KU LEUVEN

FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN

Management, Strategy and Innovation (MSI)

The comparison of various similarity measurement approaches on interdisciplinary indicators

Ying Huang, Wolfgang Glänzel, Bart Thijs, Alan L. Porter and Lin Zhang

MSI_2102

The comparison of various similarity measurement approaches on interdisciplinary indicators

Ying Huang[1], Wolfgang Glänzel[1], Bart Thijs[1], Alan L. Porter[2,3], Lin Zhang[4]

[1] Department of MSI, Centre for R&D Monitoring (ECOOM), KU Leuven, Leuven, Belgium

[2] Science, Technology & Innovation Policy, Georgia Tech, Atlanta, GA, United States,

[3] Search Technology, Inc., Norcross, GA, United States

[4] School of Information Management, Wuhan University, Wuhan, China

Abstract: How to measure the interdisciplinary is a crucial topic in Interdisciplinary research (IDR), and the integrated indicators (e.g., Rao-Stirling) that combine three distinct components (variety, balance and disparity) has become one of the most promising attempts. Among the three components, variety and balance play relatively straightforward roles in diversity assessment but to what extent the (dis)similarity measuring approaches may affect the interdisciplinarity indicators is seldom discussed in the literature. In this paper, we compare various similarity measurement approaches from (1) different subject classification systems, (2) different normalization of (dis)similarity measure, (3) different (dis)similarity matrices of subjects, (4) different time windows; and (5) different levels of aggregations, using the academic publications labeled "Article" in eight selected journals published during the period 2009–2018 were selected as the sample dataset. Our results corroborate the following findings: First, a finer classification system with more subject categories increases the possibility that one cites sources from different subject categories. Second, different normalization approaches may lead to obviously different interdisciplinarity results, and such a finding is supported by the relatively low correlations between the interdisciplinarities calculated by Salton's Cosine and Ochiai's Cosine. Third, on the basis of Salton's cosine normalization, the interdisciplinary values obtained by different settings are highly correlated, especially in terms of different citation similarity matrices (cited, citing and cross-citation) and, in general, with different time windows. Fourth, results based on an aggregated dataset tend to overly expand the 'interdisciplinarity' degree of a journal, especially when the focused journal is actually 'multidisciplinary'.

Keywords: Interdisciplinary research (IDR); Similarity; Rao-Stirling; True Diversity;

Introduction

Interdisciplinary research (IDR) has been increasingly viewed as the way to advance fundamental understanding and, also, to solve problems whose solutions are beyond the scope of a single field of research practice for challenging contemporary scientific and societal problems (Porter & Rafols, 2009). The time appears ripe for new approaches to measure IDR, as advances since the 1980s in science and engineering information resources come together with potent computing power and analytical software (Porter et al., 2006). Several methods have been discussed, and these approaches enable characterization of various research elements and units (e.g., papers, journals, researchers, collections of researchers or institutes and fields) in terms of their degree of interdisciplinarity using various bibliometric information sources (e.g., published journals, cited references, or citing publications).

The concept and measurement of interdisciplinarity are originally borrowed from the 'diversity' in ecology to evaluate species richness and species evenness. In the early, the Shannon index, the Simpson index, and the Gini-Simpson index are proposed to measure diversity (Jost, 2006). C Radhakrishna Rao (1982) proposed the disparity as the third element to characterize diversity. Stirling (1998) translated diversity measures into a framework for the measurement of "interdisciplinarity" in science policy and research evaluation.

The basic common consensus seems to be that interdisciplinarity concerns the integration of knowledge generated in different disciplines and that it has three distinct components: the number of disciplines cited (variety), the distribution of citations among disciplines (balance), and how similar or dissimilar these categories are (disparity)(Leinster & Cobbold, 2012; Stirling, 2007). Several studies have composed these three aspects into one single indicator, such as Rao-Stirling (C. Radhakrishna Rao, 1982; Stirling, 2007), which is modified to DIV by measuring "balance" and "variety" independently (Leydesdorff, 2018; Leydesdorff et al., 2019). The indicator was refined by (Rousseau, 2019) that proposed by transforming to DIV* = N.DIV to meets the "effective number requirement" of Leinster and Cobbold (2012) and thus measures "true" diversity. Other studies have considered one or more of the three components separately, investigating their separate relationships with citation impact and citation delay (Wang et al., 2015). In addition, network coherence, an indicator to measure the intensity of similarity relations within a dataset, was proposed to reflects the novelty of its knowledge integration (Rafols & Meyer, 2010). However, the indicators and solutions, which have been proposed to measure the extent of IDR, do not only provide opportunities but also have their limitations and drawbacks in representing IDR, and these indicators may even produce conflicting results. Wang and Schneider (2019) examined the validity and relations between these measures, and they come to the deviant results when comparing measures that supposedly should capture similar features or dimensions of the concept of interdisciplinarity.

Among the three components, variety and balance seem to play relatively straightforward roles in the process of diversity assessment. Disparity relying on a specific distance or similarity metrics between pairs of disciplines refers to the degree to which disciplines of different similarities may be distinguished. One can measure disparity in terms of the distances between elements. However, the measurement of disparity is sensitive to the choice of unit distance or proximity (Leydesdorff & Ivanova). Different dissimilarity or similarity measures may, however, result in divergent results with respect to the degree of interdisciplinarity. To what extent the similarity measuring approaches may affect the interdisciplinarity indicators is seldom discussed in the literature.

In this paper, we address the above research question from four respects, respectively, comparisons among (1) different subject classification systems; (2) different normalization of (dis)similarity measure; (3) different (dis)similarity matrices of subjects; (4) different time windows; and (5) different levels of aggregations.

Data and Framework

Data

In this study, academic publications labeled "Article" in eight selected journals published during the period 2009–2018 were downloaded from Clarivate Analytics Web of Science (WoS). The eight selected journals were chosen to represent specialties with assumingly different extent of specialization and interdisciplinarity: *Bioinformatics* (focus more on the field of natural sciences) and *Energy Policy* (focus more on the field of social sciences) have, in general, quite broad coverage and a high degree of interdisciplinarity; *Harvard Law Review* (Social Sciences Citation Index) and *Journal of Differential Equations* (Science Citation Index Expanded) are specialized and has a relatively low degree of interdisciplinarity; *Nature* and *Science* are well-known multidisciplinary journals with a broad coverage embracing practically all fields of the sciences, social sciences and humanities publishing both specialized and interdisciplinary articles; *Journal of Informetrics* and *Scientometrics* are specialized journal devoted to the subdisciplines of library and information science, which are, however, as such assumed to be interdisciplinary.

Table 1. The WoS category and research areas of the selected journals

Abbreviation Name	WoS Category (WC)	Research Area
Nature	Multidisciplinary Sciences	Science & Technology - Other Topics
Science	Multidisciplinary Sciences	Science & Technology - Other Topics
Bioinformatics	Biochemical Research Methods; Biotechnology & Applied Microbiology; Mathematical & Computational Biology	Biochemistry & Molecular Biology; Biotechnology & Applied Microbiology; Computer Science; Mathematical & Computational Biology; Mathematics
Energy Policy	Economics; Energy & Fuels; Environmental Sciences; Environmental Studies	Business & Economics; Energy & Fuels; Environmental Sciences & Ecology
Harvard Law Rev	Law	Government & Law
J Differ Equations	Mathematics	Mathematics
J Informetrics	Computer Science, Interdisciplinary Applications; Information Science & Library Science	Computer Science; Information Science & Library Science
Scientometrics	Computer Science, Interdisciplinary Applications; Information Science & Library Science	Computer Science; Information Science & Library Science

Note: Research areas constitute a subject categorization scheme that is shared by all Web of Science product databases. Journals and books covered by Web of Science WoS Core Collection (WoS) are assigned to at least one Web of Science category (WC). Each Web of Science category is mapped to one research area.

In total, 37,953 research articles were used as the basis of the study. The annual change of the number of published records in these four journals during 2009–2018 is given in Figure 1. All the data are extracted from the in-horse dataset of ECOOM, and the year used here is the rather than the publication year is shown in the Web of Science Core Collection platform. Notably, 326 records have no cited references or that have not been assigned with any WCs, and 37627 records are used as the sample dataset in this study.

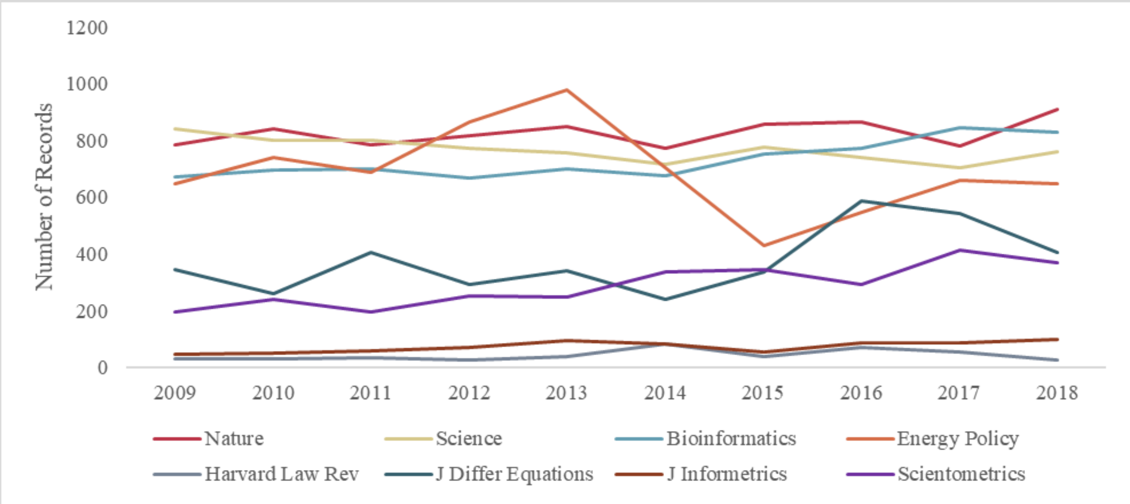


Figure 1. The annual publication records in eight selected journals during 2009-2018.

Framework

Following our research questions, namely how the results are affected by applying the five different schemes, we propose the analytic framework shown in Figure 2. We analyze the effect of similarity measurement on interdisciplinary indicators by different settings of the subject classification system, normalization measures, subject similarity matrix, time window and aggregation levels.

