ISSI 2023 Proceedings, v1, pp 685–700
DOI: 10.5281/zenodo.8253488 https://cns-iu.github.io/workshops/2023-07-02_issi/

The global disruption index (GDI): An incorporation of citation cascades in the disruptive index

Alex Jie Yang[1], Yuhao Wang[2], Jia Kong[3], Qi Zhang[4], Haotian Hu[5], Hao Wang[6], Sanhong Deng[7]

[1] *alexjieyang@outlook.com,* [2] *1668724679@qq.com,* [3] *kongjia@smail.nju.edu.cn,* [4] *qi.zhang@smail.nju.edu.cn,* [5] *hhtdlam@126.com,* [6] *ywhaowang@nju.edu.cn,* [7] *sanhong@nju.edu.cn*
School of Information Management, Nanjing University, 22 Hankou Rd., Nanjing (China), 210093

Abstract
The disruption index (DI) proposed by Funk and Owen Smith (2017), and improved by Wu et al. (2019) has been widely studied and applied in bibliometrics recently. This paper examines the relationship between disruptiveness and influence and finds that highly influential publications tend to be more disruptive. However, the original formula of the disruption index (DI) shows inaccurate results in our sample data. We identify three limitations of DI and propose the global disruption index (GDI), which considers multiple generations of cited or citing documents in citation cascades, to address these limitations. Our empirical study, based on a dataset of 134,003 papers published in the field of Information Science Library Science from 1900 to 2021, and all articles published in *Scientometrics*, shows that GDI has a significant correlation with the citation count and the number of disruptive citations. We argue that GDI is a more appropriate metric for measuring the disruptiveness of scientific publications and has great significance for academic evaluation and scientific development in the future. Furthermore, our study suggests that the number of consolidated citations also facilitates publications' disruptiveness. These findings have important implications for understanding the nature of scientific impact and for developing more accurate and comprehensive bibliometric measures.

Introduction
Ever since Garfield (1963) proposed the impact factor (IF), citation-based evaluation metrics have become an essential part of scientific research. At a time of tight budgets and intense competition, evaluation metrics that rely on citation counts play a critical role in university ranking, peer reviews, and journal evaluations. Nevertheless, the overemphasis on citation counts has led to a host of problems, including academic misconduct and a lack of research innovation (An et al., 2022; Fortunato et al., 2018; Waltman, 2016; Wang & Barabási, 2021; Zeng et al., 2017). Furthermore, citation-based indicators are insufficient in measuring the novelty of scientific research, which means that appropriate metrics for evaluating innovation in research have been lacking for a long time. Many scholars have therefore criticized the traditional evaluation indicators and called for new metrics for assessing the level of innovation in publications. This call for new metrics is reflected in *the Leiden Manifesto* (Hicks et al., 2015) and *the San Francisco Declaration* (Cagan, 2013).

Joseph Alois Schumpeter (1912) first introduced the fundamental concept and idea of innovation, which involves the establishment of production functions. Thomas Kuhn (1962) argued that scientific development heavily relies on the destruction of existing paradigms, which aligns with the idea of disruptive innovation (Bower, 1995). Scientific progress often stems from anomalies that cannot be explained by existing paradigms, and breakthroughs come from novel ideas or methods that solve these anomalies. The disruptive index (DI) introduced by Funk and Owen Smith (2017) is a new metric that uses local citation structures to reflect the novelty of research and the extent to which it is subversively innovative, meaning that it changes or significantly affects the research paradigm of science. Wu et al. (2019) applied DI to bibliometrics and found that milestone papers, such as Nobel-Prize-winning publications, tend to have higher DI, while inheritance research typically receives relatively low DI scores. According to Leydesdorff and Bornmann (2021), the key distinction between DI and other

bibliometric measures is that it examines the local structure of the citation network, including the combination of prior and posterior generations and takes into account the evolutionary dynamics and crucial transitions of scientific publications.

Nevertheless, recent studies have revealed the inconsistencies and limitations of the DI metric (Bornmann et al., 2020a, 2020b; Leydesdorff & Bornmann, 2021; Leydesdorff et al., 2021). Chen (2021) suggested that scientific research and technologies can be evaluated along two dimensions: disruptiveness and consolidation, which are also separated by the multidimensional citation impact framework proposed by Bu et al. (2021). As a result, a single indicator that combines both dimensions may be inappropriate (Leydesdorff et al., 2021). Wang et al. (2022) discovered that many scientific works with high DI values were not ground-breaking research. DI is determined solely by the quantity of the first-order citing documents and first-order references, and the frequency with which those references have been cited, so not all publications with a DI of 1 are ground-breaking achievements (Leydesdorff & Bornmann, 2021). Li et al. (2022) suggested that metrics in the first-order citation networks are inadequate for characterizing the multivariate sequential interactions between documents. Ullah et al. (2021) demonstrated that higher-order models perform better at ranking scientific entities than one-step or zero-step models. Wang et al. (2022) argued that disruptively innovative research is generally heterogeneous in practice. However, the assessment of heterogeneous scientific works may be imprecise and biased if based only on local citation information (Prathap, 2019).

Through our empirical investigation, we have found that highly prominent articles tend to be more disruptive but may not be consolidated. However, we discovered that the original DI formula yields incorrect results, as we describe in Section 4. Upon further inspection, we identified three limitations of the DI formula, which we discuss in Section 2. To address these issues, we combined the notion of the citation cascade (Min et al., 2021) to create the global disruptive index (GDI), which incorporates high-orders of citation information. We then conducted an empirical study to evaluate the consistency and effectiveness of GDI, and found that it has a significant correlation with the citation count and the number of disruptive citations. We also examined whether the number of consolidated citations affects the disruptive innovation of articles, and our results suggest that consolidated citations may have a slight positive effect on publications' disruptiveness, as indicated by GDI.

Methodology

The disruptive index is defined as (Wu et al., 2019):

$$DI = \frac{N_G - N_R}{N_G + N_R + N_E} \quad (1)$$

where N_R, N_G and N_E denote the count of publications citing FP and its references, citing FP without citing its references, and citing FP's references without citing FP, respectively.

Figure 1 illustrates the disruptive index (DI), with the number of squares representing the number of disruptive citations and the number of pentagons indicating the number of consolidated citations. DI values range from -1 to 1, with a DI of 1 indicating that none of the focal paper's cited publications are connected to its citing references, and a DI of -1 indicating the opposite. Higher DI values indicate greater disruptiveness of the paper (Azoulay, 2019), meaning that it challenges or upsets existing paradigms in citation networks. However, it's important to note that DI only considers the first citation generation for both citing references and cited documents, which makes it difficult to account for the heterogeneity of articles.

