

# How similar are field-normalized scores from different free or commercial databases calculated for large German universities?

Thomas Scheidsteger <sup>\*</sup>, Robin Haunschild <sup>\*</sup>, and Lutz Bornmann <sup>\*,\*\*</sup>

<sup>\*</sup>*t.scheidsteger@fkf.mpg.de; r.haunschild@fkf.mpg.de; l.bornmann@fkf.mpg.de*  
0000-0001-8351-2498; 0000-0001-7025-7256; 0000-0003-0810-7091  
IVS-CPT, Max Planck Institute for Solid State Research, Stuttgart, Germany

<sup>\*\*</sup>*lutz.bornmann@gv.mpg.de*  
0000-0003-0810-7091

Science Policy and Strategy Department, Administrative Headquarters of the Max Planck Society, Germany

We compared field-normalized citation scores from the freely available bibliographic database OpenAlex with those from three commercial databases (Web of Science, Scopus, and Dimensions). The scores have been calculated for nearly 335,000 publications published by 48 German universities in four OECD subject areas between 2013 and 2017. We found varying but overall strong agreement according to Lin's concordance coefficient. We also calculated (aggregated) mean normalized citation scores for the 48 universities and found that the agreements across different databases are low. On the one hand, the results suggest that comparisons of universities using field-normalized citation scores across different databases should be avoided. On the other hand, the difference of the concordance coefficients on the paper and university level is a good example for the problem of ecological fallacy in bibliometrics: The mean impact is not representative for the single papers' impact in the set.

## 1. Introduction

Research evaluation using bibliometric methods is frequently based on commercial bibliographic databases that have similar approaches to select journals for including papers in the database and to ensure the quality of included papers: Web of Science (WoS) and Scopus (Baas, Schotten, Plume, et al., 2020; Birkle, Pendlebury, Schnell, et al., 2020). Dimensions is also a commercial bibliographic database which provides an alternative to WoS and Scopus including many more publications (Herzog, Hook, & Konkiel, 2020). With the emergence of Microsoft Academic Graph (MAG) in 2015, a free bibliographic database with an outstanding coverage (Sinha, Shen, Song, et al., 2015; Wang, Shen, Huang, et al., 2020) emerged. Since Microsoft decided to discontinue MAG, the successor database OpenAlex was started by Priem, Piwowar, and Orr (2022). With the many databases that are available in principle for research evaluation purposes, the question arose whether all databases come to similar results in a certain research evaluation situation.

In a first case study, Scheidsteger, Haunschild, Hug, et al. (2018) analyzed the publications of a computer science institute with a well-maintained publication list. They chose a bibliometric standard indicator (a field-normalized indicator) and tested whether the indicator scores are similar across two different databases. Thus, they investigated the convergent validity of field-normalized indicator scores that have been generated based on MAG and WoS data. The results were encouraging (i.e., the values were in a good agreement) and motivated the present study with a significantly enlarged publication set from 48 German Universities that cover a broad range of subject areas (and not only computer science, as in the first case study). Field-normalized citation scores were calculated based on data from four different databases – three commercial databases and OpenAlex. In this follow-up study of the case study by Scheidsteger, et al. (2018), we are interested in the convergent validity of the field-normalized scores from

the different databases: do we receive the same or similar field-normalized scores when the same indicator is used or not?

## 2. Data and Methods

### 2.1. Selection of data sources

As sources of bibliometric data we used the above mentioned three commercial databases and the free database OpenAlex. The WoS data had been released in October 2021 and the Scopus data in April 2021. From Dimensions we used a data dump from January 2022 and from OpenAlex a snapshot from February 2022.

### 2.2. Field-normalized citation scores

For the comparison of field-normalized scores across four databases, we used the normalized citation score (NCS) (Waltman, van Eck, van Leeuwen, et al., 2011). It is one of the most popular approaches to field-normalize citation counts (van Wijk & Costas-Comesaña, 2012). In principle, any other field-normalized indicator could have been used in this study such as percentiles (Bornmann & Williams, 2020). The NCS is calculated as follows: the citation count of each paper is divided by the average citation count of similar papers (i.e. the reference set). Similar papers are usually defined as papers from the same field, publication year, and document type. The NCS is formally defined as

$$NCS_i = \frac{c_i}{e_i}$$

with  $c_i$  denoting the citation count of a focal paper  $i$  and  $e_i$  denoting the expected citation rate of similar papers (Lundberg, 2007; Rehn, Kronman, & Wadskog, 2007). In many cases, papers in the databases are assigned not only to one, but to multiple fields. In this case, we calculated several NCS values for each paper. To obtain a single NCS for each paper, the multiple NCS values were averaged (Haunschild & Bornmann, 2016).

