

10° EBBC

Encontro Brasileiro de
Bibliometria e Cientometria



COMPARATIVE METHODOLOGICAL ANALYSIS OF BIBEXCEL AND BIBLIOMETRIX IN SCIENTIFIC CO- AUTHORSHIP NETWORKS: evidence from the Brazilian Synchrotron

César da Silva Robusti

<https://orcid.org/0000-0003-4857-4874>
cesar.robusti@uscsonline.com.br
Universidade Municipal de São Caetano
do Sul (USCS)

Milton Carlos Farina

<https://orcid.org/0000-0003-0551-8282>
milton.farina@online.uscs.edu.br
Universidade Municipal de São Caetano
do Sul (USCS)

Silvio Popadiuk

<https://orcid.org/0000-0003-1089-4928>
silvio.popadiuk@mackenzie.br
Universidade Presbiteriana Mackenzie
(UPM)

Abstract:

This study compares Bibexcel and Bibliometrix in co-authorship network analysis, focusing on transparency, automation, and reproducibility. Using 364 Web of Science articles linked to the Brazilian Synchrotron (LNLS), the research evaluates data extraction, preprocessing, and visualization workflows. Results indicate that Bibexcel provides greater transparency and control, while Bibliometrix offers automation and efficiency. The comparison highlights methodological implications for open science and scientific sovereignty, emphasizing how the choice of tool affects reproducibility, accessibility, and interpretation of collaborative networks.

Keywords: Bibexcel. Bibliometrix. Co-authorship networks. Reproducibility. Bibliometrics.

EIXO TEMÁTICO:

Técnicas, Ferramentas e
Infraestruturas

MODALIDADE:

Resumo expandido

DOI: 10.22477/x.ebbc.1029

COMO CITAR ESTE ARTIGO:

ROBUSTI, César da Silva;
FARINA, Milton Carlos; PO-
PADIUK, Silvio. Comparative
methodological analysis of
Bibexcel and Bibliometrix
in scientific co-authorship
networks: evidence from the
Brazilian Synchrotron. *In*:
EBBC - Encontro Brasileiro de
Bibliometria e Cientometria.
10. **Anais [...]**. Curitiba, 2026.
DOI: 10.22477/x.ebbc.1029.
Disponível em: <https://doi.org/10.22477/x.ebbc.1029>.



1 INTRODUCTION

Scientific production has expanded significantly due to the digitization of information and the growth of international databases. In this context, bibliometrics has become essential for analyzing collaboration patterns and knowledge structures (Donthu et al., 2021; Aria and Cuccurullo, 2017).

Despite the increasing use of computational tools, there is still a lack of systematic comparative studies assessing how differences in transparency, automation, and preprocessing affect reproducibility and interpretation of co-authorship networks. Among the software used for this type of analysis, Bibexcel and the Bibliometrix package stand out.

This study seeks to fill this gap, providing a methodological comparison between the software, contributing to the debate on transparency and scientific sovereignty. This study aims to comparatively evaluate the effects of workflow transparency, automation, and preprocessing differences between Bibexcel and Bibliometrix on the reproducibility and interpretability of co-authorship network analyses.

2 REVIEW OF LITERATURE

Bibliometrics has its roots in authors such as Price (1963), who conceptualized the exponential growth of scientific output and established the quantitative study of science as a field of research, alongside the formulation of empirical laws, including Lotka's law, Bradford's law, and Zipf's law, which describe patterns of productivity, information dispersion, and word frequency, respectively (Machado Junior et al., 2016).

By employing quantitative methods, bibliometrics can be used to analyze scientific

output, enabling the identification of publication patterns, collaboration networks, and the thematic structure of fields of knowledge (Donthu et al., 2021). Recent studies underscore the importance of transparent and reproducible methods, especially in light of the growing volume of scientific data.