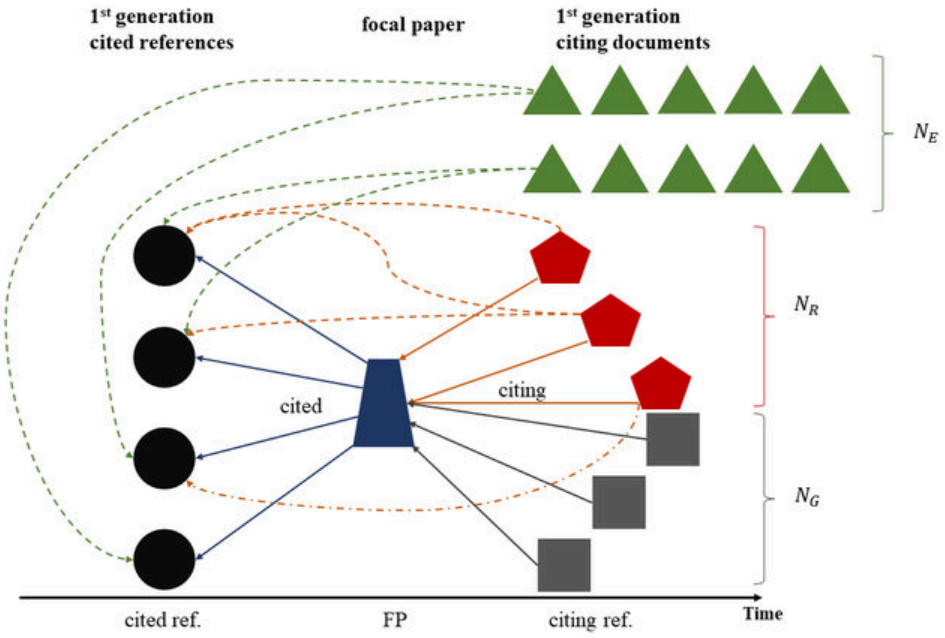


Figure 1. The visualization of the disruptive index of the focal publication in citation cascade, which in only the local citations (the first generation) are considered. The number of triangles, pentagons and squares represent N_E, N_R and N_G, respectively.

It is important to note that the citation number of the focal paper (FP) equals the sum of its citing publications (N_G) and its cited documents (N_R). DI divides the FP's citation publications into two types, representing the levels of disruptiveness and consolidation, respectively. However, the numerator of the original DI formula, N_G-N_R, is problematic, as it subtracts the consolidation component from the total citation count twice. To address this issue, Leydesdorff et al. (2021) revised the formula as follows:

$$DI^* = \frac{N_G}{N_G+N_R+N_E} \quad (2)$$

where DI* indicates the level of disruptiveness of the focal publication.

Another issue with the original DI formula is the value of N_E, which appears to be incorrect. Empirical results show that N_E is, on average, larger than its counterparts by an order of magnitude (see Section 4, Table 1, and Figure 3). This phenomenon occurs because all publications in N_G and N_R are connected to a single paper, the focal paper, while those in N_E are connected to sometimes a substantial number of papers. As a result, DI is argued to be inconsistent and uneven in many cases (Leydesdorff & Bornmann, 2021).

Here, we identify three limitations/issues of DI, that is,

i. DI does not take into account high-order citations in multi-generations of citation cascades.
ii. The numerator (N_G-N_R) is inappropriate for duplicate subtractions, since N_G itself can already denote the level of disruptiveness.
iii. N_E is typically much larger than its counterparts in the formula due to nodes in N_E connecting to more than one node in the citation network. Therefore, standardization is needed to address this issue.

The citation cascade, a new notion introduced by Min et al. (2021), involves small-scale citation networks that originate from a focal publication. As the focal paper is cited by other papers, these papers are, in turn, cited by subsequent papers, and so on. Hu et al. (2011) suggested that

the citation cascade provides an effective means of assessing a publication's influence from the perspective of indirect influence, by considering multiple levels of citation generations. Metrics based on global information in the citation cascades tend to outperform local metrics (Xu et al., 2020).

To address the aforementioned three issues, we propose the Global Disruption Index (GDI) which takes into consideration the general structure of the citation network, including k generations. Similar to DI* (Leydesdorff et al., 2021), we use N_G to denote disruptiveness. Instead of using the original N_E, we use a normalized value $nomlizedN_E$ which is calculated by taking the average of the number of FP's references and applying the Laplacian smoothing method (see formula 4). Therefore, GDI is defined as follows:

$$GDI(Global\ disruption\ index) = \frac{\sum_{i=1}^{k} N_G^i + k}{\sum_{i=1}^{k} N_R^i + \sum_{i=1}^{k} N_G^i + nomlizedN_E + (2k+1)} \quad (3)$$

$$nomlizedN_E = \frac{\sum_{i=1}^{k} N_E^i + k}{\sum_{i=1}^{k} N_C^i + k} \quad (4)$$

where k denotes k generations of cited or citing documents, i refers to the i^{th} generation of cited documents, N_R^i represents the number of cited documents in the i^{th} generation that cite at least one of the focal paper's references directly or indirectly within i generations, N_G^i refers to the number of cited documents in the i^{th} generation that cite none of the focal paper's references directly or indirectly within i generations, and N_E represents the cited documents that not included in N_G^i or N_R^i of the focal paper's citing references within 2 generations.

Additionally, Laplace smoothing is used to handle missing values, by adding a smoothing factor (k) to the numerator and a larger factor (k+1) to the denominator. This approach is commonly used in probability and statistics to prevent zero probabilities and improve the stability of calculations, especially when dealing with sparse data. In our study, Laplace smoothing helps to account for cases where a cited or citing document has no citations or is missing from our data, allowing for more accurate computation of GDI.

Min et al. (2021) proposed two criteria for citation cascades: a paper can appear in multiple generations and can also appear multiple times within the same generation. We adopt both criteria in our definition of citation cascades.

Figure 2 illustrates the citation cascade of the focal paper, including the different cited documents within two generations. In the first generation, we have papers that directly cite the focal paper and are in turn cited by the second-generation documents. Analogously, in the first citing-generation references, we have papers that are cited by the focal paper directly and cite the second-generation cited references. Both citing-generations of references are denoted by round balls in the left part of the figure, while cited documents are denoted by squares or pentagons in the right part, depending on whether they cite at least one of the focal paper's references directly or indirectly. The triangles represent the cited documents of the round balls, which are not part of the focal paper's citing or cited generations. As discussed above, we divide the number of citing references within two generations into their citation counts for the calculation of GDI. Although Figure 2 shows only two generations of cited or citing documents, extending it to k generations is straightforward.