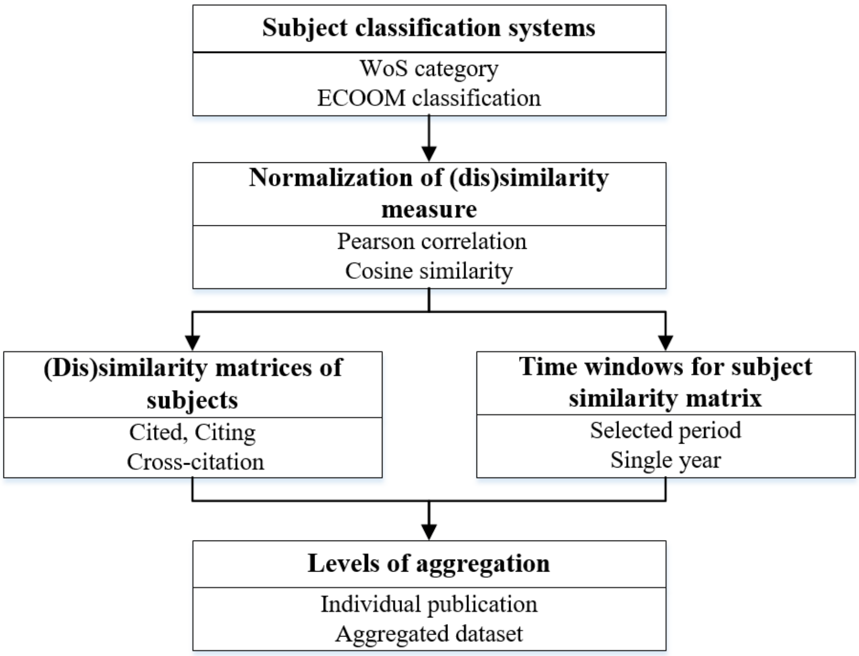


Figure 2. The analytic framework

Most current bibliometric measures of IDR largely rely on journal-based subject categories provided along with the respective bibliographic databases. The use of such journal-based subject assignment has become the main basis for scientometric measures, such as the 250+ WCs tagged to the 12,000+ journals indexed in the WoS, the 22-broad field classification scheme of the Essential Science Indicators (ESI) database by Clarivate Analytics, the Subject Category (WC), 334 subject areas listed in the Scopus of Elsevier. In addition, bibliometricians from Leuven and Budapest developed a hierarchical scheme: the Leuven–Budapest (ECOOM) subject classification (Glänzel & Schubert, 2003), with 16 major subject fields and 68 subfields, recently updated and revised to 74 subfields (Glänzel et al., 2016). As we aim to analyze the degree of interdisciplinarity, we prefer to exclude some of the extremely broad and/or multidisciplinary fields, such as "Multidisciplinary Sciences." The six newly added WCs in 2016: "Audiology & speech-language pathology," "Cell & Tissue Engineering," "Ergonomics," "Green & sustainable science & technology," "Logic," and "Nanoscience & Nanotechnology" and the former WC – "Biology, miscellaneous" – which is no longer in use since 2015, are not included in our analysis as all documents assigned to these categories have simultaneously assignment to other, narrower and more specialized subjects. Therefore, in this paper, 245 WCs (WC-245), 15 ECOOM major subject fields (Filed-15) and 73 ECOOM subfields (DISC-73) are taken into consideration.

When it comes to the normalization of a matrix, there are two problems that have to be distinguished: the methods of normalization and the type of matrix to be normalized. In the cocitation patterns, McCain (1990) noted that Pearson's correlation coefficient could be used as a similarity measure with some advantages. Ahlgren et al. (2003) questioned the use of Pearson's correlation coefficient as a similarity measure on the grounds that this measure only measures the degree of a linear dependency and is sensitive to zeros. Instead, they suggested applying Salton's cosine measure, proposed in (Salton & McGill, 1986), particularly if one aims at the visualization of the structure, as in the case of social network analysis or multidimensional scaling (MDS) (Ahlgren et al., 2003). Unlike Salton's cosine and the Pearson correlation, the Jaccard index (originally called Community Coefficient)) abstracts from the shape of the distributions and focus only on the intersection and the sum of the two sets, so it can be considered as an advantage in the co-occurrence matrix (Leydesdorff, 2008). Besides, there are some other normalizations for similarity measures, such as probabilistic activity index (PAI) (Zitt et al., 2000), focuses only on the strength of the co-occurrence relation; Association Strength (AS), the estimated co-occurrence probability of the concepts(van Eck et al., 2006), etc. Adnani et al. (2020) compared the five similarity indexes (Jaccard, Dice-Sorensson, Salton, Pearson, and Association Strength) for the three types (co-word, cocitation and co-authorship) of scientometric analysis and concluded that no consensus on the appropriateness of an index for co-word and co-authorship analyses, while Salton is the widely preferred one for cocitation. In this paper, we mainly use the Pearson's correlation coefficient, Salton's cosine and Ochiai's cosine, an index is based on binary or scalar values (also known as the Ochiai index) (Ochiai, 1957). Ochiai's cosine was used in the study of Zhang et al. (2016), and its formulate is:

$$S_{ij} = \frac{c_{ij} + c_{ji}}{\sqrt{(TC_i + TR_i)(TC_j + TR_j)}}$$

where *i* and *j* refer to subject fields ($i \neq j$), $c_{ij} + cji$ is equal to the total number of cross-citations between subject fields *i* and *j*; TC_k denotes the total number of citations received by subject field k ($k = i, j$) (from other subject fields) and TR_k denotes the total number of citations given by subject field *k* ($k = i, j$) (to other subject fields).

After setting the normalization measure, we need to consider the citation matrix accordingly, which comprises three structures: the directed citing, cited dimension and the undirected cross-citation relation. The "citing pattern" refers to the current knowledge base of the downloaded dataset and indicates the knowledge integration from references into publications (Leydesdorff & Rafols, 2009). The "cited pattern" refers to the structure in the (cited) archive and expresses the knowledge diffusion from focused publication to times cited (Leydesdorff, Rafols, et al., 2013). Whereas the "cross-citation pattern" merges the cited side and citing side together to present knowledge (Zhang et al., 2009; Zhang et al., 2010). In this paper, we would like to follow the above definition and compare the discrepancy of interdisciplinarity measurement based on the above three dimensions. The normalized the cited-, citing- cross-citation - matrices of the subject fields is calculated based on Salton-cosine similarity as follows:

$$S_{ij} = \frac{\sum_{k=1}^{n} c_{ik} * c_{jk}}{\sqrt{\sum_{k=1}^{n} c_{ik}^2}\sqrt{\sum_{k=1}^{n} c_{jk}^2}}$$

where i and j refer to subject fields ($i \neq j$), c_{ik} *and* c_{jk} denote the number of cited-, citing-, cross-citations received by subject field k (from other subject fields) or given by subject field k (to other subject fields).

Most of the current similarities for interdisciplinarity is conducted on the level of disciplinary category, which is mainly represented by manually assigned categories to the journals, e.g., the WCs are tagged to WoS journals that are updated on a monthly basis (Clarivate Analytics, 2019), especially for the journals included in the Emerging Sources Citation Index (ESCI) (Huang et al., 2017). Therefore, once the journals changed, the citation relation among the disciplinary categories will change as well. Besides, the strength of these citation relationships among the disciplinary categories is evolving rather than static. These phenomena are reflected in a series of overlap mapping works (Carley et al., 2017; Leydesdorff & Rafols, 2009; Rafols et al., 2010). Therefore, the different time windows for subject similarity matrix are also explored.

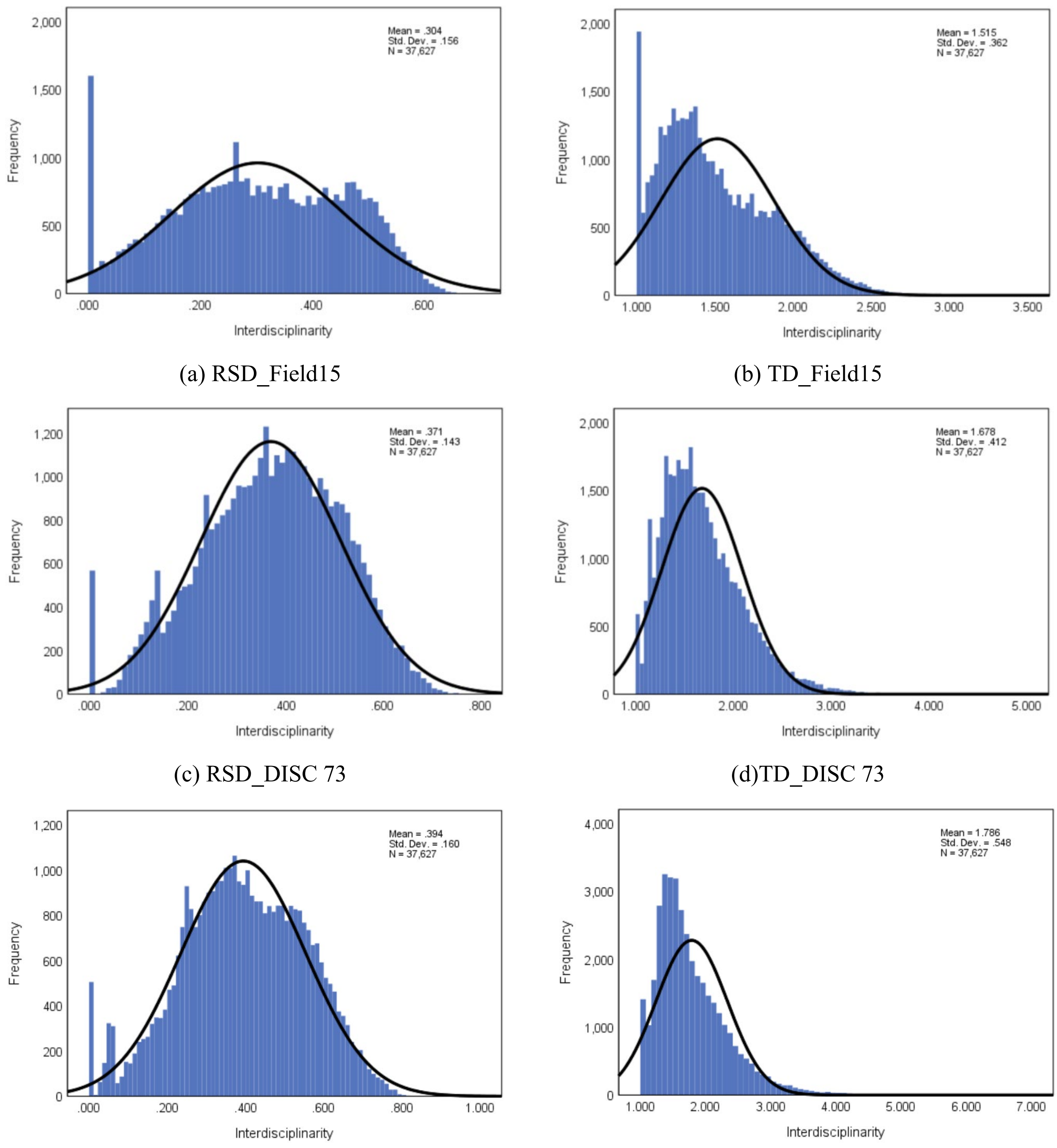
In the current study, we assume that the measures used to analyze diversity need to capture the intuitive notion of diversity, and the three components discussed should all be taken into consideration. We adopt the Rao-Stirling diversity (RSD), as a classical integrated diversity measure, and the True Diversity (TD) (Zhang et al., 2016), as a new proposed integrated measure in this study, to compare the interdisciplinarity values based on different similarity measuring approaches. Average values of interdisciplinary for the eight journals are calculated accordingly, as well as an integrated diversity value for the aggregated set of papers in the journal.