### 2.3. Subject classifications

The expected citation rates for the NCS were calculated based on different field categorization schemes in the four databases. In WoS (Birkle, et al., 2020) and Scopus (Baas, et al., 2020), journals are intellectually assigned to 252 *WoS Subject Categories (WoSSC)* and 335 *All Science Journal Classification Codes (ASJC)*, respectively. In the other two databases, subjects are assigned paper-based using different taxonomies and machine learning algorithms. Dimensions (Herzog, et al., 2020) has a two-level hierarchy of *Fields of Research* with 22 main categories and 154 sub categories. OpenAlex has a six-level hierarchy of *concepts* with 19 top-level categories and 284 second-level categories (Scheidsteger & Haunschild, 2023). In the case of Dimensions and OpenAlex, we used the second-level (sub) categories for the field-normalization because of their similar granularity compared to the journal-based schemes. Based on the different field categorization schemes in the databases, we received two groups of NCS values: the (1) scores from a journal-based classification with NCS\_WoS and NCS\_Scopus, and (2) scores from a paper-based classification with NCS\_Dimensions and NCS\_OpenAlex.

### 2.4. Publication set

For the comparison of field-normalized scores, it was necessary to have the same institutional publication set from each database. To reach this goal, we started with the WoS database that includes disambiguated publication data for German universities. We focused on the publication years from 2013 to 2017 (to have citation windows of at least five years), and the

document types article and review (i.e., only substantial publications). We only considered papers in the following OECD subject categories: Natural sciences, Engineering, Medicine, and Social Sciences. In other subject categories, the use of bibliometrics might be questionable. We restricted the publications only to those with DOIs. This focus simplified the collection of a common data set across the four databases and missed only at most 4% of the initial dataset in each publication year. In order to have reliable data across the publication years, we chose the 48 universities that published more than 3,000 papers between 2013 and 2017. The final WoS dataset consisted of 363,020 publications. The match of the WoS data with data from the other databases using the DOI resulted in a common dataset with 334,511 papers. In all databases, citations were counted until the end of 2020.

Of the common dataset, only publications could be considered in the comparisons of NCS values for which a second-level classification had been assigned in Dimensions and OpenAlex. Furthermore, we restricted the dataset to the papers that have at least 10 documents with a mean citation count of at least 1.0 in their reference set as proposed by Haunschild, Schier, and Bornmann (2016). These restrictions led to the publication numbers in Table 1 that were considered in the NCS comparisons of this study.

Table 1. Number and percentage of publications (within the common set of 334.511 DOIs) in the four databases suited for the calculation of field-normalized citation scores

Database	# publications	% publications
WoS	334,385	99.96
Scopus	334,227	99.92
Dimensions	329,709	94.73
OpenAlex	309,716	92.59

## 2.5. Mutual comparisons of databases

With four databases, we could perform six comparisons of NCS values. Additionally, we can either look at all publications at once or at each university separately. As statistical key figures to assess the similarity between two databases, we use two kinds of correlation coefficients that had already been used in the case study by Scheidsteger, et al. (2018): i) Spearman's rank correlation coefficient  $r_s$  (applicable to monotonous relations), and ii) Lin's concordance correlation coefficient  $r_{ccc}$  (Lin, 1989, 2000; Liu, 2016) which measures the degree of agreement (with confidence interval). Additionally, we use as an aggregated indicator the mean normalized citation score MNCS (Waltman, et al., 2011), defined as the average over the NCS values of a whole research unit., e.g., a whole university.

### 3. Results

#### 3.1. Results based on all publications

Table 2 shows the Spearman's rank correlation coefficient  $r_s$  for the six comparisons (and the number of considered publications). The consistently high  $r_s$  of at least 0.88 demonstrate *high correlations* (Cohen, 1988) between NCS values from the databases.

Table 2. Spearman's rank correlation coefficient  $r_s$  (below the diagonal) with respect to NCS values from four databases and numbers of considered publications (above the diagonal).