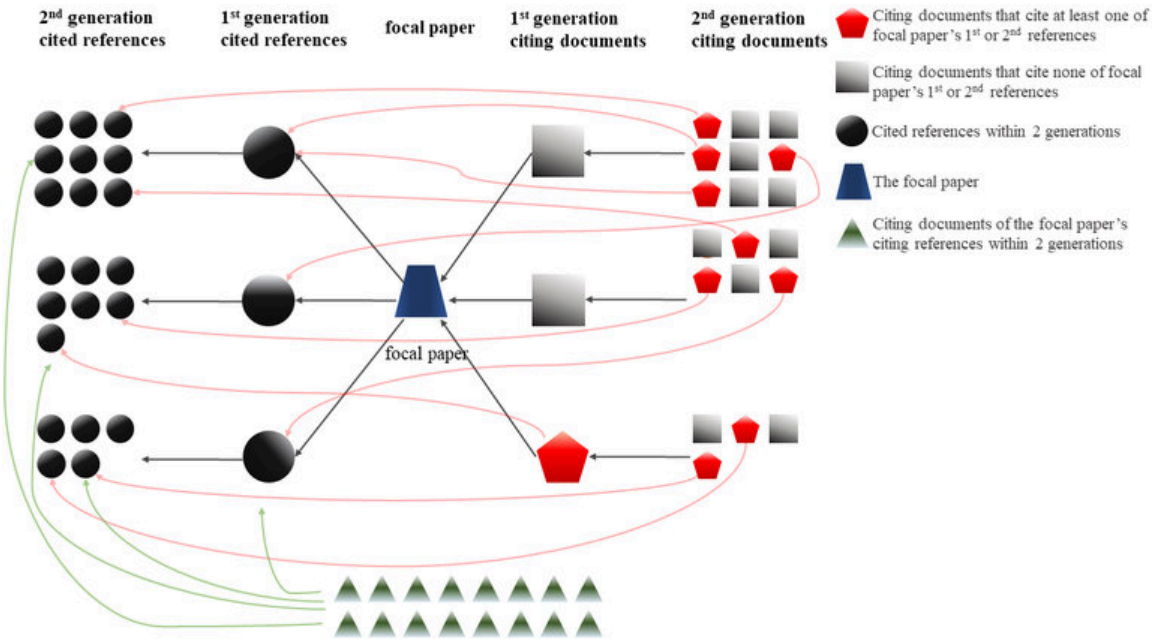


Figure 2. The focal paper's citation cascade and cited documents within 2 generations (k=2). The antiparallelogram in the middle is the focal paper. The round balls in the left represent the cited references within 2 generations. The pentagons in the right refer to the citing documents within 2 generations that cite at least one of focal paper's 1st or 2nd references. The squares in the left denote citing documents within 2 generations that cite none of focal paper's 1st or 2nd references.

Data and methods

We retrieved our data from the Web of Science database using the following sampling procedure: (i) We downloaded bibliographic data for 134,003 papers published in the field of Information Science Library Science (WoS categories) from 1900 to 2021. (ii) We then searched for all papers published in *Scientometrics* and restricted the document type to articles and the field to Information Science Library Science (WoS categories). This resulted in a sample of 6,330 articles for our analysis.

Calculating the GDI requires a significant amount of computation. To obtain reasonably reliable metric estimates, we created a local network database using bibliographic data for 134,003 papers published in the field of information science and library science from 1900 to 2021. We also constructed a large adjacency matrix for computation. While the actual citation counts in practice are much higher, our local database only contains papers in the field of information science. Therefore, during analysis, we used the in-degree to denote citation counts and the out-degree to denote reference counts.

To calculate the relevant measures, we used Python 3.9.10 and implemented programs specifically tailored to our needs. To ensure accuracy, we performed duplication elimination processing at each stage. We also created a recursive function and set a maximum of k iterations to calculate the number of k-step posterior citations. This function allowed us to calculate a maximum of k generations of referenced publications to the main paper. In addition, we constructed a local network database based on 134,003 papers published in the field of information science and library science from 1900 to 2021 to provide reasonably reliable estimates of metrics. We also created a sizable adjacency matrix for computation. Although in practice the actual citation counts are substantially bigger because our local database only contains papers in the field of information science, the in-degree denotes citation counts and

the out-degree denotes reference counts during analysis. Lastly, we used STATA-15.1 to conduct Pearson correlation analysis.

To obtain a comprehensive understanding of our sample data and mitigate the impact of outliers, we excluded the top 0.1% most highly cited papers and data with a citation count or NG of 0. After this process, we were left with 3,375 papers for further analysis. The citation count of these papers is represented by the out-degree of nodes in the global citation network, which was constructed from 134,003 papers in the field of Information Science Library Science published from 1900 to 2021.

Results and discussion

Disruptiveness and consolidation level of papers

Table 1 summarizes the key descriptive statistics for the distinct metrics in our sample of 3,375 articles, including the sample size, mean, standard deviation, the minimum value, the maximum value, variance, skewness and kurtosis of distinct metrics. To reduce the impact of outliers, we excluded the top 0.1% most cited papers and articles with zero citation count or no generations. The mean citation count of the remaining articles is 12, with a range from 2 to 206. The mean values of N_G and N_R are 7 and 5, respectively, with a maximum of 176 and 99. In contrast, the average value of N_E is 260, with a maximum of 4,723, which is an order of magnitude larger than N_G and N_R. The second-order metrics, $\sum N_R$, $\sum N_G$,, and $\sum N_E$, show the same trend, with $\sum N_E$ being much larger than the other two metrics. However, the normalized N_E metric is at the same order of magnitude as $\sum N_R$ and $\sum N_G$, making it a more suitable denominator for the GDI calculation. As discussed in Section 2, this is because nodes in NE are connected to more than one node in the network, while nodes in N_G or N_R only connect to the focal one. Figure 3 illustrates the distributions of N_G, N_R, N_E, $\sum N_G$, $\sum N_R$, $\sum N_E$, and *normalized* N_E.

Table 1 and Figure 3 show that N_G and N_R have similar distributions, with comparable standard errors (12 vs. 8), skewness (6 vs. 4), and kurtosis (50 vs. 30). The logarithmic scatter plot in Figure 3a and Figure 3d indicates that N_G and N_R increase logarithmically with citation count, which is expected since N_G+N_R equals the citation count. On the other hand, $\sum N_G$ and $\sum N_R$ have weak differences in distributions, with inconspicuous proximity in standard errors (73 vs. 26), skewness (5 vs. 4), and kurtosis (33 vs. 26). Notably, some highly cited papers have low values of $\sum N_R$, but high $\sum N_G$ values (see Figure 3d and Figure 3e), and some well-cited publications have inferior N_R values but high N_G values (see Figure 3a and Figure 3b). If we consider N_G or $\sum N_G$ as the disruptiveness level and N_R or $\sum N_R$ as the consolidation level of publications, these findings suggest that highly influential papers tend to have strong disruptiveness levels but may not necessarily have strong consolidation levels.