Results

Different subject classification systems: WC versus ECOOM classification

Several WoS-based classification systems are used in bibliometric studies, where WC and ECOOM classification representing different levels of granularity belong to the most commonly used schemes. The question of granularity is one of the crucial points in IDR studies, both from the conceptual (topic interdisciplinarity vs. field interdisciplinarity) and the methodological viewpoint (cf. Glänzel et al. (2021), where this issue is deepened and tackled in detail).

The distribution of interdisciplinarity values for individual publications measured by RSD and TD based on different subject classifications are shown in Figure 3. As expected, as the subject classification becomes more fine-grained (moving from ECOOM Field-15 over DISC-73 to WC-245), the Mean, Median and Maximum values of interdisciplinarity of all publications present an obvious increasing trend, as well as the Variance (see also Glänzel et al. (2021)). This is why, on the one hand, a finer granulation with more subject categories increases the possibility that one cites sources from different subject categories, and on the other hand, the multiple assignments, e.g., the journals are allocated into more than one subject category. Furthermore, by comparing the RSD and TD calculated for the publication dataset under study, we confirm that the Hill-type indicator TD gives more weight to variety and has better discriminative power than the classical Rao-Stirling indicator (Zhang et al., 2016). Considering the interdisciplinarity in the WC-245 has a good distribution and better distinction when comparing the papers with frequent interaction in the same broad fields, we choose the WC-245 as the benchmarking subject classification in the following analyses.



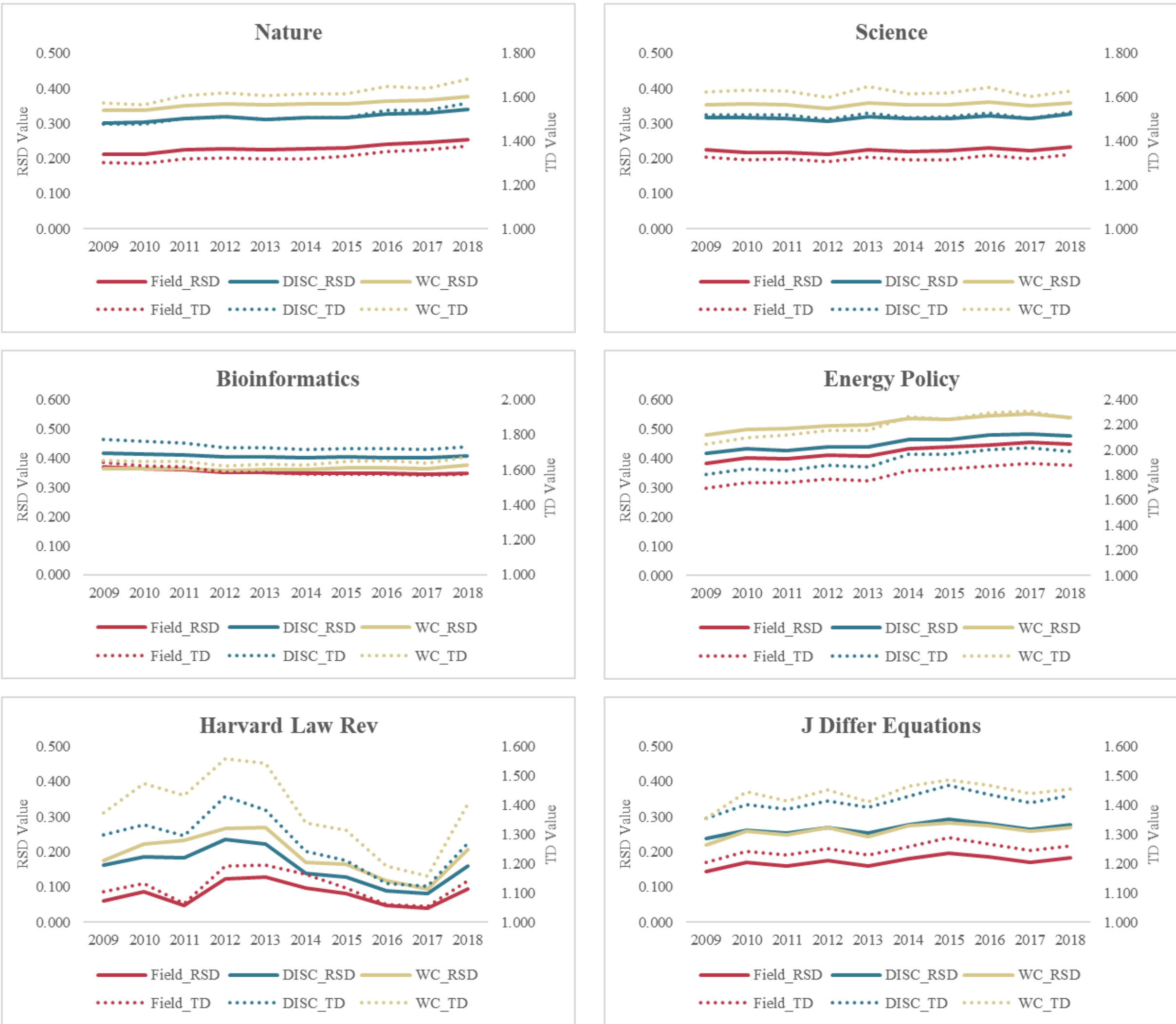
(a) RSD_Field15 (b) TD_Field15

(c) RSD_DISC 73 (d)TD_DISC 73

(e) RSD_WC-245 (f) TD_WC-245

Figure 3. The distribution of interdisciplinarity scores for individual publications measured by RSD and TD based on different subject classifications

The average values of Interdisciplinarity indicators for each journal understudy during 2006–2015 are shown in Figure 4. It is not surprising to see that legal journal *Harvard Law Rev.* and mathematics journal *J. Differ Equ.* have, in general, much lower diversities compared to the others, but the former shows some fluctuations and lower value because it may have fewer chances to cite natural science studies; *Nature* and *Science,* as two representative multidisciplinary scientific journals, show similar trends: their publications are not so interdisciplinary because more than half of Nature's papers have come from the life sciences over its recent history (Gates et al., 2019); *Bioinformatics*, a highly interdisciplinary journal covering biochemical research methods, biotechnology & applied microbiology, and mathematical & computational biology and others keep the interdisciplinary degree at a stable high level, at least in the mirror of the chosen indicators. *Energy Policy*, another highly interdisciplinary journal covering economics, energy & fuels, environmental sciences, and environmental studies, presents the most interdisciplinary feature due to its research areas are cross natural science and social science. Such characteristics are also shown in *J. Informetrics* and *Scientometrics*, two well-known journals in the fields of quantitative science studies that have bridged research traditions in information science, computer sciences, and management science (especially in science, technology, and innovation policy).



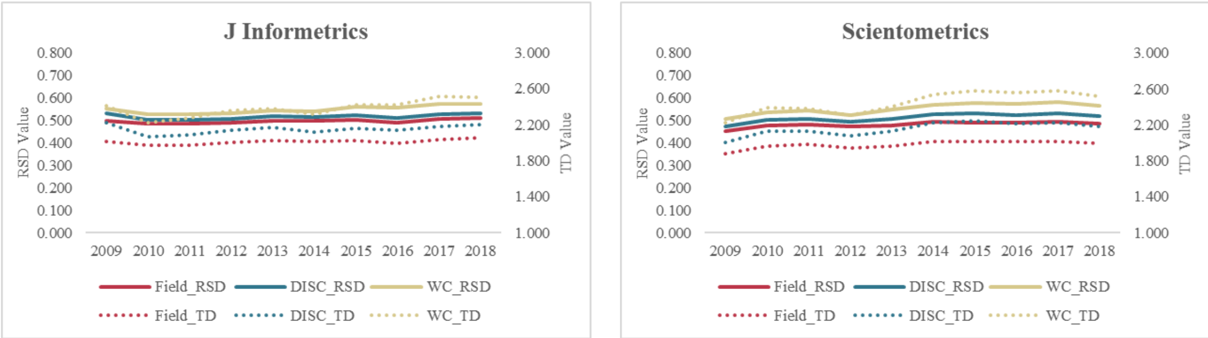
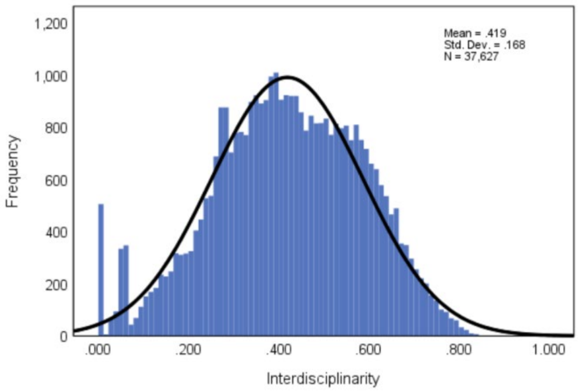


