Arts & Humanities (AH), Life Sciences & Biomedicine (LSB), Physical Sciences (PS), Social Sciences (SS), and Technology (Tech). We employed the fractional counting method to assign weights for each item to avoid inflationary effects, as academic papers may belong to more than one discipline, have multiple authors or institutions (Waltman, 2016; Waltman & van Eck, 2015). In total, our analysis comprises of 39,737,991 publications and 1,325,489,381 references.

In what follows, we provide basic formula (1) for calculating citation deflator (CDR) based on the growth of reference rate.

$$CDR_{a,t_b(t)} = \frac{\frac{\sum_{i=1}^{n} r_{a,t_b,i}}{\sum_{j=1}^{m} r_{a,t,j}}}{\frac{n}{m}} = \frac{\frac{n \cdot \overline{r_{a,t_b}}}{m \cdot \overline{r_{a,t}}}}{\frac{n}{m}} = \frac{\overline{r_{a,t_b}}}{\overline{r_{a,t}}} \tag{1}$$

where t_b is an arbitrary baseline year, which we set to t_b=2000 for the remainder of our analysis. Here a refers to the research area that is broadly associated with the publication, n is the number of papers published in baseline year t_b, m is the number of papers published in year t, $r_{a,t_b,i}$ is the number of references in paper i published in baseline year t_b in research area a, $r_{a,t,j}$ is the number of references in paper j published in year t in research area a, $\overline{r_{a,t_b}}$ is the average number of references in baseline year t_b in research area a, $\overline{r_{a,t}}$ is the average number of references in year t in research area a.

Our approach thus uses $CDR_{a,t_b(t)}$ to a rescaling of the "nominal citation rate" $\Delta NC_{p,t}$, thereby yielding a "real citation rate" $\Delta RC_{p,t}$ given by formula (2).

$$\Delta RC_{p,t} = \Delta NC_{p,t} \times CDR_{a,t_b(t)} \qquad (2)$$

Finally, the real citation count of paper p based on the benchmark year can be obtained by summing up all the real citation values of paper p in each citing year using formula (3).

$$RC_{p,t_b} = \sum_{t=t_b}^{t} \Delta RC_{p,t} \qquad (3)$$

In this example case, the target subject is a particular paper, although it could conceivably be any other unit of analysis, e.g. a journal, researcher, institution or country.

Using the data and formulas outlined above, we aim to investigate the overall trend of citation inflation and its impact on scientific evaluation. Through an analysis of citation patterns over time, we can identify the extent to which citation inflation has influenced the scientific community's perception of a paper's impact. Furthermore, we will examine the presence of citation inflation in different WC, research areas, and countries, and compare our results to those obtained by (Petersen et al., 2019). The insights gained from this research will be valuable in assessing current scientific evaluation methods and in guiding the development of more accurate and equitable citation evaluation indicators.