It is also noteworthy that the distributions of high-order parameters become more uniform, as shown in the rightmost subfigures of Figure 3d and 3e. As the generation increases, the numerical distribution becomes more discrete and uniform, which makes it easier to compare scientific publications with different citation histories.

Table 1. The descriptive statistics for the citation count and distinct metrics (DI, DI* and GDI) and the corresponding coefficients (NG, NR, NE et al.) of the sample papers.

	sample size	*mean*	*std*	*min*	*max*	*var*	*skewness*	*kurtosis*
Citation count	3375	12	17	2	206	301	5	34
N_G	3375	7	12	1	176	149	6	50
$\sum N_G(k=2)$	3375	73	147	0	1727	2.2e+4	5	33
N_R	3375	5	8	0	99	66	4	30
$\sum N_R(k=2)$	3375	26	47	0	681	2231	4	26
N_E	3375	260	292	0	4723	8.5e+4	3	26
$\sum N_E(k=2)$	3375	1404	1339	0	1.2e+4	1.8e+6	1	4
normalized $N_E(k=2)$	3375	24	15	1	140	217	2	7
DI	3375	0.12	0.29	-0.38	1.00	0.08	2.43	4.52
*DI**	3375	0.15	0.28	4.1 e-4	1.00	0.08	2.38	4.29
GDI(k=2)	3375	0.41	0.26	1.5 e-2	1.00	0.07	0.43	-0.92

Note: The value is reported as an integer or with one to two decimal places. For some large values, scientific notation is used, e.g., 1.2e+4 represents approximately 12,000.

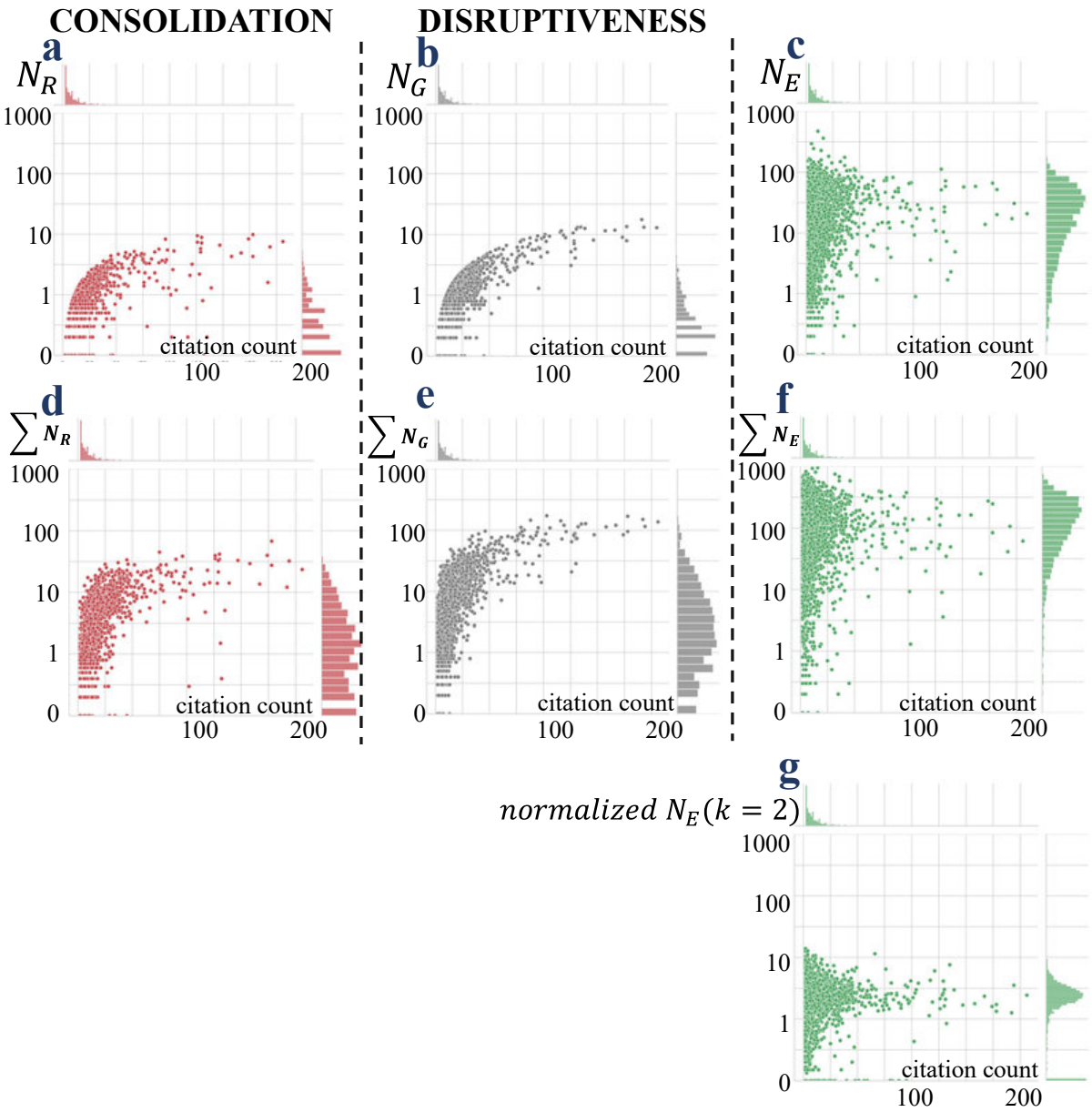


Figure 3. Scatter plots and distributions for the citation count and each parameter for papers published in *Scientometrics*, respectively, with logarithmic processing of vertical data. (a) The distributions of (a) N_R, (b) N_G, (c) N_E, (d) $\sum N_R$, (e) $\sum N_G$, (f) $\sum N_E$, and (g) *normalized* N_E, wherein N_R, N_G, N_E for 1-order DI parameters, and $\sum N_R$, $\sum N_G$, $\sum N_E$ and *normalized* N_E for 2-orders GDI parameters, respectively. Additionally, (a) N_R and (d) $\sum N_R$ denote the consolidation level while (b) N_G and (e) $\sum N_G$ represent the disruptiveness level, respectively.