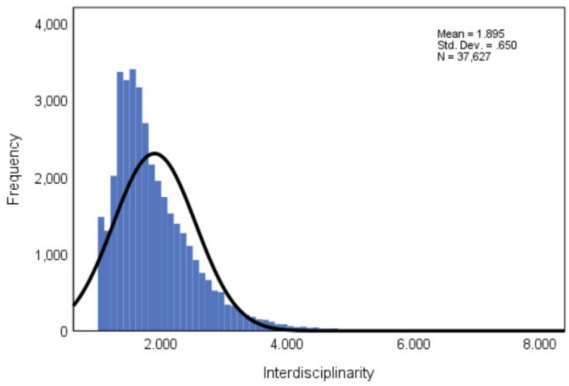
Figure 4. The average value of interdisciplinarity scores for different subject classification systems of the eight journals

Different normalization of (dis)similarity measure: Pearson versus Cosine

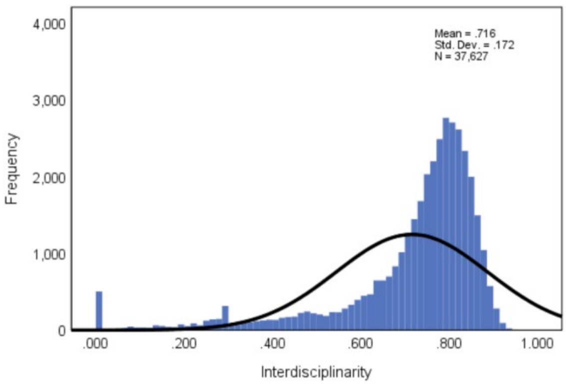
The normalization of (dis)similarity measure depends on the applied feature soft data, the relationship between the similarity index and its mathematical formulas (vector and non-vector), and the type of analysis to be performed (Adnani et al., 2020; Schneider & Borlund, 2007). When it comes to quantitatively evaluate the accuracy of relatedness measures or the resulting maps, the Pearson correlation is the most accurate raw relatedness measure (Klavans & Boyack, 2006). But they revealed that the cosine-normalized asymmetrical occurrence matrix provides more reasonable and intuitive results and more consistent with consensus science mapping. In this paper, we apply the Pearson's correlation coefficient, Salton's cosine and Ochiai's cosine as the ways to calculate the similarity between the WCs, then the interdisciplinarity of the records in the dataset. The distributions are shown in Figure 5. Compared with Pearson's coefficient and Salton's Cosine, the Ochiai's Cosine indicates a higher discriminatory power according to its' peakedness and kurtosis, both in RSD and TD. The interdisciplinarity values calculated by Pearson's coefficient and Salton's Cosine show similar distributions.



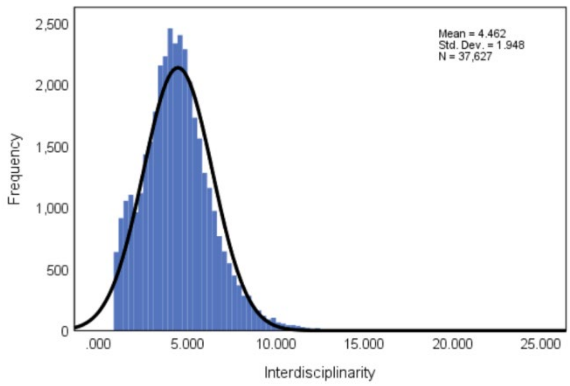
(a) RSD_Pearson coefficient



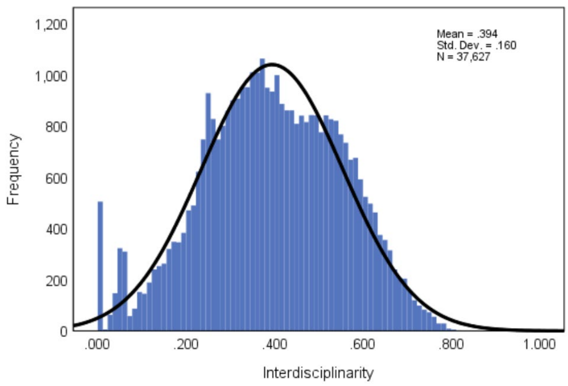
(b) TD_Pearson coefficient



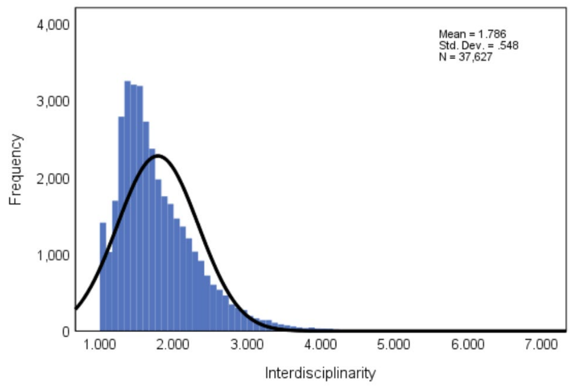
(c) RSD_ Ochiai's Cosine



(d) TD_ Ochiai's Cosine



(e) RSD_ Salton's Cosine



(f) TD_ Salton's Cosine

Figure 5. The distribution of interdisciplinarity scores for individual publications measured by RSD and TD based on normalization of (dis)similarity measure

The phenomenon is also observed in the selected eight journals, shown in Figure 6. The average interdisciplinarity calculated by the Pearson coefficient and Salton's Cosine is very close every year. Generally, the value calculated by Ochiai's cosine is much higher, but its effect on the journals with different disciplinary are different. For the high interdisciplinary journals that cite broader fields, e.g., *J. Informetrics* and *Scientometrics*, it makes no obvious changes in the order ranked by interdisciplinarity when they integrate similar subject knowledge. However, once the journals cite more diverse subject knowledge, their interdisciplinarity value calculated by the above two (dis)similarity measures may show more uncertain rank by value, e.g., *Bioinformatics* and *Energy Policy*. For the specialized journals that cite fewer references in the limited fields, e.g., *Harvard Law Rev.*, it shows more fluctuations than the journals that cited fewer references in relatively broad fields, e.g., *J Differ Equations*. A list of the diversity indicators for the eight journals can be viewed in Table 7. Overall, Ochiai's cosine will highlight the slight difference in the original citation matrix, but Salton's Cosine cares more about the citation vector, aka, citation distribution among the two subject categories.

Figure 6. The average value of interdisciplinarity scores for different normalization of (dis)similarity measure of the eight journals

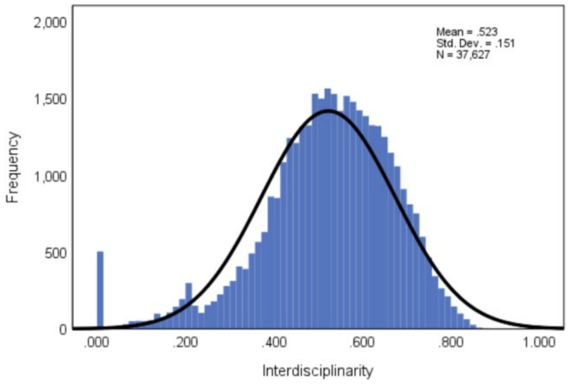
Table 7. Diversity indicators for the eight journals

Source	Num	Geni	Distance (Ochiai)	Distance (Salton)	Simpson	Shannon	TD (Ochiai)	TD (Salton)
Nature	10.380	0.556	0.961	0.548	0.767	1.847	4.581	1.618
Science	8.864	0.591	0.959	0.526	0.757	1.750	4.227	1.625
Bioinformatics	9.781	0.650	0.968	**0.519**	0.813	1.927	5.241	**1.646**
Energy Policy	9.903	0.635	0.960	**0.696**	0.793	1.855	4.882	**2.190**
Harvard Law Rev	2.861	0.775	**0.509**	0.417	0.216	0.431	1.535	1.344
J Differ Equations	4.402	0.686	**0.820**	0.492	0.592	1.108	2.205	1.440
J Informetrics	10.061	0.572	0.978	0.764	0.763	1.784	4.671	2.397
Scientometrics	10.402	0.616	0.971	0.755	0.772	1.827	5.074	2.475

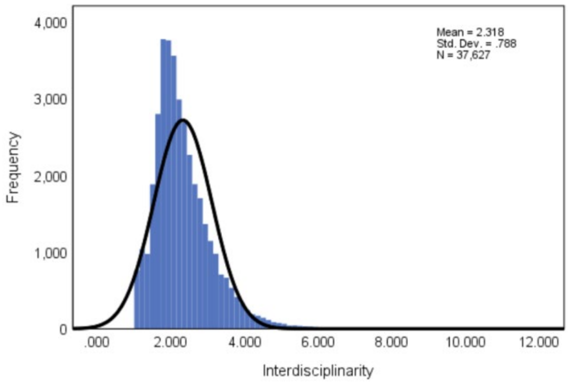
Note: The values are averaged of the individual record published in the selected journal.

Different (dis)similarity matrices of subjects: Cited versus Citing versus Cross-citation

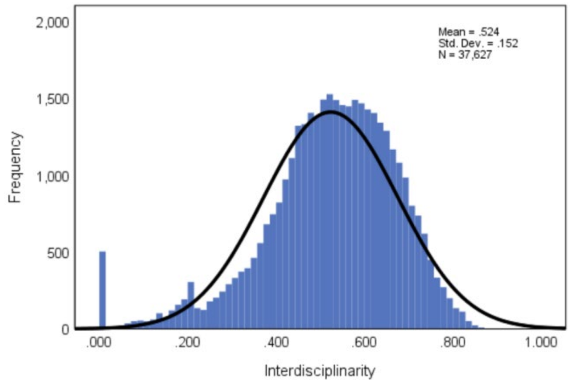
We constructed three different similarity matrices of WCs based on 'cited,' 'citing' and 'cross-citation' data, respectively, through a publication-journal-subject field classification scheme. Specifically, the cited, citing, and cross-citation matrix of individual publications is aggregated first into journal level and then into the WCs subject level. The normalized similarity matrix of all subject fields is finally calculated based on Salton's cosine similarity measure. Figure 5 presents the distribution of interdisciplinarity values for individual publications understudy, respectively using a similarity matrix of WCs based on 'cited' and 'citing' data. It is interesting to see that the two similarity matrices deriving from different citation patterns lead to almost the same interdisciplinarity value distribution. The cross-citation mode narrows the discrepancy among the subject categories because it combines the citation relationship from two directions, so the interdisciplinarity values are relatively lower 'cited' and 'citing' mode, both in RSD and TD.