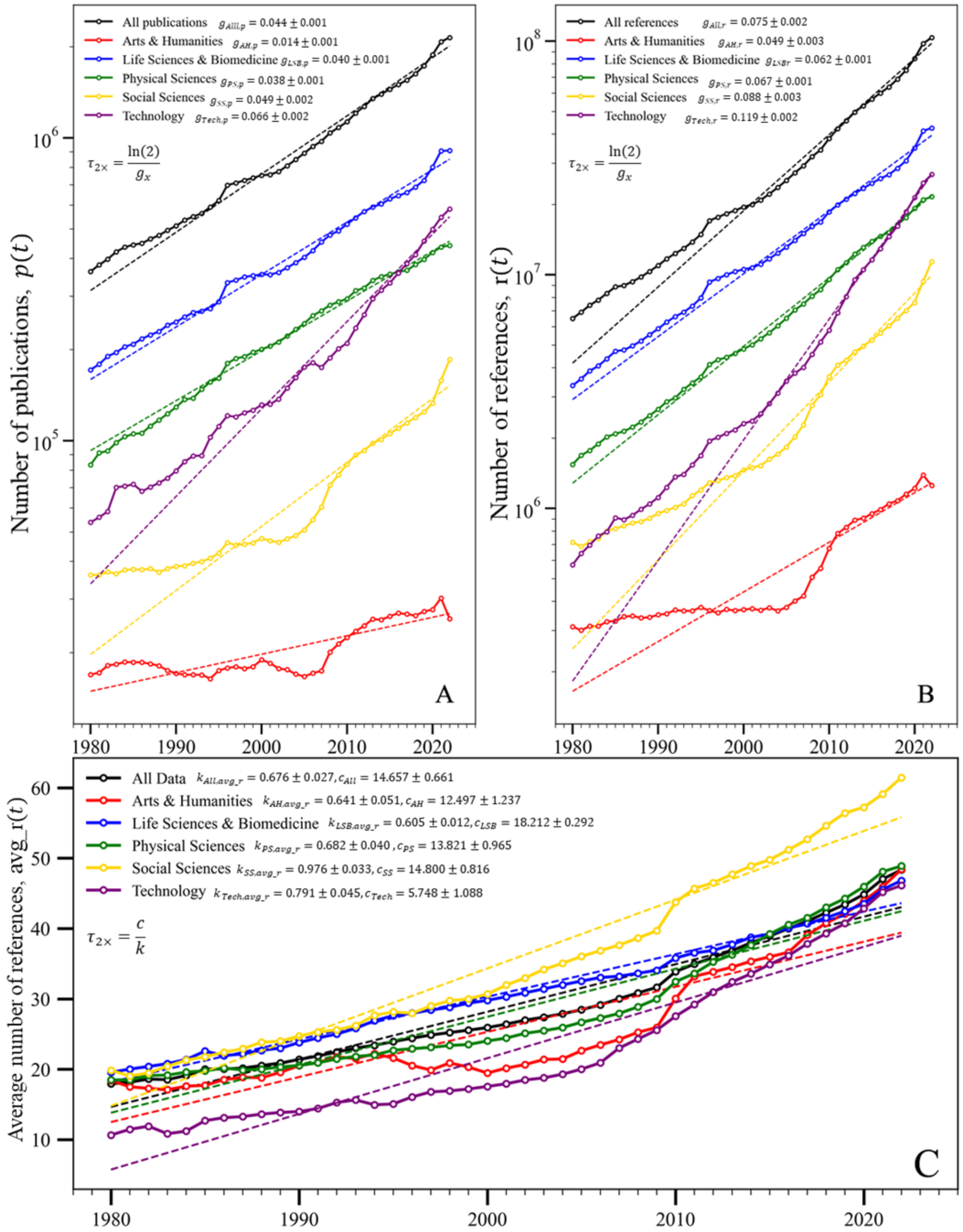
3. Results

3.1 Growth of scientific articles and references

Analyses were carried out to identify changes over time in properties related to publications, references and average references. g_x is the annual exponential growth rate associated with the generic time-dependent quantity $X(t)$. This growth rate is calculated by applying ordinary least squares regression to fit the exponential model: $\ln X(t) = \ln X(0) + g_x(t)$. For $g_x \ll 1$, then the 1-year growth in $X(t)$is roughly $100g_x$ percent. Quantities modeled in this way are the publication volume, $p_{(t)}$; the reference supply, $r_{(t)}$. The average reference supply, avg_$r_{(t)}$, was estimated by applying ordinary least squares regression to fit the linear model: $X(t) = k_x(t) * X(0) + c$. k_x is the annual linear growth rate associated with the generic time-dependent quantity. We set the value of sigma to 0.05 and the confidence level to 0.95 for parameter estimation.

In Fig 1, it is observed that all publications exhibit an annual growth rate of approximately 4.4%, whereas all references show an exponential growth rate of 7.5%. The average number of references follow an approximate linear distribution with a slope of 0.68. Fig 1 also displays growth at the research area aggregation level. Based on the growth rate of published articles, the field of Technology (6.6%) has shown the fastest growth, followed by the fields of Arts & Humanities (4.4%), while Life Sciences & Biomedicine (4%) and Physical Sciences (3.8%) show comparable growth rates. Similarly, in terms of references, Technology (11.9%,) has exhibited the most rapid growth, followed by the fields of Arts & Humanities (8.8%), with Life Sciences & Biomedicine (6.2%) and Physical Sciences (6.7%) showing roughly equivalent growth rates. However, Social Sciences (0.98) exhibit the highest growth rate in the average number of references, followed by Technology (0.79) and Physical Sciences (0.68). These results are consistent with the finding of (Petersen et al., 2019) that the total number of references produced each year is growing even faster than that of publications.