Comparisons of DI and GDI

To demonstrate the consistency and effectiveness of GDI, as well as its superiority over DI and DI*, we calculated the Pearson correlation coefficients between these three metrics and the citation count and $\sum N_G$, which represents the disruptive influence of publications to some extent. Figure 4 shows the distributions and Pearson correlations for each metric with the citation count, while Appendix Figure A1 displays the distributions and Pearson correlations for each metric with $\sum N_G$.

The results indicate that DI has no significant correlation with the citation count (Pearson correlation coefficient=0.0059, p=0.73), implying that the influence of publications with high

DI or low DI is randomly distributed among highly-cited and normally-cited papers. This contradicts the observation that well-cited publications tend to be more disruptive. We suggest that the three issues discussed in Section 2 can explain this discrepancy. Moreover, DI has a very low correlation with $\sum N_G$ (Pearson correlation coefficient=0.1266), suggesting that DI fails to measure the real level of disruptiveness under certain circumstances. This can be attributed to the lack of normalization of N_E, which makes the denominator too large, mainly affected by the number of FP's references, thus diminishing the correlation between DI and the real disruptive influence. Additionally, the repeated subtraction of N_R and the lack of consideration of high-order generations introduce random disturbance factors in DI.

DI* proposed by Leydesdorff et al. (2021) addressed the issue ii (see Section 2), but didn't take into account the issue i and the issue iii. Results in Figure 4b and Appendix Figure A1e show that DI* has a comparatively significant yet weak correlation with the citation count (Pearson correlation coefficient=0.0548, $p<0.01$). And the correlation between DI* and $\sum N_G$ is also low (the Pearson correlation coefficient=0.1673). This finding indicates that well-cited publications are, faintly, more likely to get higher DI* values. Nevertheless, the correlation coefficient is too small to perceive and compare differences between well-cited publications and normally-cited ones.

The current paper proposes GDI as a metric that addresses all three limitations of DI. Although GDI considers only two generations, it is consistent with the citation count. Figure 4c and Appendix Figure A1f show that GDI has a high and significant correlation with the citation count (Pearson correlation coefficient = 0.3583, $p<0.001$) as well as $\sum N_G$ (Pearson correlation coefficient = 0.5233, $p<0.001$). These results indicate that well-cited publications tend to have higher GDI values, while normally-cited or lowly-cited papers generally have lower GDI values. This finding is consistent with the observation that well-cited publications tend to be more disruptive.

It's worth noting that the traditional disruption indices, both DI and DI* have a negative correlation with N_R and $\sum \boldsymbol{N_R}$, while GDI a positive correlation with these parameters (see Appendix Figure A1g-A1l). This raises an interesting question of whether the number of consolidated citations impacts the disruptive innovation of articles. DI and DI* suggest that a high number of consolidated citations inhibits disruptive innovation, whereas GDI suggests the opposite, indicating that a high number of consolidated citations can actually facilitate publications' disruptiveness.

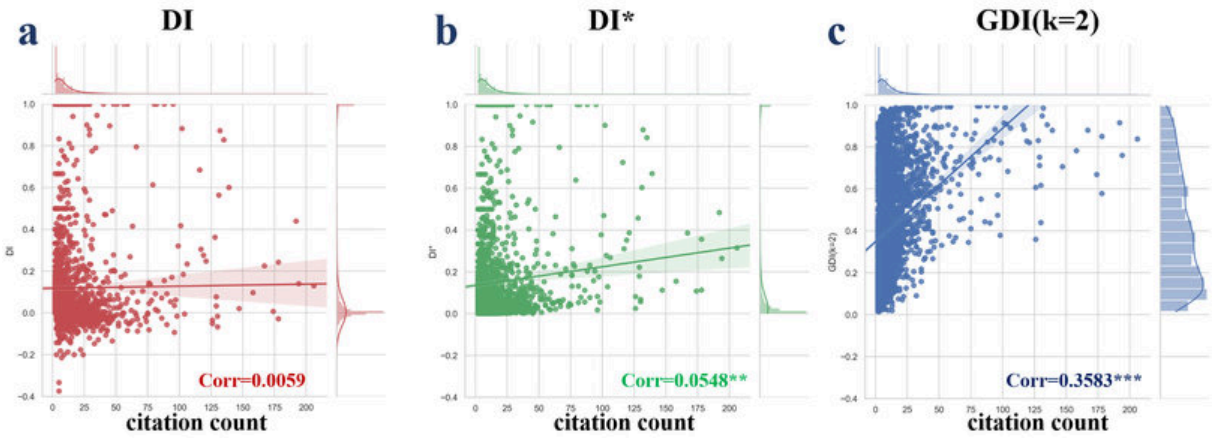


Figure 4. Scatter plots, fitting curves, distributions, and the Pearson correlations for the citation count and each metric of papers published in *Scientometrics*, respectively, wherein (a) the correlation for DI and the citation count is 0.0059 (p=0.73), (b) the correlation for the DI* and the citation count is 0.0548 (p<0.01); and (c) the correlation for GDI(k=2) and the citation count is 0.3583 (p<0.001). And these metrics' distribution histograms are illustrated in the subgraphs on the edges.

*** p<0.05, ** p<0.01, *** p<0.001.**

Disruptiveness for different citation windows

Bornmann et al. (2019) conducted a study to determine the appropriate citation window for DI, and suggested a minimum window of 5 years. In our analysis, we divided the sample data into four groups based on publication year to investigate the disruptiveness of different time windows. The distributions of GDI (k=2) for different publication years are shown in Figure 5, while the distributions of DI and DI* for different publication years are shown in Appendix Figure A2 and Figure A3, respectively.

The results show that papers published in the first 11 years (1978-1989) are the most likely to be disruptive, followed by the next 11-year interval (1989-2000). However, as the number of papers in journals has increased since 2000, the average disruptiveness has declined somewhat in the period of 2000-2011, and in the last 11 years (2011-2022), papers are less likely to be disruptive. From the perspective of citation length and distribution, DI and DI* values are evenly distributed on both sides of the average value on the exponential axis, with earlier papers (with longer citation windows) having significantly larger average values. In other words, in general, DI or DI* values increase exponentially with the length of the citation window. In contrast, GDI (k=2) distributions for publications with different citation windows show entirely different patterns, with publications with longer citation windows being much more likely to be disruptive.

Our study reveals a novel finding that older publications tend to be more disruptive, which is consistent with the preferential attachment model (Barabasi & Albert, 1999) and first-mover advantage. According to the preferential attachment model, newly published papers tend to cite the "first" paper in the field, often referred to as the classic or milestone paper. As a result, these classic papers may accrue a growing number of citations over time, which in turn increases their GDI scores. The tendency to cite classic papers and neglect recent references could explain why older publications tend to be more disruptive.