Database	WoS	Scopus	Dimensions	OpenAlex
WoS	<i>1</i>	334,135	316,809	309.647
Scopus	0.93	<i>1</i>	316.672	309.504
Dimensions	0.88	0.89	<i>1</i>	294.811
OpenAlex	0.88	0.88	0.91	<i>1</i>

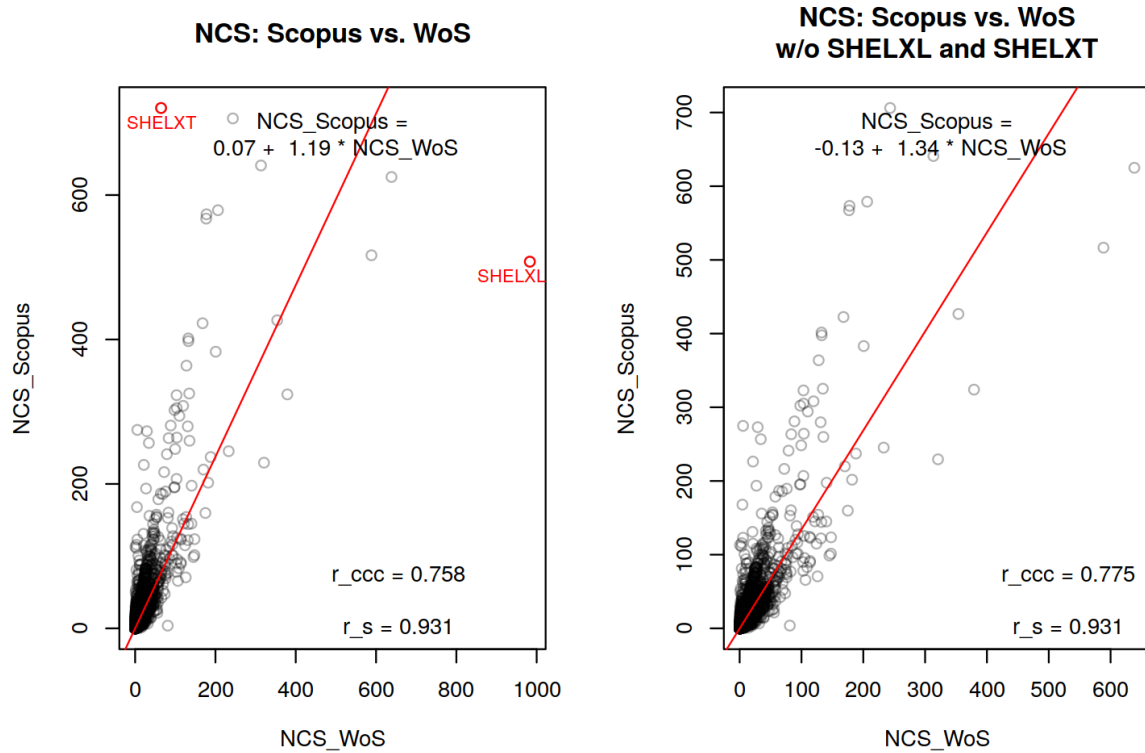
Table 3 displays Lin's concordance correlation coefficient  $r_{ccc}$  for the comparisons together with the associated confidence intervals (confidence level 95%). According to Koch and Spörl (2007), values of  $r_{ccc}$  between 0.8 and 1.0 mean an *almost complete agreement* which is only reached by the comparisons of Dimensions with OpenAlex and Scopus, respectively. The other comparisons reach values between 0.6 and 0.8 pointing to a *strong agreement*.

Table 3. Lin's concordance correlation coefficient  $r_{ccc}$  (below the diagonal) together with the respective confidence interval (above the diagonal) with respect to NCS values from four databases ( $r_{ccc}$  values higher than 0.8 are printed in bold).

Database	WoS	Scopus	Dimensions	OpenAlex
WoS	<i>1</i>	[0.7567; 0.7591]	[0.7703; 0.7723]	[0.6769; 0.6793]
Scopus	0.758	<i>1</i>	[0.8537; 0.8555]	[0.7842; 0.7866]
Dimensions	0.771	<b>0.855</b>	<i>1</i>	[0.8756; 0.8772]
OpenAlex	0.678	0.785	<b>0.876</b>	<i>1</i>

Scatter plots allow graphical assessments of comparisons between NCS values from different databases. As an example, Figure 1 shows two scatter plots for the comparison of Scopus with WoS. On the left, all documents are included. The outcomes of a linear regression and the correlation coefficients are also displayed. Two example outlier papers with the highest NCS values in either of the databases WoS and Scopus and with very different NCS values in the respective other database are marked with red dots and labels. These are papers on numerical methods and software tools, a genre that often reaches very high citation counts, and labelled with their abbreviated names: SHELXL (Sheldrick, 2015a) has in WoS an NCS of 983 (*WoSSCs* Chemistry, Multidisciplinary; Crystallography) and in Scopus an NCS of 508 (*ASJC*: Physical and Theoretical Chemistry; Inorganic Chemistry; Condensed Matter Physics; Materials Chemistry) whereas SHELXT (Sheldrick, 2015b) has an NCS of only 65 in WoS with the same *WoSSCs* as SHELXL, but a very different NCS of 721 and probably related a different *ASJC* with Structural Biology. On the right side, these two papers are excluded which changes the slope of the linear regression as well as Lin's concordance coefficients.