Figure 1. Growth of publications, references and average number of references



3.2 Growth of citation inflation

The method proposed for accounting for citation inflation is analogous to the method used for adjusting real prices for monetary inflation (Orphanides & Solow, 1990). Instead of using the publication rate as a citation deflator, we use the average number of references that is similar to a consumer price index as a price deflator, to convert "nominal" citation rates into "real" citation rates, so that they have common units and are comparable across time. This originates from a fundamental assumption that a cited literature carries value, and the magnitude of its value is related to the relative number of references.

The doubling time of all publications, references and average references is roughly 16 years,10 years and 22 years (Fig1). If we use the CDP that proposed by (Petersen et al., 2019), every 16 years the “real value” of a citation is halved in relative terms. While if we use the CDR that proposed in this paper, the “real value” of a citation is halved in 22 years. Compared to CDR, CDP overemphasizes the value of early citations, resulting in their over-amplification, and the value of recent citations is overly underestimated (Fig 2). This could be due to the exponential growth of publications, which increases rapidly, while the average reference per article grows linearly at a slower rate.

Figure 2. A comparison of two citation inflators based on all data

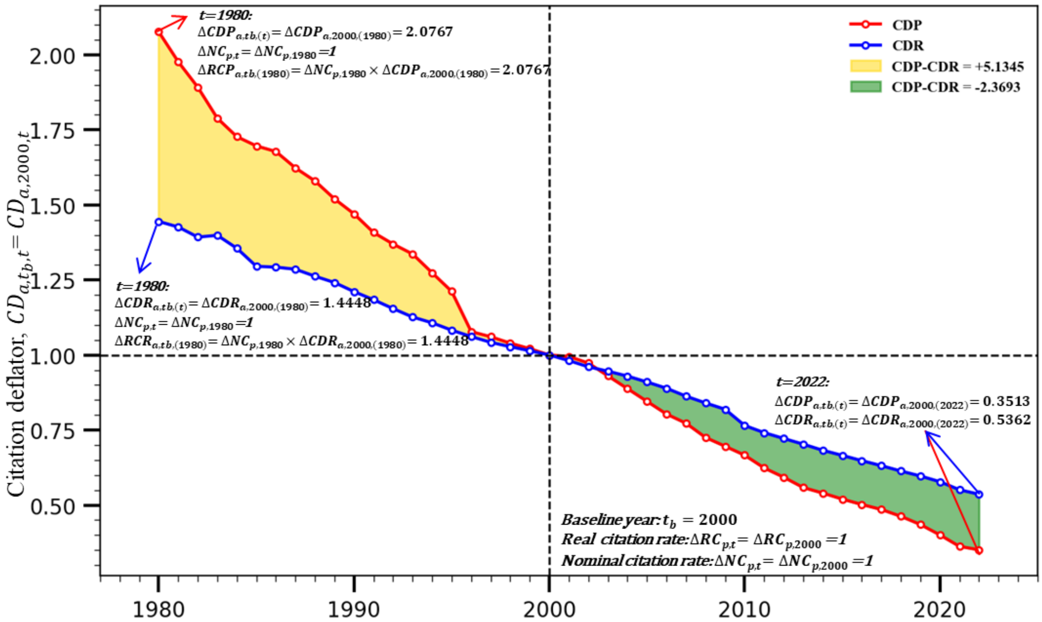
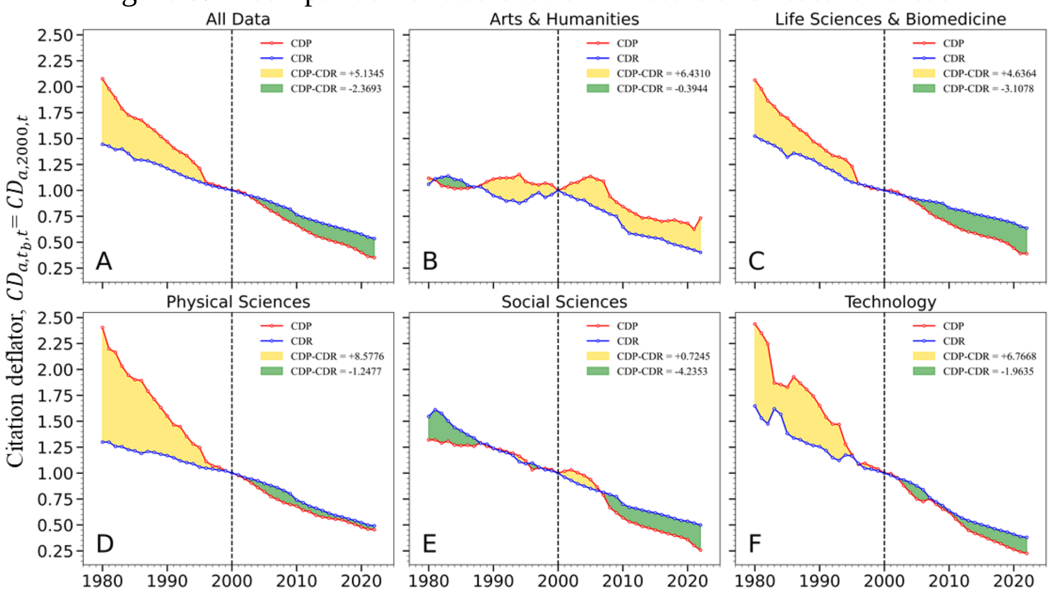