In turn, scientific collaboration networks consist of actors and the relationships established among them, in which nodes represent individuals or institutions connected by social ties (Emirbayer and Goodwin, 1994). In this context, co-authorship networks are widely used in scientometric studies to analyze patterns of collaboration (Katz and Martin, 1997). Co-authorship is considered one of the most explicit forms of scientific collaboration and a reliable indicator of knowledge production (Glänzel and Schubert, 2004). However, challenges such as the disambiguation of authors' names can affect the accuracy of network structures, requiring careful data preprocessing (Kang et al., 2009).

In this context, tools such as Bibliometrix and Bibexcel have significantly facilitated the analysis of these empirical laws and other quantitative methods in bibliometrics through integrated, user-friendly computational environments.

Bibexcel, developed by Persson, Danell, and Schneider (2009), was widely used in previous decades, as it allowed for the organization and transformation of bibliographic data into matrices that could be exported to network analysis software such as Gephi or UCINET. Its operation requires greater manual intervention, particularly in the standardization of authors and the construction of co-occurrence matrices.

Bibliometrix, in turn, is an R package that integrates data import, processing, and analysis steps into a single environment (Aria and Cuccurullo, 2017). The Biblioshiny interface makes it easy for researchers without programming knowled-

ge to use, offering various analyses with just a few clicks. From a methodological standpoint, the choice of software impacts the interpretation of results and the reproducibility of the research.

3 METHODOLOGICAL PROCEDURES

This study is characterized as a comparative methodological research between Bibexcel and the Bibliometrix, seeking to analyze advances, limitations, and opportunities of these tools for SNA. In the first stage, a functional exploration of both software packages was carried out, considering their specific capabilities for generating co-authorship networks. Bibliometrix was operated through the Biblioshiny interface, which facilitates visualization, although it requires initialization in the R environment.

In the second stage, an empirical comparison was conducted using the same bibliographic corpus composed of 364 articles indexed in Web of Science (WoS) published between 2016 and 2025, with at least one author affiliated with the Brazilian Synchrotron Light Laboratory (LNLS) and linked to research that used the electron accelerator, that is, the affiliation field was used to search for articles.

Data were exported in plain text format and were used exclusively for methodological purposes, restricted to the generation of SNA. Bibexcel was used for manual preprocessing, including normalization of authors and matrix construction. Bibliometrix was used for automated processing and visualization.

The metrics analyzed included the number of authors, co-authorships, and the overall structure of the networks, allowing for the comparison of convergences and divergences resulting from different data extraction, processing, and representation processes.

4 DATA ANALYSIS

The comparison between Bibexcel and Bibliometrix begins with the data extraction and processing. In Bibliometrix, searches can be performed on WoS, and records are imported and converted in the R environment into a bibliometric data frame, applying automatic normalizations that are not very transparent to the user. Bibexcel, on the other hand, operates as a pre-processing tool: it directly imports the plaintext file, preserves the original structure of the fields, and allows for cleaning, standardization, and manual generation of lists and matrices, which can be exported to Ucinet, Pajek, or Gephi. Its flow is procedural and incremental, offering greater control over each step.

This distinction can be interpreted in light of bibliometric workflow structures, as discussed by Aria and Cuccurullo (2017), who emphasize the importance of integrated and automated pipelines for large-scale analysis.