Figure 1: Scatter plot of the NCS values of Scopus vs. WoS *with* two outlier papers marked in red and *without* them.



### 3.2. Analysis of outlier effects

The results for SHELXT and SHELXL point out that outliers may have a significant influence on the correlation between NCS values from different databases. In order to curtail a possible distorting effect of outliers on the correlation coefficients, we compared the datasets excluding outliers. In this study, we defined outliers as papers with NCS values among the top 1% in either database.

Spearman's correlation coefficients are very similar but Lin's concordance coefficients changed more pronounced. The differences are indicated in Table 4.

Table 4: Effect of removing the top 1% NCS values on Lin's concordance coefficient for the six database comparisons (scores above 0.8 are printed in bold face).

	Scopus vs. WoS	Dimensions vs. WoS	OpenAlex vs. WoS	Dimensions vs. Scopus	OpenAlex vs. Scopus	OpenAlex vs. Dimensions
Lin's $r\_ccc$ for all papers	0.76	0.77	0.68	<b>0.85</b>	0.79	<b>0.88</b>
Lin's $r\_ccc$ without top 1% papers	<b>0.87</b>	0.70	0.62	0.78	0.72	<b>0.86</b>
Change in Lin's $r\_ccc$	0.11	-0.07	-0.06	-0.08	-0.07	-0.02

Lin's concordance coefficient increases by 0.1 for the comparison of WoS and Scopus, thereby improving the agreement from *strong* to *almost complete*. For the comparison of Dimensions

and OpenAlex, the coefficient decreases slightly by about 0.02, but for the comparison of Dimensions and Scopus the decrease by 0.08 leads to a change from *almost complete* to *strong* agreement. The other three comparisons led to similar results (independent of inclusion or exclusion of outliers).

### 3.3. Results for the 48 German universities separately

The same evaluations on a per-paper basis have been done for all 48 universities separately. Because of many collaborations between German universities the 334,511 DOIs occur 424,267 times in total in the separate evaluations of the universities. At first, we look at the spread of the correlation coefficients collected in Table 5.

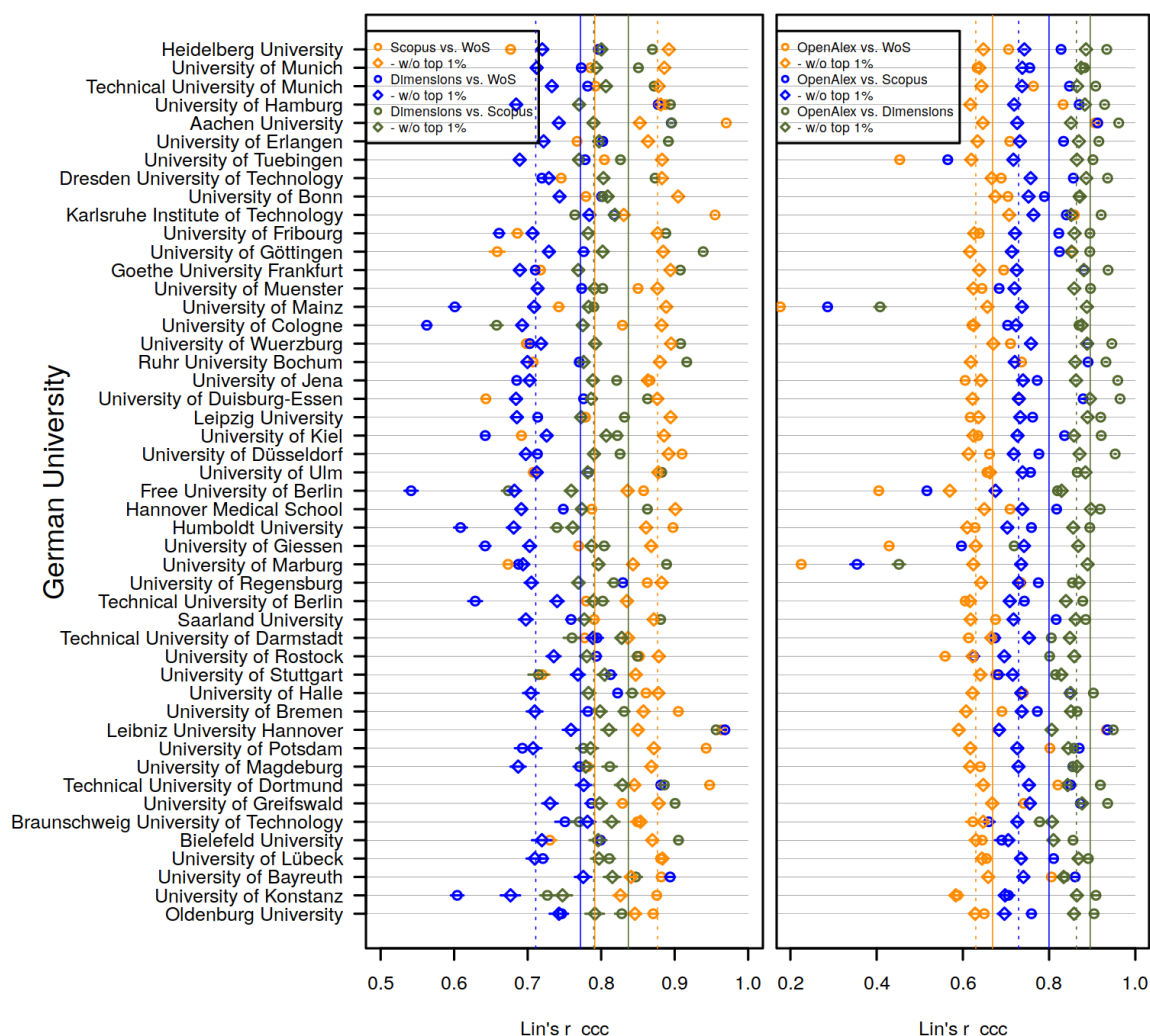
Table 5. Min-max-intervals of the correlation coefficients of the 48 universities separately. Spearman's  $r_s$  is given beneath the diagonal, Lin's  $r_{ccc}$  above the diagonal.