Figure 3. A comparison of two citation inflators at 5 research areas



In addition, we also investigated the differences between CDP and CDR in different research areas (Fig 3). The similarity between the two metrics is that they both indicate citation inflation, which means a decrease in the value of citations. Regarding the differences, similar to the overall trend, compared to CDR, CDP overestimates the early citation value while underestimating the recent citation value in the Life Sciences & Biomedicine, Physical

Sciences, and Technology. However, E and F fields show distinct deviations from the overall pattern. In the Arts & Humanities, the overestimation period is longer, while in the Social Sciences, the underestimation period is longer. This may be because the growth rate of publications in Arts & Humanities is significantly higher than the growth rate of citations per article, while in Social Sciences, it is the opposite.

Figure 4. A comparison of two citation inflators at top16 WC with the most publications

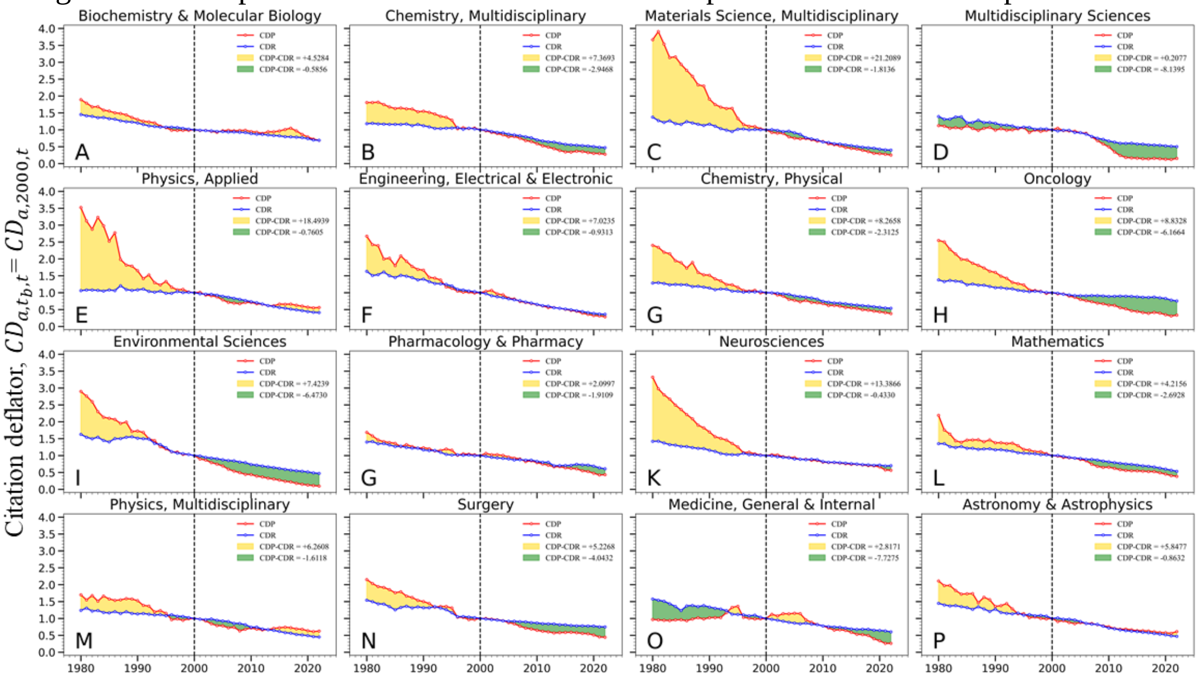
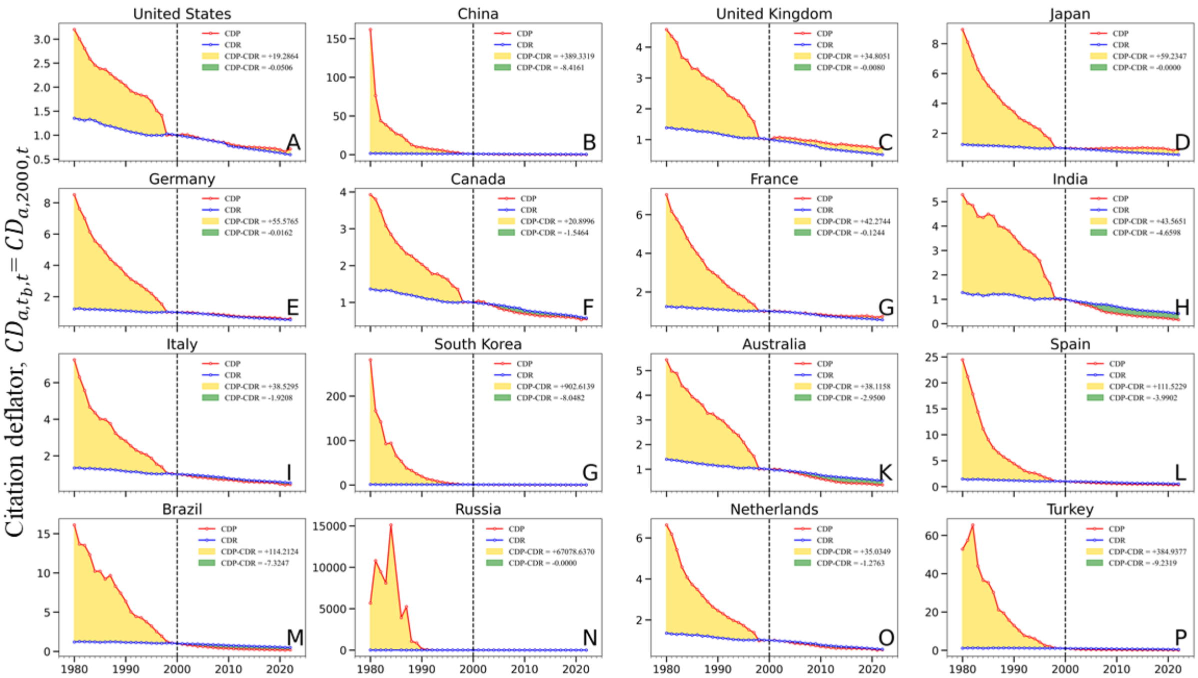