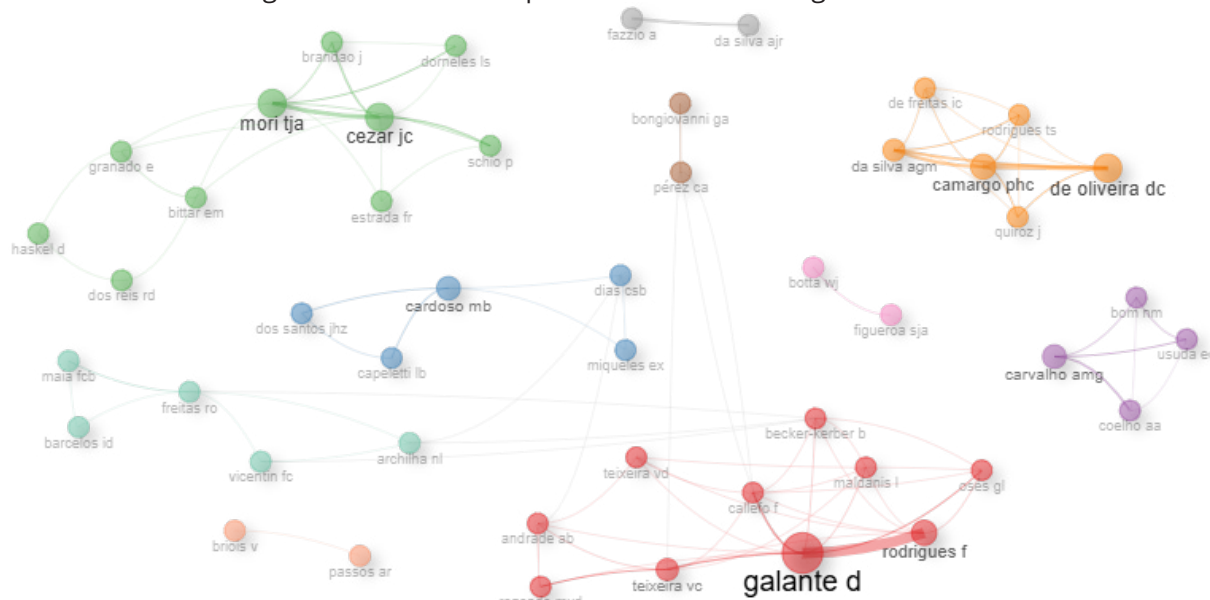
Both software programs use the same initial extraction from WoS, with definition of strings, Boolean operators, time frame, and document type. Bibliometrix accepts .txt, .bib, or EndNote formats from WoS, while Bibexcel works exclusively with .txt. Visualization in Bibliometrix occurs via Biblioshiny and integrations with igraph, operating on already aggregated graphs, in Bibexcel, the researcher manually constructs matrices such as Article $\times$ Author and Author $\times$ Author, ensuring transparency and reproducibility external to the software.

Bibexcel exports matrices to Ucinet (DL format) and lists of nodes and edges to Gephi, allowing for weight adjustments, normalizations (Jaccard, Salton, Cosine), and detailed structural analyses.

For illustrative purposes, the co-authorship network generated in Bibliometrix (Fi-

Figure 1) shows groupings of authors and sizes proportional to productivity.

Figure 1 - Co-authorship network created using Bibliometrix



Source: Drawn up by the authors using Bibliometrix.

Table 1 shows the 15 most central authors, with clusters identified by modularity algorithms and measures of betweenness and closeness. The betweenness measure indicates the role of structural bridge, while the closeness measure reflects the average proximity of an author to the others.