Database	WoS	Scopus	Dimensions	OpenAlex
WoS	1	[0.64; 0.97]	[0.54; 0.97]	[0.18; 0.93]
Scopus	[0.91; 0.95]	1	[0.66; 0.96]	[0.29; 0.94]
Dimensions	[0.86; 0.91]	[0.87; 0.90]	1	[0.41; 0.97]
OpenAlex	[0.86; 0.90]	[0.86; 0.90]	[0.89; 0.92]	1

The values of Spearman's  $r_s$  consistently show a *high* to *perfect correlation*. Lin's  $r_{ccc}$  displays a more diverse behavior. In order to assess the distribution of  $r_{ccc}$  values across the universities and possibly detect outliers, stripcharts are a helpful means. The left stripchart of Figure 4 compares the three commercial databases with one another: In each case the distributions of  $r_{ccc}$  values are relatively homogeneous; moreover, we consistently have a *strong* to *almost complete agreement*, with the exception of only two universities in Dimensions vs. WoS (Free University of Berlin with  $r_{ccc}=0.54$  and University of Cologne with  $r_{ccc}=0.56$ ). The stripchart also displays the distributions of  $r_{ccc}$  with each university's top 1% papers removed. This measure reduced the spread of  $r_{ccc}$  values across the universities drastically. E.g., the comparison Dimensions vs. WoS now has no longer  $r_{ccc}$  values below 0.67 and also none over 0.8. In case of Scopus vs. WoS, *all*  $r_{ccc}$  values can be seen as pointing to an *almost complete agreement*, and for Dimensions vs. Scopus the values only range between 0.75 and 0.83.

The greatest spread of  $r_{ccc}$  values in Table 5 occurs at comparisons including the free database OpenAlex. Comparing OpenAlex with the three commercial ones in the right stripchart of Figure 4 reveals in each case several outliers separated from a more or less homogeneous majority field. Two of them are among the most extreme outliers in each comparison: The University of Mainz and the University of Marburg have very low  $r_{ccc}$  values of about 0.2 (OpenAlex vs. WoS), about 0.3 (OpenAlex vs. Scopus) and about 0.4 (OpenAlex vs. Dimensions). Again, the removal of top 1% papers changes the picture drastically. There are now only three universities with values slightly below 0.6 remaining for OpenAlex vs. WoS: Free University of Berlin with  $r_{ccc}=0.57$ , University of Konstanz with 0.58, and Leibniz University Hannover with 0.59. For OpenAlex vs. Dimension, *all*  $r_{ccc}$  values even point to an *almost complete agreement*.