Figure 5. A comparison of two citation inflators at top16 country with the most publications



We analysed the differences of the two metrics at a more granular level, namely the differences on the level of WC disciplines. As shown in Fig 4, most disciplines exhibit trends similar to the overall pattern. When the growth rate of publications in a particular discipline is excessively high in the early stage, the value of early citations may be overly amplified, leading to the possible ineffectiveness of the CDP indicator, such as Materials Science & Multidisciplinary, Physics, Applied, Neurosciences. Furthermore, when analysing at the national level, CDP indicators may sometimes fail significantly (Fig5), such as Russia, China, South Korea. This may be due to some historical contexts, language systems, limitations in the coverage of WOS, and other factors that result in some countries having relatively few English-language publications in their early history. However, CDR measures the decay of citation value based on the average number of references per article, which is independent of the volume of publications. Therefore, it can be used to supplement the shortcomings of CDP. CDR can more accurately measure the relative value decay of citations, thereby improving the accuracy of citation metrics.

3.3 Inflation-corrected citation value and impact measures

The impact of scientific growth varies depending on the unit of analysis. In terms of publications, the growth leads to a significant decrease in the visibility of previous publication cohorts relative to recent ones due to the doubling period of roughly 2 decades(Pan et al., 2018). Meanwhile, the growth also affects the estimation of citation impact for unit of analysis when tallying citation counts over long periods. In such cases, the total citation impact strongly depends on whether nominal or deflated values are used, as shown in Fig6-8. NC represents the aggregate number of citations based on the nominal citation rate, while RC1 and RC2 correspond to the total citation counts based on CDP and CDR, respectively.