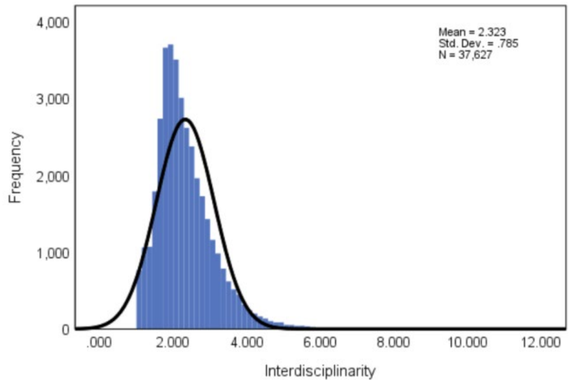
(a) RSD_Cited



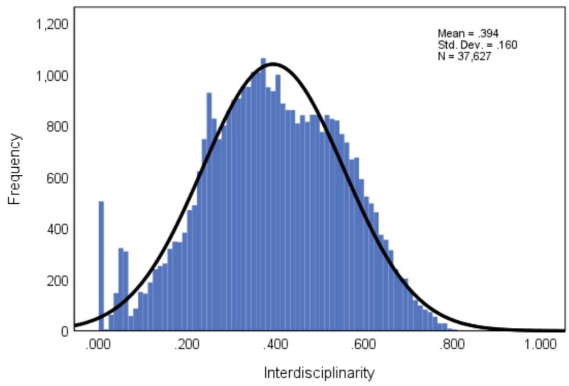
(b) TD_Cited



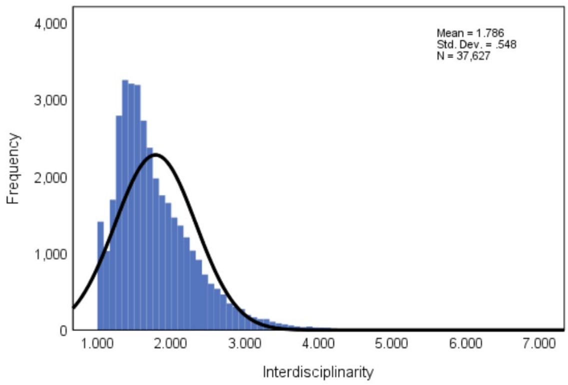
(c) RSD_Citing



(d) TD_Citing



(e) RSD_Cross-citation



(f) TD_Cross-citation

Figure 7. The distribution of interdisciplinarity scores for individual publications using the similarity matrix of WCs based on 'cited' and 'citing' data.

To further explore the influence of subject similarities based on different citation patterns, we present the average values of interdisciplinarity for the eight journals under study by year during 2009-2018 in Figure 8. No matter, which citation pattern we apply for the subject similarity calculation, the curves of interdisciplinarity values almost overlap for both RSD and TD, even though these eight journals have quite different citation characteristics. It may indicate that different similarity matrices of subject categories based on various citation patterns (cited versus citing) have a very limited impact on the interdisciplinarity index calculation. Again, the cross-citation based similarity matrices show a relatively lower value in all eight journals.

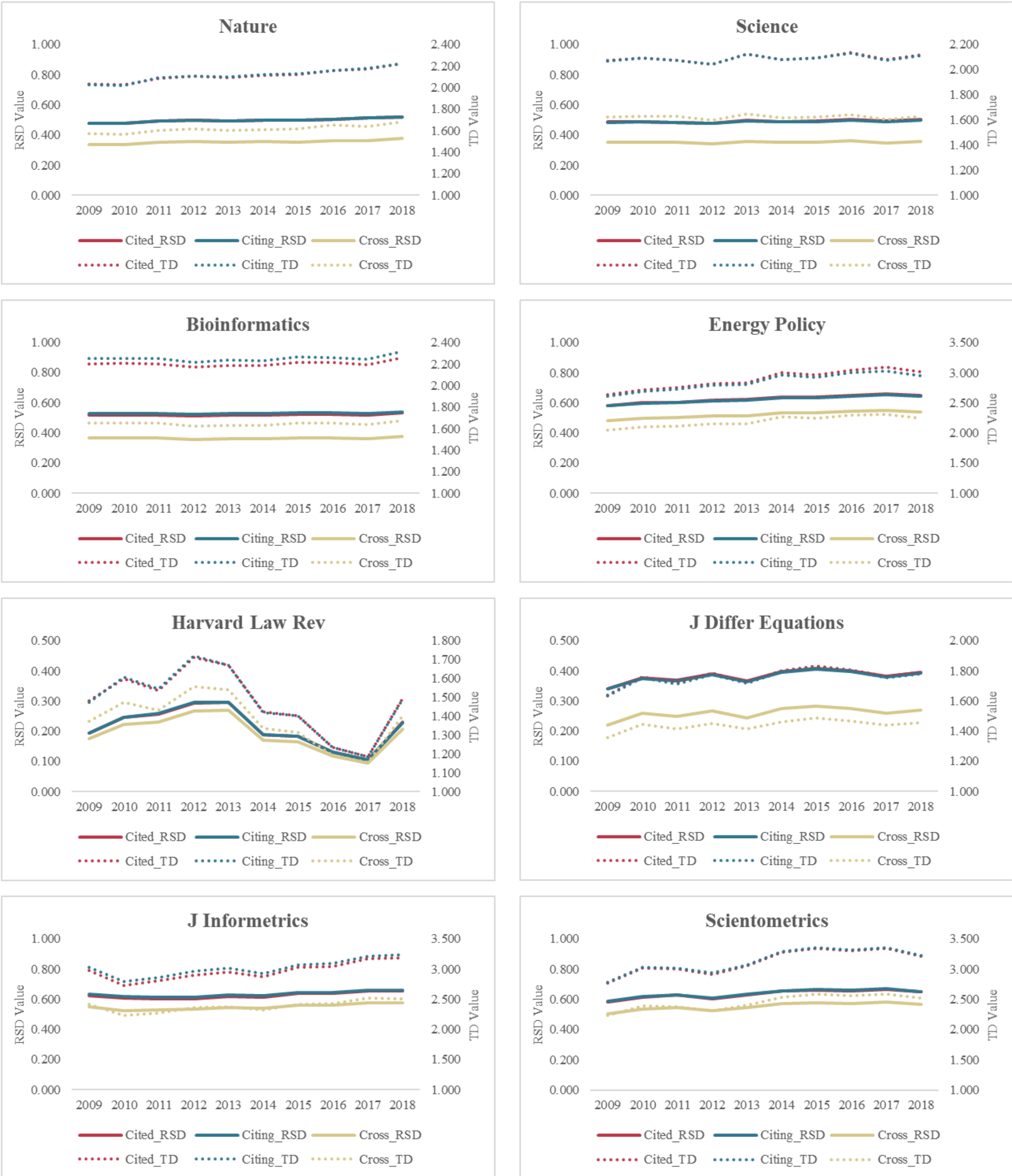
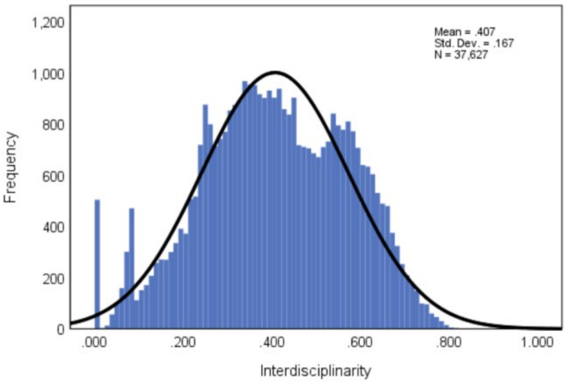


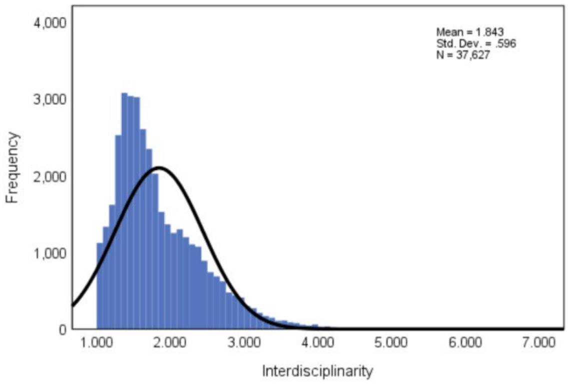
Figure 8. The average value of interdisciplinarity scores for different citation matrices of the eight journals

Different time windows for subject similarity matrix: 2009 versus 2018 versus 2009-2018

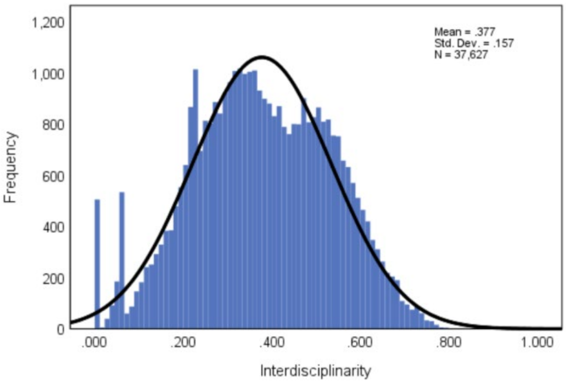
Another problem appearing in measuring IDR is how to choose benchmarking years or the time windows for the subject similarity matrix calculation. On the one hand, scholars from different institutions may not always have full access to the scientific data platform; on the other hand, many researchers are inclined to use up-to-date data as the information source. In this study, we respectively choose 2009, 2018 and the period of 2009-2018 as benchmarks to calculate the WCs subject cross-citation similarity matrices. The distribution of interdisciplinarity values for individual publications using subject similarity matrices based on different time windows is shown in Figure 9. It is easy to understand that the interdisciplinarity value must be updated once the similarity matrix changed, but the distributions show no observable deviation between the results based on the subject similarity matrix in the above years/ period.