Figure 4: Lin's  $r_{ccc}$  with confidence intervals for 48 universities (ordered by publication output) in mutual comparisons of the three commercial databases on the left and of OpenAlex with the other ones on the right considering either all documents or without the top 1% papers in each database. Vertical lines indicate the median over all universities with all documents (solid) and without the top 1% papers (dashed line), respectively.



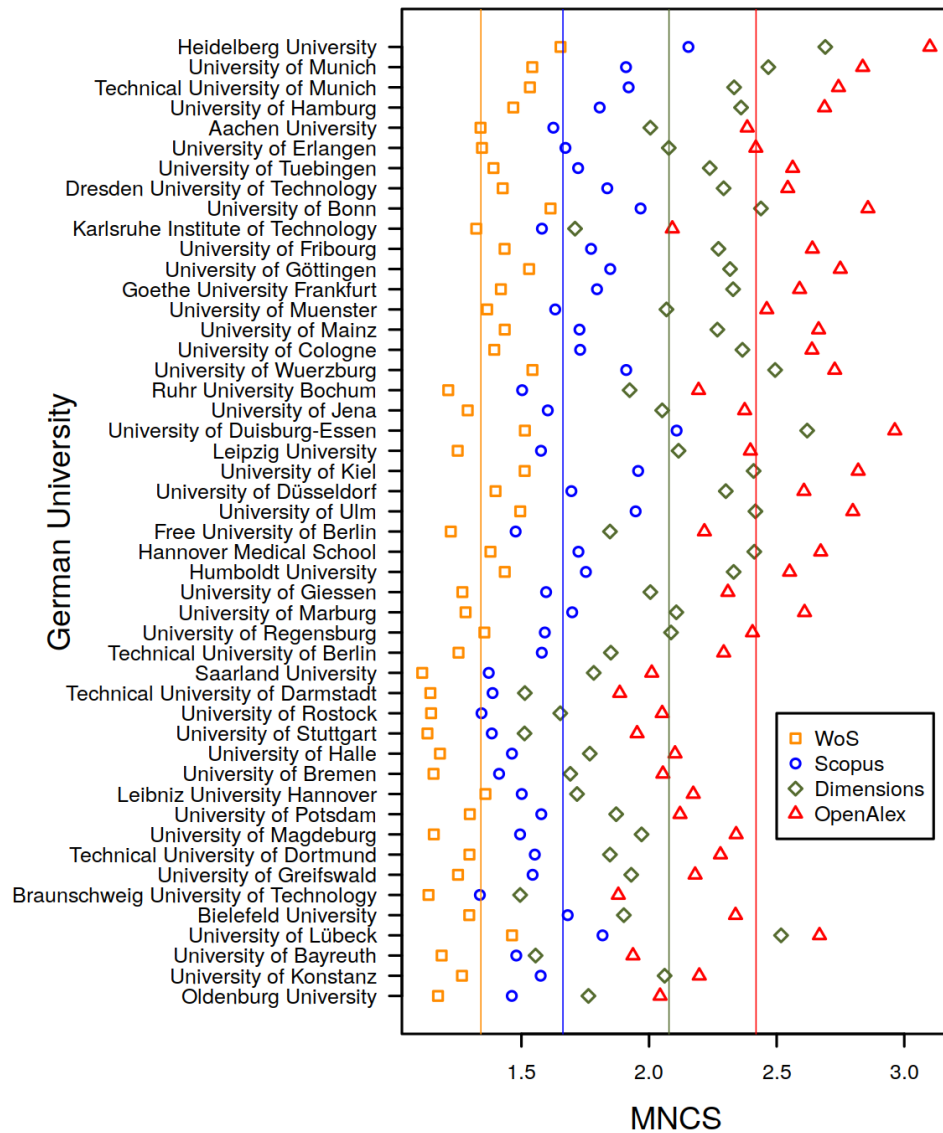
### 3.4. MNCS across universities

In research evaluation, rankings of universities play an important role, usually sorted by the associated MNCS values. Figure 5 shows the respective values for the 48 universities based on data from the four databases. The visual impression of a high correlation is corroborated by Spearman's correlation coefficients listed in Table 6. The highest  $r_s$  values occur between WoS and Scopus resp. OpenAlex and Dimensions. So, viewed through the MNCS indicator, the differences between the universities are similarly represented in the four databases. However, the concordance coefficients between the MNCS values of the German universities are rather low. Although the concordance coefficients calculated for the single papers are high (they indicate at least a strong agreement), many coefficients for the aggregated values are low.