The difference between the traditional measures calculated using nominal citation rates, NC_i, and their deflated counterparts, RCP_i and RCR_i, reflects the magnitude of the measurement error due to citation inflation. The ratios calculated by formula(4) quantify this difference within each unit of analysis, with 100(ρ−1) corresponding to the percent difference relative to the nominal value .

$$\rho_{i,CDP} = {}^{NC_i}/_{RCP_i}, \rho_{i,CDR} = {}^{NC_i}/_{RCR_i} \ and \ \rho_{i,RCP,RCR} = {}^{RCP_i}/_{RCR_i} \quad (4)$$

Fig. 6 also serves as a visual guide in associating the differences in the nominal and real curves, and the summary difference quantified by ρ. Regardless of the indictors employed, it is evident that citation values are inflated, with CDP overestimating the degree of citation inflation. For example, CDP estimates a citation inflation rate of 120% for China's total citations, while CDR estimates a rate of 64%. CDP exaggerates the inflation rate by 35% compared to CDR.

We further analysed the impact of citation inflation on the longitudinal evaluation of journal. Nature and Science are two of the most prominent and influential scientific journals in the world, that publish articles on all areas of science and technology. Because Nature and Science belongs to the research area of Technology, we use two distinct deflators, $CDP_{Tech,2000(t)}$ and $CDR_{Tech,2000(t)}$, respectively. As shown in Fig7, the similarity in the annual citation growth rates between the two journals may have contributed to the comparable levels of citation inflation measured by the two indicators, with CDP slightly overestimating the degree of inflation compared to CDR. For these two journals, the citation inflation rate is estimated to be around 50%.

Finally, we investigated the impact of citation inflation on the longitudinal evaluation of publications. We selected the top three highly cited articles from both Nature (Iijima, 1991; O'Regan & Grätzel, 1991; Watts & Strogatz, 1998) and Science(Barabási & Albert, 1999; Kirkpatrick et al., 1983; Novoselov et al., 2004), and calculated their nominal and adjusted citation counts, as shown in Figure 8. Fig8 (A, B) shows that even when applying the same citation time window and field-specific inflation factor, the citation inflation rate of these two articles is still different because they receive different citation rates each year. Additionally, as shown in Fig8 (A, D), with the adjustment of the citation inflation factor (CDP or CDR), the ranking of articles based on actual total citation counts may differ from that based on nominal total citation counts. For example, paper D has a higher nominal total citation count than paper A, but paper D's actual total citation count is lower than that of paper A's. This simple example highlights how it is important to account for the different timing of citations – in addition to discipline(Waltman & van Eck, 2015) and fixed citation window(J. Wang, 2013).

Figure 6. The impact of citation inflation on the longitudinal evaluation of countries

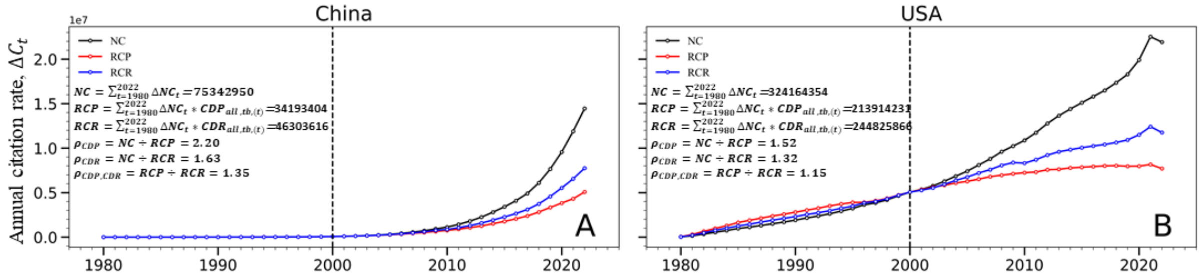


Figure 7. The impact of citation inflation on the longitudinal evaluation of journals