Overall, early-published papers tend to be more disruptive as they are more likely to challenge existing scientific paradigms. However, this conclusion is not absolute, as some recent papers (published in 2011-2022) also exhibit high GDI values.

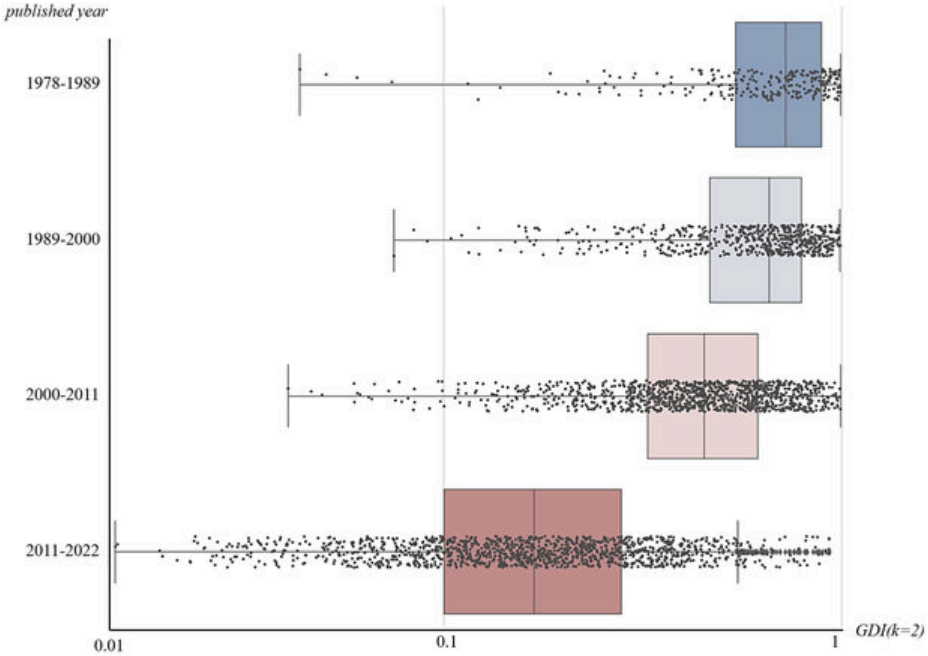


Figure 5. Box diagram for the distribution of GDI(k=2) for different published years, for papers published in 1978-1989, 1989-2000, 2000-2011 and 2011-2022, respectively, with logarithmic processing of horizontal data. Colored regions represent 95% confidence intervals.

Conclusion

The current study investigated the relationship between disruptiveness and influence and found that highly influential papers tend to be more disruptive, yet may not be consolidated, while the original formula of DI produced incorrect results (see Section 4). We identified three limitations of DI that account for this phenomenon (see Section 2) and introduced the global disruptive index (GDI) to address these issues. GDI considers high-order citation generations and incorporates information on heterogeneity in each generation for both cited and citing documents. We also employed normalization and Laplace smoothing to adjust relevant parameters. Our empirical study of 134,003 papers published from 1900 to 2021 in the field of Information Science Library Science demonstrated the consistency of GDI and tested its performance across different citation windows. The results showed a significant correlation between GDI and the citation count and the number of disruptive citations, and revealed that early-published papers with longer citation windows tend to be more disruptive. It is also worth discussing whether the number of consolidated citations affects the disruptive innovation of articles, and GDI suggests that it weakly facilitates publications' disruptiveness. Based on these findings, we argue that GDI can address the limitations of DI and accurately represent the level of scientific output's disruptiveness. We believe that GDI will become a valuable supplement to conventional metrics.

The contributions of this work include identifying the relationship between influence and disruptiveness, highlighting the limitations of DI, and proposing GDI as a more appropriate metric for measuring the disruptiveness of scientific publications. The results have implications for academic evaluation and scientific development, as well as for future research in the area of bibliometrics.

While the current study has provided valuable insights into the relationship between disruptiveness and influence, there are several limitations to our research. First, our analysis was limited to the local citation network of the Information Science and Library Science field from the Web of Science database, which may not fully capture the inter-domain impact of publications. Second, although GDI considers up to k generations of citation documents, we only calculated up to 2 generations in our empirical study. Future research should explore the effects of considering higher-order generations. Third, while we demonstrated the consistency of GDI with citation count, future research should examine the convergent validity and consistency of disruptiveness with peer reviews or expert databases, such as F1000prime.

In addition, future research could explore how GDI can be extended to higher levels of scientific entities, such as journals, authors, and institutions. However, the computational complexity of GDI with higher orders of generation may require the use of distributed computing systems. Finally, the robustness of GDI in different orders of generation and in various discipline fields also warrants further investigation.

Overall, the global disruptive index (GDI) is a promising metric that can account for the heterogeneity of disruptively innovative works. However, further research is needed to fully explore its potential and limitations.

Acknowledgements
This research was supported by the National Social Science Fund of China (19BTQ062) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_0076). Data and Python codes for visualization can be found via https://github.com/AlexJieYang/ISSI2023-GDI.

References

An, X., Sun, X., & Xu, S. (2022). Important citations identification with semi-supervised classification model [Article]. *Scientometrics*, *127*(11), 6533-6555. https://doi.org/10.1007/s11192-021-04212-6

Azoulay, P. (2019). Small research teams 'disrupt' science more radically than large ones. *Nature*, *566*, 330-332. https://doi.org/10.1038/d41586-019-00350-3

Barabasi, 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

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020a). Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers. *Quantitative Science Studies*, *1*(3), 1242-1259. https://doi.org/10.1162/qss_a_00068

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020b). Disruptive papers published in Scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by Wu, Wang, and Evans (2019). *Scientometrics*, *123*(2), 1149-1155. https://doi.org/10.1007/s11192-020-03406-8

Bornmann, L., & Tekles, A. (2019). Disruption index depends on length of citation window [Article]. *Profesional De La Informacion*, *28*(2), Article e280207. https://doi.org/10.3145/epi.2019.mar.07

Bower, J. L., & Christensen, C. M. (1995). *Disruptive technologies: catching the wave* (Vol. 73(1)). Harvard Business Review.