Table 6: Spearman's  $r_s$  and Lin's  $r_{ccc}$  for the comparison of the MNCS values for 48 German universities between two databases

	Scopus vs. WoS	Dimensions vs. WoS	OpenAlex vs. WoS	Dimensions vs. Scopus	OpenAlex vs. Scopus	OpenAlex vs. Dimensions
<b>Spearman's <math>r_s</math></b>	0.96	0.89	0.92	0.93	0.95	0.96
<b>Lin's <math>r_{ccc}</math></b>	0.32	0.12	0.06	0.37	0.16	0.60

Figure 5: Mean normalized citation scores across the 48 German universities (ordered by publication output) in the four databases. The vertical solid lines represent the respective mean values across the universities.





#### 4. Discussion

In order to answer the research question from the title, we calculated correlation coefficients for ten mutual comparisons of NCS values for a common set of nearly 335,000 publications from 48 German universities in four databases. The results for Lin's concordance correlation coefficient  $r_{ccc}$  show that *all* comparisons reveal *almost complete* but at least *strong agreement*. We tried to assess the distorting effect of outliers by removing the papers within the top 1% of NCS values in the databases. These additional analyses led to a small decrease in most cases with coefficients indicating at least *strong agreement*.

Looking at the 48 German universities separately, we found in nearly all cases a *strong* to *almost complete agreement* between the three commercial databases, but several cases of very low  $r_{ccc}$  values in comparisons of the free database OpenAlex with the three commercial ones. But the removal of the top 1% papers with extreme outliers in most cases led to *strong* or *almost complete agreement*. In terms of the aggregated indicator MNCS, we found a highly correlated representation of inter-university differences between all databases. However, the concordance between the MNCS values of the German universities is rather low. Both results indicate that relative comparisons between different universities are valid only within either of the tested databases. On the one hand, the results reveal that MNCS values which have been calculated using data from different databases should not be compared. On the other hand, the difference of the concordance coefficients on the paper and university level is a good example for the problem of ecological fallacy in bibliometrics: The mean impact is not representative for the single papers' impact in the set.

We conclude that the suitability of OpenAlex for bibliometric evaluations is similar to that of the established commercial databases when problems of distorting effects in the high NCS realm are taken into account. and even more promising given the fact that since the OpenAlex snapshot from January 2022 changes, e.g., of the subject classification, have been implemented (Priem & Piwowar, 2022).

#### Open science practices

As the present study is explicitly a comparison between commercial bibliographic databases and a free one, the data of the first category are per se closed but the data of OpenAlex are freely available for all purposes (especially the data dumps called snapshots, see <https://docs.openalex.org/download-all-data/openalex-snapshot>). The software for evaluation and visualization is based on the open source language R (<https://www.R-project.org>) and can be provided on request.

#### Acknowledgments

The WoS data used in this study were retrieved from an in-house WoS database developed and maintained by the Max Planck Digital Library and derived from the Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (AHCI) provided by Clarivate (Philadelphia, Pennsylvania, USA). The database contains disambiguated and unified address information for German research institutes and universities developed by the I<sup>2</sup>SoS Bibliometrics Team at University of Bielefeld (<https://www.uni-bielefeld.de/einrichtungen/i2sos/bibliometrie/index.xml>) and provided by the German "Kompetenznetzwerk Bibliometrie" (KB, Competence Centre for Bibliometrics, <http://www.bibliometrie.info>). The Scopus data produced by Elsevier were also provided by the KB. We thank Digital Science for providing us with Dimensions data for temporary storage in an in-house PostgreSQL database for use in this research. We thank OurResearch for

preserving and continually building upon the MAG corpus and providing the OpenAlex freely for all purposes.

### Author contributions

T.S.: Conceptualization, Data curation, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing

R.H.: Conceptualization, Methodology, Visualization, Writing – review & editing

L.B.: Conceptualization, Methodology, Writing – review & editing

### Competing interests

The authors declare that they have no competing interests.

### Funding information

The KB database with affiliation information for German research institutions has been sponsored by the German Federal Ministry of Education and Research (BMBF) under the funding code “16WIK2101A”.