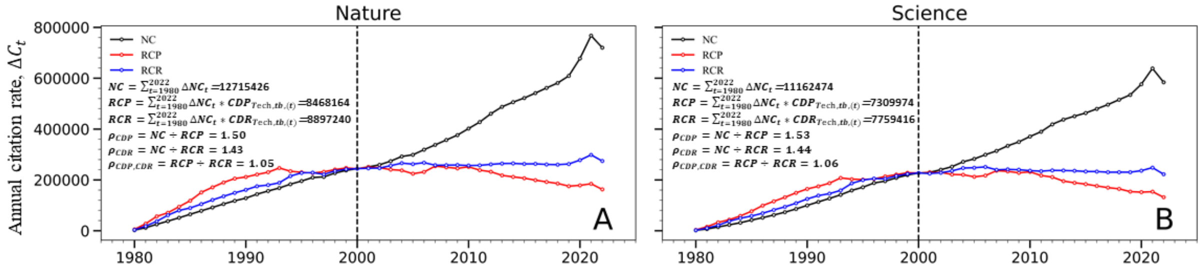
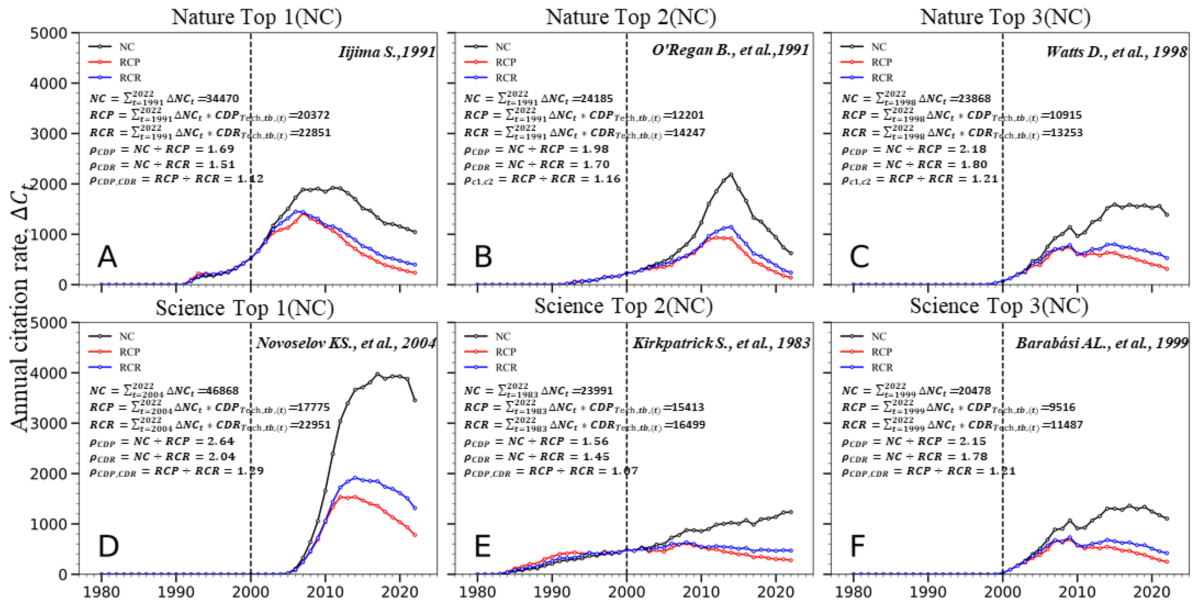


Figure 8. The impact of citation inflation on the longitudinal evaluation of publications



4. Conclusions

Citation inflation is a statistical bias that affects the quantitative evaluation of science, but it is often overlooked. The primary causes of citation inflation are the exponential increase in publications and reference list lengths, resulting in more citations being produced today than in the past. To address this issue, we developed a statistical method to deflate citations by factoring out this growth. The citation inflator we propose diverges from that presented by (Pan et al., 2018). in that it gauges the decay of citation value through the growth rate of references per article, rather than solely through the growth rate of publications. Our method accounts for the different real value of citations occurring at different times and is not susceptible to being invalidated by the rapid growth of publications in a particular field. However, (Pan et al., 2018) proposed a citation inflator based on the implicit assumption that the annual citation rate per article remains constant, which is not always the case in reality, especially for longer time spans As such, one drawback of our method is that it requires selecting a target research area a and calculating the average number of references for all publications in the research area a, which for researchers or institutions may be challenging to access meta data of all publications.

Our statistical method is necessary because citation rates and the real value of citations change over time. Our method shows that every 22 years, the "real value" of a citation is halved. Not converting nominal citation values into real citation values results in significant mis-measurement of scientific impact. The citation inflator addresses citation evaluation biases that fixed citation time windows and field normalization cannot resolve. The growth of the average number of references is significant because many careers and institutions span 22-year doubling periods, leading to measurement errors associated with citation inflation. Therefore, a citation deflator index is needed to precisely correct for inflationary temporal bias and could be easily incorporated into popular bibliometrics databases such as Web of Science, Scopus, and Google Scholar.