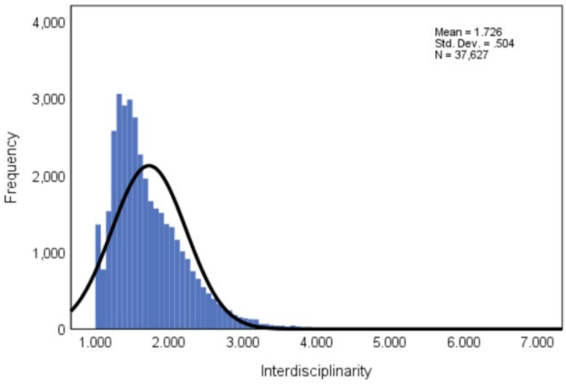
(a) RSD_PY2009



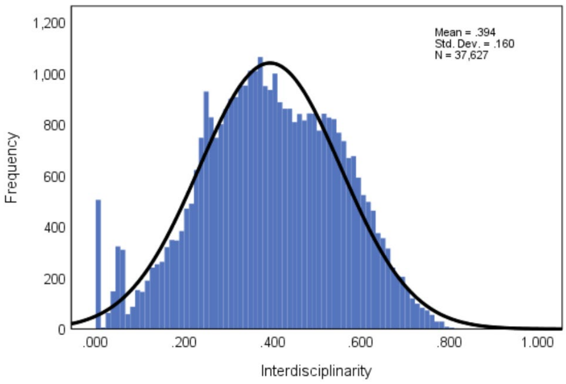
(b) TD_PY2009



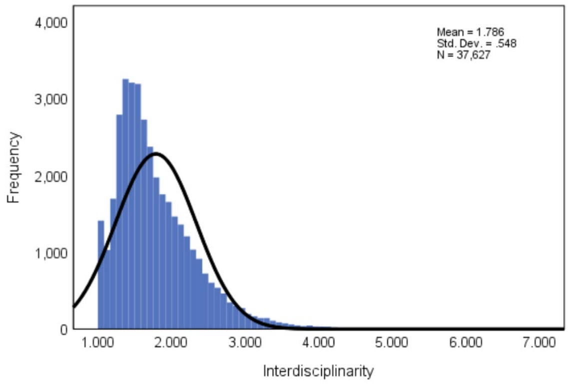
(c) RSD_PY2018



(d) TD_PY2018



(e) RSD_PY2009-2018



(f) TD_PY2009-2018

Figure 9. The distribution of interdisciplinarity values for individual publications using subject similarity matrices based on different time windows (2009 vs. 2018 vs. 2009-2018).

Figure 10 further presents comparisons of variation on the journal level. The interdisciplinarity calculated based on the similarity matrices constructed both in the time-interval (2009-2018) and in the single-year (2009 or 2018) shows high coincidence for the cases of these eight journals.

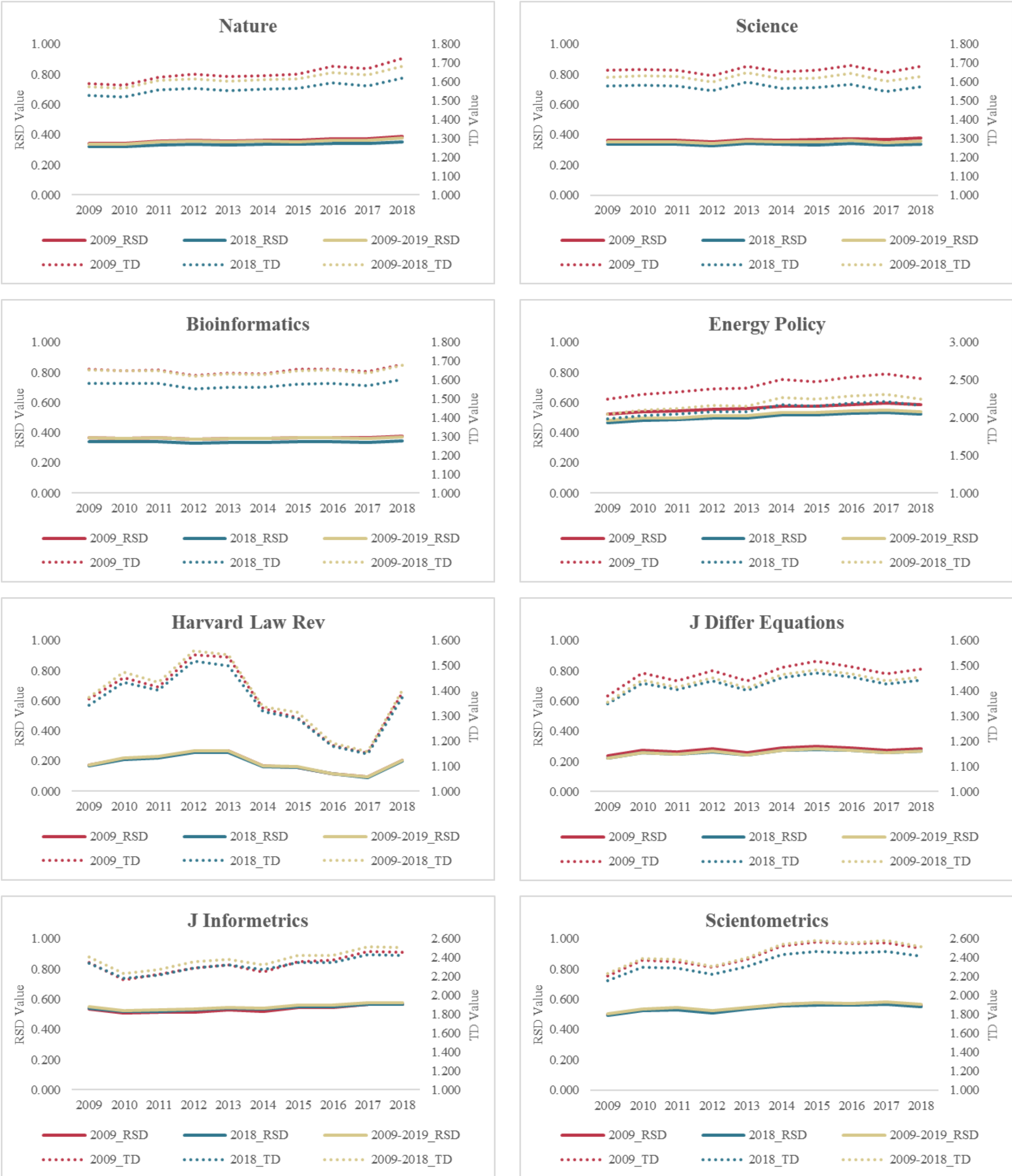


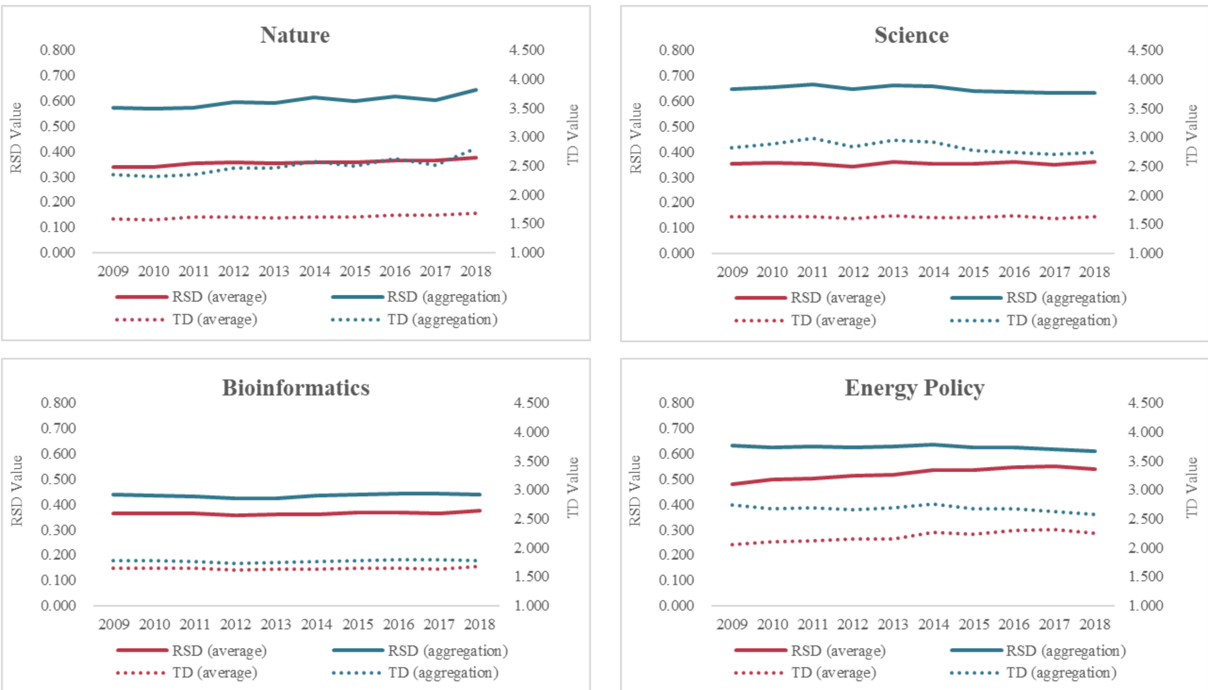
Figure 10. The average value of interdisciplinarity scores for different time windows for the subject similarity matrix of the eight journals

Different levels of aggregation: individual publications versus aggregated datasets

In order to assess the degree of interdisciplinarity for a journal during a particular time period, the most instinctive way is to average the interdisciplinarity values of each individual paper by classifying its references into one or more disciplines. Alternatively, it is possible to calculate one integrated diversity value for the aggregated set of papers in the journal. In particular, creating one matrix of all the cited WCs for that set of papers results in an overall diversity value for the journal under study. The interdisciplinary values of the eight journals during the period of 2009 to 2018 are shown in Figure 11.

Overall, the interdisciplinarity values of the aggregated datasets for these eight journals are higher than the average interdisciplinarity of individual publications (There is an exception, the aggregated interdisciplinarity is lower than the average interdisciplinarity in 2013 of *Harvard Law Rev.,* a very specialized journal, because the aggregation may change the balance distribution of the cited WCs distribution). Results based on an aggregated dataset tend to overly expand the 'interdisciplinarity' degree of a journal, especially when the focused journal is actually 'multidisciplinary', such as *Nature* and *Science*. The aggregated values directly generated from the aggregated datasets cannot make good distinctions between the two cases: a set of highly interdisciplinary papers and a combination of single-disciplined papers distributed in various subject fields. Therefore, the average interdisciplinarity of individual publications seems to better describe the actual interdisciplinarity of the corresponding journal.