### References

- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377-386. doi: [https://doi.org/10.1162/qss\\_a\\_00019](https://doi.org/10.1162/qss_a_00019).
- Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, 1(1), 363-376. doi: [https://doi.org/10.1162/qss\\_a\\_00018](https://doi.org/10.1162/qss_a_00018).
- Bornmann, L., & Williams, R. (2020). An evaluation of percentile measures of citation impact, and a proposal for making them better. *Scientometrics*, 124(2), 1457-1478. doi: <https://doi.org/10.1007/s11192-020-03512-7>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Publishers.
- Haunschild, R., & Bornmann, L. (2016). The proposal of using scaling for calculating field-normalized citation scores. *El Profesional de la información*, 25(1), 1699-2407. doi: <https://doi.org/10.3145/epi.2016.ene.02>.
- Haunschild, R., Schier, H., & Bornmann, L. (2016). Proposal of a minimum constraint for indicators based on means or averages. *Journal of Informetrics*, 10(2), 485-486. doi: <https://doi.org/10.1016/j.joi.2016.03.003>.
- Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies*, 1(1), 387-395. doi: [https://doi.org/10.1162/qss\\_a\\_00020](https://doi.org/10.1162/qss_a_00020).
- Koch, R., & Sporl, E. (2007). Statistical methods for comparison of two measuring procedures and for calibration: Analysis of concordance, correlation and regression in the case of measuring intraocular pressure. *Klinische Monatsblätter Fur Augenheilkunde*, 224(1), 52-57. doi: <https://doi.org/10.1055/s-2006-927278>.
- Lin, L. I. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, 45(1), 255-268. doi: <https://doi.org/10.2307/2532051>.
- Lin, L. I. (2000). A Note on the Concordance Correlation Coefficient. *Biometrics*, 56(1), 324-325. doi: <https://doi.org/10.1111/j.0006-341X.2000.00324.x>.
- Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., Tu, X.M. (2016). Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, 28(2), 6. doi: <https://doi.org/10.11919/j.issn.1002-0829.216045>.

- Lundberg, J. (2007). Lifting the crown - citation z-score. *Journal of Informetrics*, 1(2), 145-154. doi: <https://doi.org/10.1016/j.joi.2006.09.007>.
- Priem, J., & Piwowar, H. (2022). openalex-concept-tagging. Retrieved June 27, 2022, from <https://github.com/ourresearch/openalex-concept-tagging>
- Priem, J., Piwowar, H., & Orr, R. (2022). *OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts*. Paper presented at the 26th International Conference on Science, Technology and Innovation Indicators (STI 2022), Granada, Spain. doi: <https://doi.org/10.5281/zenodo.6936227>.
- Rehn, C., Kronman, U., & Wadskog, D. (2007). *Bibliometric indicators – definitions and usage at Karolinska Institutet*. Stockholm, Sweden: Karolinska Institutet University Library.
- Scheidsteger, T., & Haunschild, R. (2023). Which of the metadata with relevance for bibliometrics are the same and which are different when switching from Microsoft Academic Graph to OpenAlex? *Profesional de la información*, 32(2). doi: <https://doi.org/10.3145/epi.2023.mar.09>.
- Scheidsteger, T., Haunschild, R., Hug, S., & Bornmann, L. (2018). *The concordance of field-normalized scores based on Web of Science and Microsoft Academic data: A case study in computer sciences*. Paper presented at the 23th International Conference on Science, Technology and Innovation Indicators (STI 2018), Leiden. The Netherlands. <https://hdl.handle.net/1887/65358>.
- Sheldrick, G. M. (2015a). Crystal structure refinement with SHELXL. *Acta Crystallographica Section C: Structural Chemistry*, 71(1), 3-8. doi: <https://doi.org/10.1107/S2053229614024218>.
- Sheldrick, G. M. (2015b). SHELXT – Integrated space-group and crystal-structure determination. *Acta Crystallographica Section A: Foundations and Advances*, 71(1), 3-8. doi: <https://doi.org/10.1107/S2053273314026370>.
- Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., Hsu, B.-J. P., & Wang, K. (2015). *An Overview of Microsoft Academic Service (MAS) and Applications*. Paper presented at the 24th International Conference on World Wide Web (WWW '15 Companion), Florence, Italy. doi: <https://doi.org/10.1145/2740908.2742839>.
- van Wijk, E., & Costas-Comesaña, R. (2012). *Bibliometric Study of FWF Austrian Science Fund 2001-2010/11*. Leiden, the Netherlands: Center for Science and Technology Studies (CWTS). URL: <https://zenodo.org/record/17851>.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37-47. doi: <https://doi.org/10.1016/j.joi.2010.08.001>.
- Wang, K., Shen, Z., Huang, C., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft Academic Graph: When experts are not enough. *Quantitative Science Studies*, 1(1), 396-413. doi: [https://doi.org/10.1162/qss\\_a\\_00021](https://doi.org/10.1162/qss_a_00021).