Bu, Y., Waltman, L., & Huang, Y. (2021). A multidimensional framework for characterizing the citation impact of scientific publications. *Quantitative Science Studies*, *2*(1), 155-183. https://doi.org/10.1162/qss_a_00109

Cagan, R. (2013). The San Francisco Declaration on Research Assessment. *Disease Models & Mechanisms*, *6*(4), 869-870. https://doi.org/10.1242/dmm.012955

Chen, J., Shao, D., & Fan, S. (2021). Destabilization and consolidation: Conceptualizing, measuring, and validating the dual characteristics of technology. *Research Policy*, *50*(1), Article 104115. https://doi.org/10.1016/j.respol.2020.104115

Fortunato, S., Bergstrom, C. T., Boerner, K., Evans, J. A., Helbing, D., Milojevic, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabasi, A.-L. (2018). Science of science. *Science*, *359*(6379), Article eaao0185. https://doi.org/10.1126/science.aao0185

Funk, R. J., & Owen-Smith, J. (2017). A Dynamic Network Measure of Technological Change. *Management Science*, *63*(3), 791-817. https://doi.org/10.1287/mnsc.2015.2366

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, *14*(3), 195-201. https://doi.org/10.1002/asi.5090140304

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429-431. https://doi.org/10.1038/520429a

Hu, X., Rousseau, R., & Chen, J. (2011). On the definition of forward and backward citation generations. *Journal of Informetrics*, *5*(1), 27-36. https://doi.org/10.1016/j.joi.2010.07.004

Kuhn, T. S. (1962). Historical structure of scientific discovery. *Science*, *136*(3518), 760-764.

Leydesdorff, L., & Bornmann, L. (2021). Disruption indices and their calculation using web-of-science data: Indicators of historical developments or evolutionary dynamics? *Journal of Informetrics*, *15*(4), Article 101219. https://doi.org/10.1016/j.joi.2021.101219

Leydesdorff, L., Tekles, A., & Bornmann, L. (2021). A proposal to revise the disruption index. *El profesional de la información*. https://doi.org/10.3145/epi.2021.ene.21

Li, X., Zhao, C., Hu, Z., Yu, C., & Duan, X. (2022). Revealing the character of journals in higher-order citation networks. *Scientometrics*, *127*(11), 6315-6338. https://doi.org/10.1007/s11192-022-04518-z

Min, C., Chen, Q. Y., Yan, E. J., Bu, Y., & Sun, J. N. (2021). Citation cascade and the evolution of topic relevance. *Journal of the Association for Information Science and Technology*, *72*(1), 110-127. https://doi.org/10.1002/asi.24370

Prathap, G. (2019). Letter to the editor: Revisiting the h-index and the p-index. *Scientometrics*, *121*(3), 1829-1833. https://doi.org/10.1007/s11192-019-03240-7

Schumpeter, J. A. (1912). *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Ullah, A., Wang, B., Sheng, J., Long, J., Khan, N., & Sun, Z. (2021). Identifying vital nodes from local and global perspectives in complex networks. *Expert Systems with Applications*, *186*, 115778. https://doi.org/10.1016/j.eswa.2021.115778

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

Wang, D., & Barabási, A.-L. (2021). *The science of science*. Cambridge University Press.

Wang, S., Ma, Y., Mao, J., Bai, Y., Liang, Z., & Li, G. (2022). Quantifying scientific breakthroughs by a novel disruption indicator based on knowledge entities. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24719

Wu, L. F., Wang, D. S., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378-+. https://doi.org/10.1038/s41586-019-0941-9

Xu, S. Q., Mariani, M. S., Lu, L. Y., & Medo, M. (2020). Unbiased evaluation of ranking metrics reveals consistent performance in science and technology citation data. *Journal of Informetrics*, *14*(1). https://doi.org/10.1016/j.joi.2019.101005

Zeng, A., Shen, Z. S., Zhou, J. L., Wu, J. S., Fan, Y., Wang, Y. G., & Stanley, H. E. (2017). The science of science: From the perspective of complex systems. *Physics Reports-Review Section of Physics Letters*, *714*, 1-73. https://doi.org/10.1016/j.physrep.2017.10.001