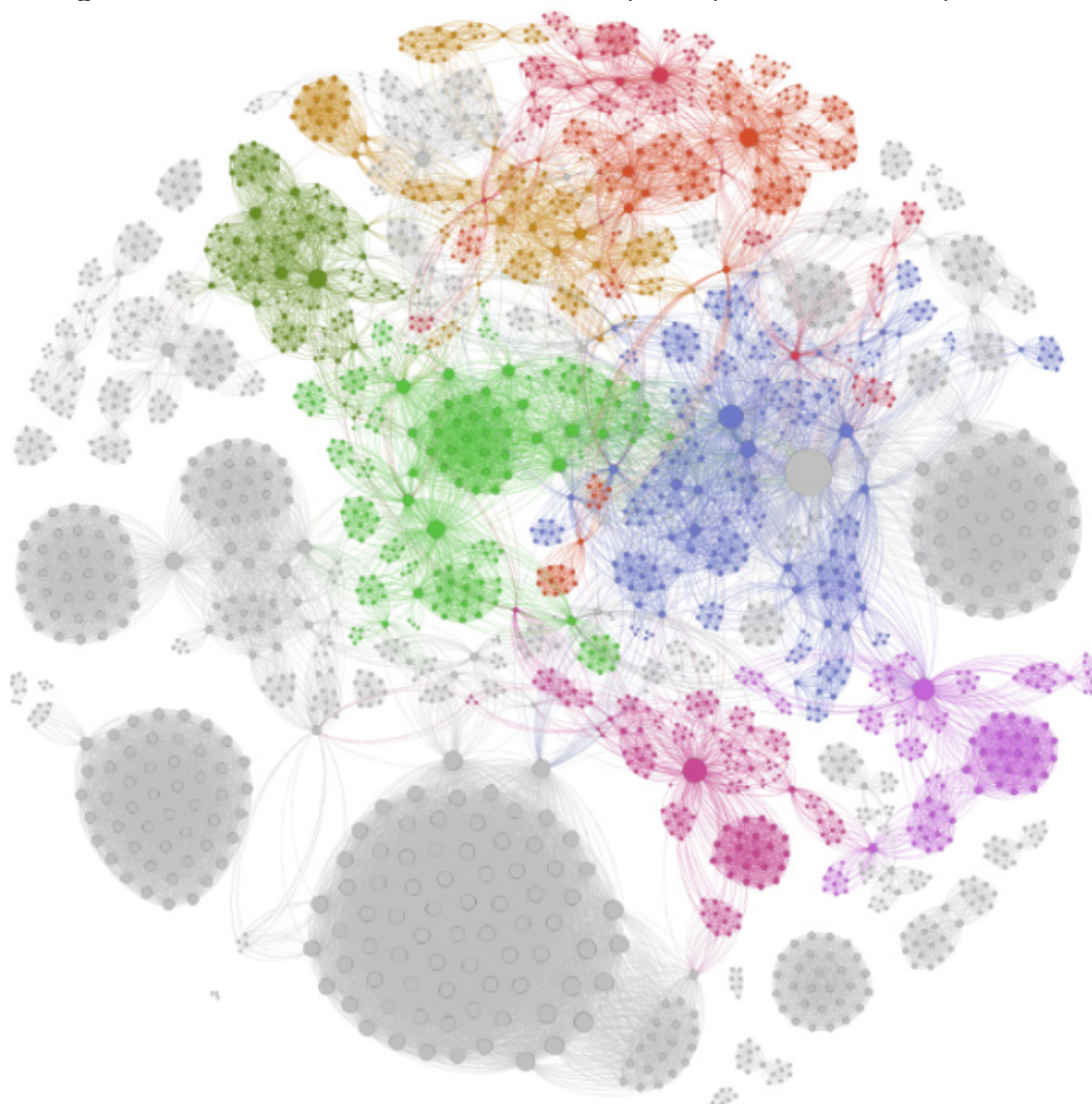
Table 1 - Centrality measures obtained from Bibliometrix

NODE	CLUSTER	BETWEENNESS	CLOSENESS
dias csb	2	69,671	0,021
becker-kerber b	1	56,716	0,024
galante d	1	56,168	0,025
andrade ab	1	47,4	0,022
cardoso mb	2	38	0,016
freitas ro	9	38	0,02
archilha nl	9	28,831	0,022
pérez ca	6	20	0,018
teixeira vc	1	16,436	0,021

Source: authors.

Figure 2, produced with data processed in Bibexcel and visualized in Gephi, shows greater flexibility in the choice of algorithms and parameters, resulting in variations in the number of communities and centrality values. Structural differences between networks, such as edge weighting, inclusion of low-frequency authors, and treatment of homonyms, explain discrepancies between Bibliometrix, Gephi, Ucinet, and Pajek. In addition, each software uses its own algorithmic routines, which can alter betweenness and closeness values even when the underlying matrix is similar.

Figure 2 - Clusters obtained via Bibexcel and Gephi output of co-authorship clusters



Source: Drawn up by the authors using Gephi.

Table 2 shows the output of the Ucinet software with measures of degree centrality, proximity, and intermediation in the network formed by 2,226 authors. The differences with the values in Table 1 do not indicate error, but rather different scales or different normalizations or possible weighting. What matters is the relative position of the author in the network, compared to the others.