Acknowledgments
Gege Lin expresses her gratitude to CWTS for hosting her as a Guest Researcher, which provided her with access to the CWTS database and facilitated the completion of this research.

Author contributions
Gege Lin: Conceptualization; Formal Analysis; Methodology; Visualization; Writing – original draft
Guoqiang Liang: Conceptualization; Funding acquisition; Writing – review & editing
Haiyan Hou, Zhigang Hu: Funding acquisition; Supervision; Writing – review & editing

Competing interests
The authors declare that they have no competing interests.

Funding information
This work was supported by the National Natural Science Foundation of China (71974030 & 72204014) and Beijing Natural Science Foundation (9232002).

References

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science, 286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Burton, R. E., & Kebler, R. W. (1960). The “half-life” of some scientific and technical literatures. *American Documentation, 11*(1), 18–22. https://doi.org/10.1002/asi.5090110105

Dangzhi Zhao, Andreas Strotmann, & Andreas Strotmann. (2020). Deep and narrow impact: Introducing location filtered citation counting. *Scientometrics*. https://doi.org/10.1007/s11192-019-03280-z

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., & Uzzi, B. (2018). Science of science. *Science, 359*(6379), eaao0185.

Huang, S., Qian, J., Huang, Y., Lu, W., Bu, Y., Yang, J., & Cheng, Q. (2022). Disclosing the relationship between citation structure and future impact of a publication. *Journal of the Association for Information Science and Technology*, *73*(7), 1025–1042. https://doi.org/10.1002/asi.24610

Iijima, S. (1991). Helical microtubules of graphitic carbon. *Nature*, *354*(6348), 56–58. https://doi.org/10.1038/354056a0

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, *220*(4598), 671–680. https://doi.org/10.1126/science.220.4598.671

Marks, R. J. (2022, December 7). *How the Inflation of Journal Citations Impacts Academia*. Mind Matters. https://mindmatters.ai/2022/12/how-the-inflation-of-journal-citations-impacts-academia/

Miura, T., Asatani, K., & Sakata, I. (2021). Large-scale analysis of delayed recognition using sleeping beauty and the prince. *Applied Network Science*, *6*(1), 48.

Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V., & Firsov, A. A. (2004). Electric Field Effect in Atomically Thin Carbon Films. *Science*, *306*(5696), 666–669. https://doi.org/10.1126/science.1102896

O’Regan, B., & Grätzel, M. (1991). A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO2 films. *Nature*, *353*(6346), 737–740. https://doi.org/10.1038/353737a0

Orphanides, A., & Solow, R. M. (1990). Money, inflation and growth. *Handbook of Monetary Economics*, *1*, 223–261.

Pan, R. K., Petersen, A. M., Pammolli, F., & Fortunato, S. (2018). The memory of science: Inflation, myopia, and the knowledge network. *Journal of Informetrics*, *12*(3), 656–678. https://doi.org/10.1016/j.joi.2018.06.005

Petersen, A. M., Fortunato, S., Pan, R. K., Kaski, K., Penner, O., Rungi, A., Riccaboni, M., Stanley, H. E., & Pammolli, F. (2014). Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences*, *111*(43), 15316–15321.

Petersen, A. M., Pan, R. K., Pammolli, F., & Fortunato, S. (2019). Methods to account for citation inflation in research evaluation. *Research Policy*, *48*(7), 1855–1865. https://doi.org/10.1016/j.respol.2019.04.009

Teixeira da Silva, J. A., & Dobránszki, J. (2018). Citation inflation: The effect of not correcting the scientific literature sufficiently, a case study in the plant sciences. *Scientometrics*, *116*(2), 1213–1222. https://doi.org/10.1007/s11192-018-2759-x

Thelwall, M. (2019). Are classic references cited first? An analysis of citation order within article sections. *Scientometrics*, *120*(2), 723–731.

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-41695-z

Van Raan, A. F. (2004). Sleeping beauties in science. *Scientometrics*, *59*(3), 467–472.

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007

Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*(12), 2378–2392.

Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, *9*(4), 872–894. https://doi.org/10.1016/j.joi.2015.08.001

Wang, D., Song, C., & Barabási, A.-L. (2013). Quantifying long-term scientific impact. *Science*, *342*(6154), 127–132.

Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, *94*(3), 851–872.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Yin, Y., & Wang, D. (2017). The time dimension of science: Connecting the past to the future. *Journal of Informetrics*, *11*(2), 608–621.