Appendix

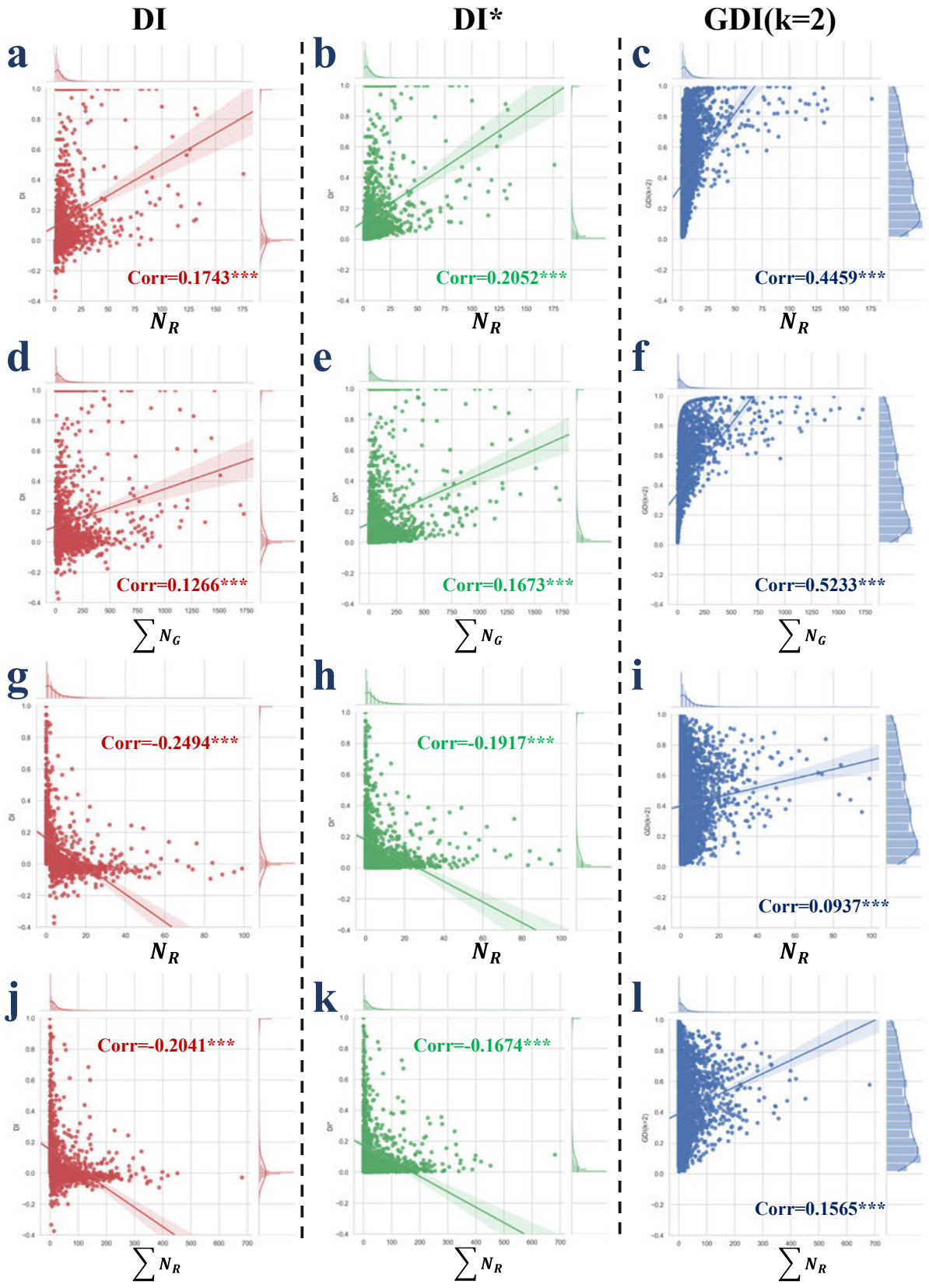


Figure A1. Scatter plots, fitting curves, distributions, and the Pearson correlations for N_G, N_R, $\sum N_G$ and $\sum N_R$ with each metric of papers published in *Scientometrics*, respectively, wherein the correlation for (a) DI and N_G is 0.1743 (p<0.001), (b) DI* and N_G is 0.2052 (p<0.001), (c) GDI(k=2) and N_G is 0.4459 (p<0.001), (d) DI and $\sum N_G$ is 0.1266 (p<0.001), (e) DI* and $\sum N_G$ is

0.1673 (p<0.001), (f) GDI(k=2) and $\sum N_G$ is 0.5233 (p<0.001), (g) DI and N_R is -0.2494 (p<0.001), (h) DI* and N_R is -0.1917 (p<0.001), (i) GDI(k=2) and N_R is 0.0937 (p<0.001), (j) DI and $\sum N_R$ is -0.2041 (p<0.001), (k) DI* and $\sum N_R$ is -0.1674 (p<0.001), (l) GDI(k=2) and $\sum N_R$ is 0.1565 (p<0.001). And these metrics' distribution histograms are illustrated in the subgraphs on the edges.

*** p<0.05, ** p<0.01, *** p<0.001.**

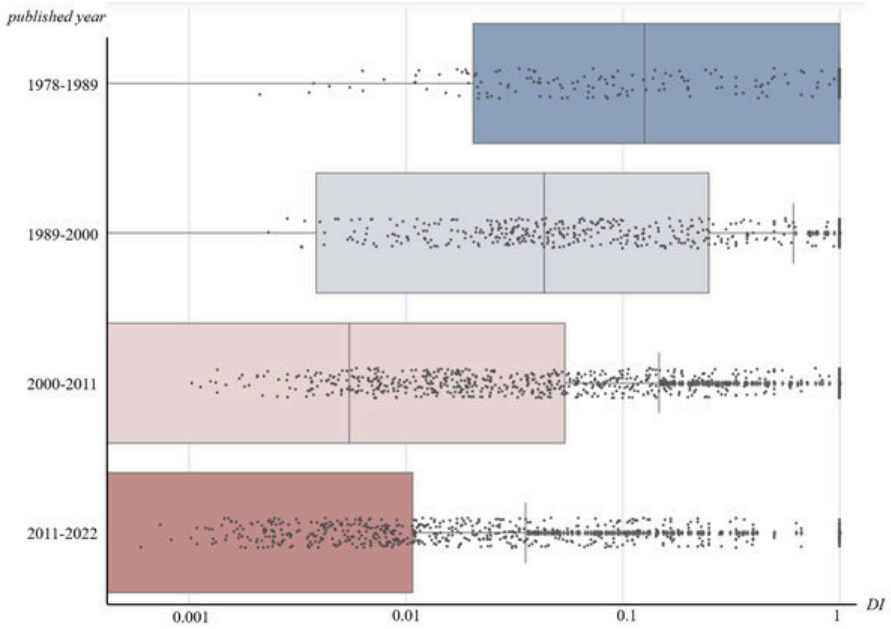


Figure A2. Box diagram for the distribution of DI for different published years, for papers published in 1978-1989, 1989-2000, 2000-2011 and 2011-2022, respectively, with logarithmic processing of horizontal data. Colored regions represent 95% confidence intervals.

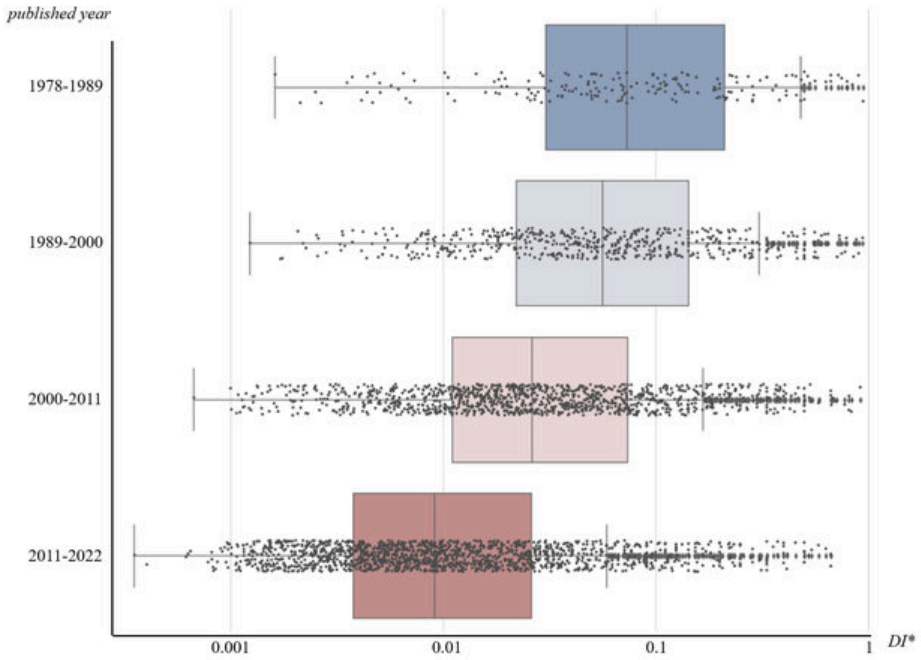


Figure A3. Box diagram for the distribution of DI* for different published years, for papers published in 1978-1989, 1989-2000, 2000-2011 and 2011-2022, respectively, with logarithmic processing of horizontal data. Colored regions represent 95% confidence intervals.