At this point, we have to point to an important limitation of using aggregated datasets as basic units for the analysis of IDR. While individual documents always integrate the knowledge used to conduct the research, the results of which are published in the documents in question, so that the documents reflect a missing, lower or larger extent of interdisciplinarity, the distinction between interdisciplinarity and multi-disciplinarity may become difficult at higher levels of aggregations, for instance, in the case of journals, university faculties, research institutes since, at this level, knowledge from specialized “mono-disciplinary” research may be juxtaposed without integration according to the possibly multidisciplinary profile of the unit (journal or research institutions). The two arguments (conceptual and quantitative) clearly support the use of individual publications. However, it doesn’t mean it is meaningless to explore the interdisciplinarity in the aggregation level: it can be applied to track whether the journals publish diverse topics or to analyze whether the interdisciplinary teams conduct 'multidisciplinary' studies rather than IDR, etc. For example, *J. Informetrics* and *Scientometrics* have similar interdisciplinarity in the average individual level, but *Scientometrics* show more multidisciplinary characteristics for publishing diverse topics. Below, we will have a closer look at the correlation of IDR measures at different levels of aggregation.



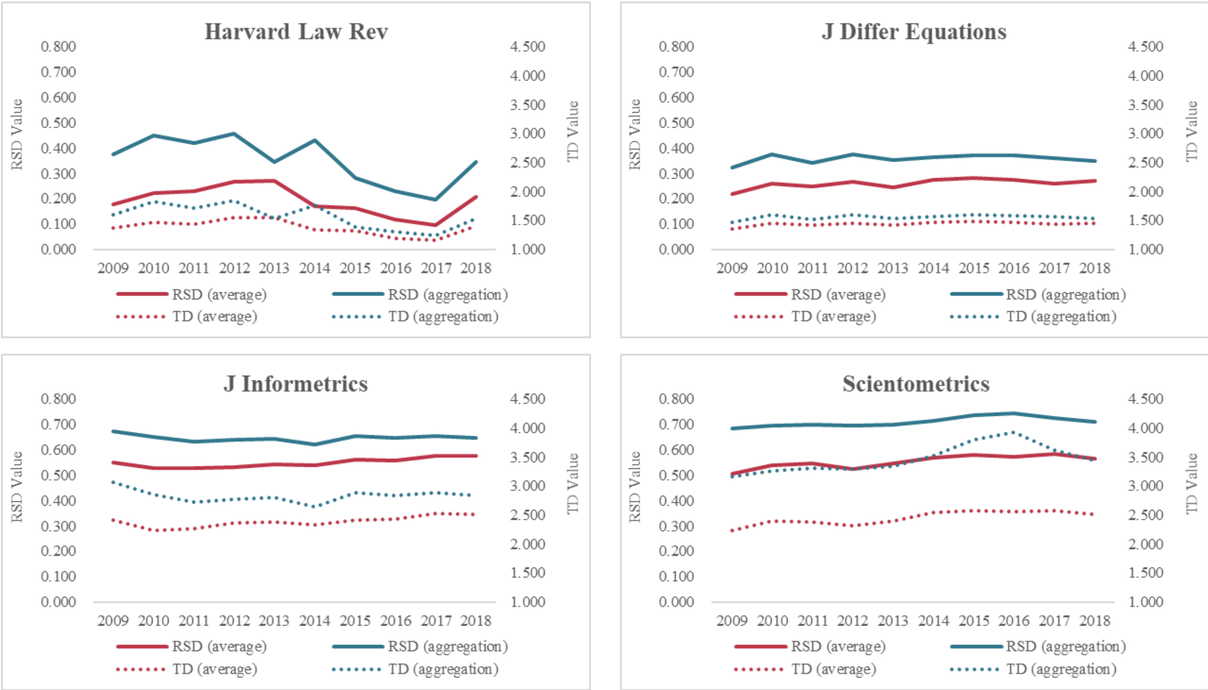


Figure 11. The average value of interdisciplinarity scores for different levels of aggregation of the eight journals

When we come to the Spearman correlations of the two analysis objects, Table 3 indicates that the average value of interdisciplinarity at the level of individual articles show significant but medium correlations to the interdisciplinarity on the aggregated level. At the same time, RSD and TD are highly correlated, while the high discriminatory power may make TD a better indicator to measure diversity. i.e., interdisciplinarity.

Table 3. Spearman correlations of interdisciplinarity based on two different levels of aggregation

	Variates	1	2	3
1	RS_Diversity (average)			
2	True_Diversity (average)	.991**		
3	RS_Diversity (aggregation)	.765**	.807**	
4	True_Diversity (aggregation)	.765**	.807**	1.000**

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

Conclusions and Discussion

Based on the dataset and corresponding indicators, we further build Spearman correlations among the set of variates, shown in Table 4. Overall, the correlations among these variates are high and significant. Based on detailed analyses, we propose the following ideas to conclude our preliminary results. First, a finer classification system with more subject categories increases the possibility that one cites sources from different subject categories. We can select them according to the subject integration we want to observe or make a combination to track the integration across the subjects. Second, different normalization approaches may lead to obviously different interdisciplinarity results, and such a finding is supported by the relatively low correlations between the interdisciplinarities calculated by Salton's Cosine and Ochiai's Cosine. Third, on the basis of Salton's cosine normalization, the interdisciplinary values obtained by different settings are highly correlated, especially in terms of different citation similarity matrices (cited, citing and cross-citation) and, in general, with different time windows. In other words, using different settings can definitely produce different values, but the distribution of interdisciplinarity is almost the same. Fourth, results based on an aggregated dataset tend to overly expand the 'interdisciplinarity' degree of a journal, especially when the focused journal is actually 'multidisciplinary'. The average interdisciplinarity of individual publications seems to better describe the actual interdisciplinarity of the corresponding journal, but the aggregation approach also provides the possibility to help judge whether a journal is a multidisciplinary journal or not by combing with other indicators.

Table 4. Spearman correlations among various variates

	Variates	1	2	3	4	5	6	7	8
1	DISC_TD								
2	Field_TD	.923**							
3	Pearson_TD	.879**	.814**						
4	Ochiai_TD	.573**	.507**	.609**					
5	Cited_TD	.850**	.776**	.974**	.734**				
6	Citing_TD	.859**	.789**	.973**	.739**	.996**			
7	WC2009_TD	.853**	.793**	.986**	.589**	.963**	.955**		
8	WC2018_TD	.861**	.802**	.995**	.568**	.962**	.959**	.983**	
9	Baseline (Salton, WC, Cross_2009-2018)	.873**	.808**	.999**	.594**	.972**	.969**	.988**	.997**

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

When we intend to measure the interdisciplinarity by using integrated indicators, such as RSD and TD, which include variety, balance, and disparity, it is important to choose the similarity measurement approach based on the research scope and research question; different settings may lead to different results. We think this research is beneficial to help explore the study of interdisciplinarity, not only the knowledge integrations from the bibliographic references (Wang et al., 2017) and knowledge diffusion from the citing article cohorts (Carley & Porter, 2012), but also the collaboration diversity from the affiliation information (Abramo et al., 2018; Zhang et al., 2018), and topics diversity from the semantic characteristics (Bu et al., 2021; Hackett et al., 2021). Furthermore, it can also contribute to various fields, including technology convergence (Gong & Keller, 2003), multi-program funding (Huang et al., 2016; Nichols, 2014), as well as science mapping visualization methods (Huang et al., 2021; Leydesdorff, Carley, et al., 2013; Rafols et al., 2010). These interests depend on measures of the similarity among different subject categories to study the strength of interdisciplinary integration.

Furthermore, most of the current interdisciplinary indicators rely on the defined discipline scheme but seldomly based on the fine-grained publication-level clustering, e.g., the work conducted by (Waltman & van Eck, 2012). In addition, the matrices used in this paper are extracted from the direct citation, and the matrices using cocitation and bibliographic coupling remains to be further explored.

Acknowledgments

The research underlying this study is done within the framework of the project “Interdisciplinarity & Impact” (2019-2023) funded by the Flemish Government. The authors also would like to acknowledge support from the National Natural Science Foundation of China (Grant Nos. 72004169;

71974150;71573085), the National Social Science Foundation of China (Grant No. 18VSJ087) and the National Laboratory Center for Library and Information Science in Wuhan University. The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the supporters.

Appendix

Table A1. The statistics of the interdisciplinary indicators by different similarity matrices

NO	Index	N (Valid)	Mean	Median	Std. Deviation	Variance	Skewness	Kurtosis	Min	Max
1	DISC_RSD	37627	0.371	0.378	0.143	0.021	-0.242	-0.376	0.000	0.772
2	DISC_TD	37627	1.678	1.607	0.412	0.170	0.971	1.253	1.000	4.389
3	Field_RSD	37627	0.304	0.303	0.156	0.024	-0.122	-0.863	0.000	0.676
4	Field_TD	37627	1.515	1.436	0.362	0.131	0.694	-0.153	1.000	3.089
5	Pearson_RSD	37627	0.419	0.421	0.168	0.028	-0.222	-0.361	0.000	0.872
6	Pearson_TD	37627	1.895	1.728	0.650	0.422	1.549	3.718	1.000	7.833
7	Ochiai_RSD	37627	0.716	0.769	0.172	0.030	-2.105	4.715	0.000	0.953
8	Ochiai_TD	37627	4.462	4.333	1.948	3.795	0.789	1.882	1.000	21.134
9	Cited_RSD	37627	0.523	0.536	0.151	0.023	-0.828	1.179	0.000	0.905
10	Cited_TD	37627	2.318	2.154	0.788	0.620	1.395	3.594	1.000	10.498
11	Citing_RSD	37627	0.524	0.538	0.152	0.023	-0.852	1.161	0.000	0.904
12	Citing_TD	37627	2.323	2.165	0.785	0.616	1.347	3.438	1.000	10.432
13	WC2009_RSD	37627	0.407	0.406	0.167	0.028	-0.169	-0.539	0.000	0.853
14	WC2009_TD	37627	1.843	1.682	0.596	0.355	1.257	2.034	1.000	6.822
15	WC2018_RSD	37627	0.377	0.374	0.157	0.025	-0.086	-0.508	0.000	0.837
16	WC2018_TD	37627	1.726	1.597	0.504	0.254	1.338	2.607	1.000	6.147
17	Baseline (Salton, WC, Cross, 2009-2018)_RSD	37627	0.394	0.394	0.160	0.026	-0.164	-0.392	0.000	0.851
18	Baseline (Salton, WC, Cross, 2009-2018)_TD	37627	1.786	1.650	0.548	0.301	1.405	2.986	1.000	6.709

References

Abramo, G., D'Angelo, C. A., & Zhang, L. (2018). A comparison of two approaches for measuring interdisciplinary research output: The disciplinary diversity of authors vs the disciplinary diversity of the reference list. *Journal of Informetrics, 12*(4), 1182-1193. https://doi.org/10.1016/j.joi.2018.09.001

Adnani, H., Cherraj, M., & Bouabid, H. (2020). Similarity indexes for scientometric research: A comparative analysis [Article]. *Malaysian Journal of Library & Information Science, 25*(3), 31-48. https://doi.org/10.22452/mjlis.vol25no3.3

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology, 54*(6), 550-560. https://doi.org/10.1002/asi.10242

Bu, Y., Li, M., Gu, W., & Huang, W.-b. (2021). Topic diversity: A discipline scheme-free diversity measurement for journals. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24433

Carley, S., & Porter, A. L. (2012). A forward diversity index. *Scientometrics, 90*(2), 407-427. https://doi.org/10.1007/s11192-011-0528-1

Carley, S., Porter, A. L., Rafols, I., & Leydesdorff, L. (2017). Visualization of disciplinary profiles: Enhanced science overlay maps. *Journal of Data and Information Science, 2*(3), 68-111. https://doi.org/10.1515/jdis-2017-0015

Clarivate Analytics. (2019). *Web of Science: Master Journal List*. Retrieved Jan. 9 from https://support.clarivate.com/ScientificandAcademicResearch/s/article/Web-of-Science-Master-Journal-List

[Record #730 is using a reference type undefined in this output style.]

Glänzel, W., Huang, Y., & Thijs, B. (2021). *Improving the precision of disparity measurement in studies of interdisciplinary research* The 18th International conference on Scientometrics & Informetrics, Leuven.

Glänzel, W., & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics, 56*(3), 357-367. https://doi.org/10.1023/A:1022378804087

Glänzel, W., Thijs, B., & Chi, P.-S. (2016). The challenges to expand bibliometric studies from periodical literature to monographic literature with a new data source: the book citation index. *Scientometrics, 109*(3), 2165-2179. https://doi.org/10.1007/s11192-016-2046-7

Gong, G., & Keller, W. (2003). Convergence and polarization in global income levels: a review of recent results on the role of international technology diffusion. *Research Policy, 32*(6), 1055-1079. https://doi.org/10.1016/S0048-7333(02)00136-1

Hackett, E. J., Leahey, E., Parker, J. N., Rafols, I., Hampton, S. E., Corte, U., Chavarro, D., Drake, J. M., Penders, B., Sheble, L., Vermeulen, N., & Vision, T. J. (2021). Do synthesis centers synthesize? A semantic analysis of topical diversity in research. *Research Policy, 50*(1), 104069. https://doi.org/10.1016/j.respol.2020.104069

Huang, Y., Glänzel, W., & Zhang, L. (2021). Tracing the development of mapping knowledge domains. *Scientometrics*. https://doi.org/10.1007/s11192-020-03821-x

Huang, Y., Zhang, Y., Youtie, J., Porter, A. L., & Wang, X. (2016). How Does National Scientific Funding Support Emerging Interdisciplinary Research: A Comparison Study of Big Data Research in the US and China. *Plos One, 11*(5), e0154509. https://doi.org/10.1371/journal.pone.0154509

Huang, Y., Zhu, D. H., Lv, Q., Porter, A. L., Robinson, D. K. R., & Wang, X. F. (2017). Early insights on the Emerging Sources Citation Index (ESCI): an overlay map-based bibliometric study. *Scientometrics, 111*(3), 2041-2057. https://doi.org/10.1007/s11192-017-2349-3

Jost, L. (2006). Entropy and diversity. *Oikos, 113*(2), 363-375. https://doi.org/10.1111/j.2006.0030-1299.14714.x

Klavans, R., & Boyack, K. W. (2006). Identifying a better measure of relatedness for mapping science. *Journal of the American Society for Information Science and Technology, 57*(2), 251-263. https://doi.org/10.1002/asi.20274

Leinster, T., & Cobbold, C. A. (2012). Measuring diversity: The importance of species similarity. *Ecology, 93*(3), 477-489. https://doi.org/10.1890/10-2402.1

Leydesdorff, L. (2008). On the normalization and visualization of author cocitation data: Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology, 59*(1), 77-85. https://doi.org/10.1002/asi.20732

Leydesdorff, L. (2018). Diversity and interdisciplinarity: how can one distinguish and recombine disparity, variety, and balance? [Article]. *Scientometrics, 116*(3), 2113-2121. https://doi.org/10.1007/s11192-018-2810-y

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics, 94*(2), 589-593. https://doi.org/10.1007/s11192-012-0784-8

Leydesdorff, L., & Ivanova, I. (2021). The measurement of "interdisciplinarity" and "synergy" in scientific and extra-scientific collaborations. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24416

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology, 60*(2), 348-362. https://doi.org/10.1002/asi.20967

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal–journal citations. *Journal of the American Society for Information Science and Technology, 64*(12), 2573-2586. https://doi.org/10.1002/asi.22946

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. *Journal of Informetrics, 13*(1), 255-269. https://doi.org/https://doi.org/10.1016/j.joi.2018.12.006

McCain, K. W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science, 41*(6), 433-443. https://doi.org/https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<433::AID-ASI11>3.0.CO;2-Q

Nichols, L. G. (2014). A topic model approach to measuring interdisciplinarity at the National Science Foundation. *Scientometrics, 100*(3), 741-754. https://doi.org/10.1007/s11192-014-1319-2

Ochiai, A. (1957). Zoogeographical Studies on the Soleoid Fishes Found in Japan and its Neighbouring Regions-III. *Nippon Suisan Gakkaishi, 22*, 522-525.

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics, 81*(3), 719. https://doi.org/10.1007/s11192-008-2197-2

Porter, A. L., Roessner, J. D., Cohen, A. S., & Perreault, M. (2006). Interdisciplinary research: meaning, metrics and nurture. *Research Evaluation, 15*(3), 187-195. https://doi.org/10.3152/147154406781775841

Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics, 82*(2), 263-287. https://doi.org/10.1007/s11192-009-0041-y

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology, 61*(9), 1871-1887. https://doi.org/10.1002/asi.21368

Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology, 21*(1), 24-43. https://doi.org/10.1016/0040-5809(82)90004-1

Rao, C. R. (1982). Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhyā: The Indian Journal of Statistics, Series A, 44*(1), 1-22.

Rousseau, R. (2019). On the Leydesdorff-Wagner-Bornmann proposal for diversity measurement. *Journal of Informetrics, 13*(3), 906-907. https://doi.org/https://doi.org/10.1016/j.joi.2019.03.015

Salton, G., & McGill, M. J. (1986). *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc.

Schneider, J. W., & Borlund, P. (2007). Matrix comparison, Part 2: Measuring the resemblance between proximity measures or ordination results by use of the mantel and procrustes statistics [https://doi.org/10.1002/asi.20642]. *Journal of the American Society for Information Science and Technology, 58*(11), 1596-1609. https://doi.org/https://doi.org/10.1002/asi.20642

Stirling, A. (1998). On the economics and analysis of diversity. *SPRU Electronic Working Paper Series, 28*.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface, 4*(15), 707-719. https://doi.org/10.1098/rsif.2007.0213

van Eck, N. J., Waltman, L., van den Berg, J., & Kaymak, U. (2006). Visualizing the computational intelligence field [Article]. *Ieee Computational Intelligence Magazine, 1*(4), 6-10. https://doi.org/10.1109/ci-m.2006.248043

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. https://doi.org/10.1002/asi.22748

Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and Impact: Distinct Effects of Variety, Balance, and Disparity. *Plos One, 10*(5), e0127298. https://doi.org/10.1371/journal.pone.0127298

Wang, Q., & Schneider, J. W. (2019). Consistency and validity of interdisciplinarity measures. *Quantitative Science Studies, 1*(1), 239-263. https://doi.org/10.1162/qss_a_00011

Wang, X., Wang, Z., Huang, Y., Chen, Y., Zhang, Y., Ren, H., Li, R., & Pang, J. (2017). Measuring interdisciplinarity of a research system: detecting distinction between publication categories and citation categories. *Scientometrics, 111*(3), 2023-2039. https://doi.org/10.1007/s11192-017-2348-4

Zhang, L., Glänzel, W., & Liang, L. (2009). Tracing the role of individual journals in a cross-citation network based on different indicators. *Scientometrics, 81*(3), 821-838. https://doi.org/10.1007/s11192-008-2245-y

Zhang, L., Janssens, F., Liang, L., & Glänzel, W. (2010). Journal cross-citation analysis for validation and improvement of journal-based subject classification in bibliometric research. *Scientometrics, 82*(3), 687-706. https://doi.org/10.1007/s11192-010-0180-1

Zhang, L., Rousseau, R., & Glänzel, W. (2016). Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. *Journal of the Association for Information Science and Technology, 67*(5), 1257-1265. https://doi.org/10.1002/asi.23487

Zhang, L., Sun, B., Chinchilla-Rodríguez, Z., Chen, L., & Huang, Y. (2018). Interdisciplinarity and collaboration: on the relationship between disciplinary diversity in departmental affiliations and reference lists. *Scientometrics, 117*(1), 271-291. https://doi.org/10.1007/s11192-018-2853-0

Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science. *Scientometrics, 47*(3), 627-657. https://doi.org/10.1023/A:1005632319799

MANAGEMENT, STRATEGY AND INNOVATION (MSI)
Naamsestraat 69 bus 3535
3000 LEUVEN, Belgium
tel. + 32 16 32 67 00
msi@econ.kuleuven.be
https://feb.kuleuven.be/research/MSI/