Table 2 - Centrality measures obtained from UCINET

AUTOR	DEGREE	CLOSENESS	BETWEENESS
Cardoso Mb	129.000	10.972.000	273.574.594
Rodrigues F	128.000	10.315.000	103.528.477
Pérez Ca	115.000	10.896.000	265.552.094
Mori Tja	97.000	12.517.000	127.574.258
De Oliveira Dc	96.000	11.608.000	183.384.266
Carvalho M	93.000	10.712.000	161.840.094
Freitas Ro	90.000	11.199.000	130.395.555
Cezar Jc	85.000	12.020.000	160.884.469
Callefo F	84.000	10.735.000	106.094.625

Source: authors.

From an epistemological perspective, Bibliometrix can be characterized as a “black-box” system, while Bibexcel promotes transparency and analytical control. This distinction impacts reproducibility and interpretability. These measures are consistent with the use of co-authorship as a proxy for scientific collaboration, as established by Katz and Martin (1997), and further developed by Glänzel and Schubert (2004).

5 CONCLUSIONS

This study demonstrates that Bibexcel and Bibliometrix represent distinct methodological paradigms: transparency versus automation. These differences have

methodological and strategic implications, especially in the context of open science and scientific sovereignty, reinforcing the need for the choice of tool to consider analytical objectives, transparency, and suitability to the research context.

The study contributes by proposing a comparative analytical framework, advancing discussions on transparency and reproducibility, and offering practical guidance for researchers extending the methodological guidelines proposed by Donthu et al. (2021) and demonstrating empirically how different analytical workflows can lead to distinct representations of the same scientific field.

Furthermore, improving transparency in bibliometric analyses supports better public decision-making regarding investments in large-scale scientific infrastructures, such as the LNLS, since it brings to light all the work that has been carried out in this type of infrastructure. This study is limited to WoS data and a single institutional context. Future studies should include multiple databases and comparative cases.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through Research Productivity Grants awarded to Milton Carlos Farina (Grant No. 306453/2025-8) and Silvio Popadiuk (Grant No. 309949/2022-0).

REFERENCES

ARIA, M.; CUCCURULLO, C. bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959–975, 2017. DOI: 10.1016/j.joi.2017.08.007. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157717300500>. Acesso em: 28 de abr. 2026.

DONTHU, N.; KUMAR, S.; MUKHERJEE, D.; PANDEY, N.; LIM, W. M. How to conduct

a bibliometric analysis: an overview and guidelines. **Journal of Business Research**, v. 133, p. 285–296, 2021. DOI: 10.1016/j.jbusres.2021.04.070. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0148296321003155>. Acesso em: 28 de abr. 2026.

EMIRBAYER, M.; GOODWIN, J. Network analysis, culture, and the problem of agency. **American Journal of Sociology**, v. 99, n. 6, p. 1411–1454, 1994. DOI: 10.1086/230450. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/230450>. Acesso em: 30 mar. 2026.

GLÄNZEL, W.; SCHUBERT, A. Analysing scientific networks through co-authorship. In: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U. (ed.). **Handbook of quantitative science and technology research**. Dordrecht: Springer, 2004. DOI: 10.1007/1-4020-2755-9_12. Disponível em: https://link.springer.com/chapter/10.1007/1-4020-2755-9_12. Acesso em: 30 mar. 2026.

KANG, I. S.; NA, S. H.; LEE, S.; JUNG, H.; KIM, P.; SUNG, W. K.; LEE, J. H. On co-authorship for author disambiguation. **Information Processing & Management**, v. 45, n. 1, p. 84–97, 2009. DOI: 10.1016/j.ipm.2008.06.006. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0306457308000907>. Acesso em: 30 mar. 2026.

KATZ, J. S.; MARTIN, B. R. What is research collaboration? **Research Policy**, v. 26, n. 1, p. 1–18, 1997. DOI: 10.1016/S0048-7333(96)00917-1. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0048733396009171>. Acesso em: 30 mar. 2026.

LEYDESDORFF, L. Caveats for the use of citation indicators in research and journal evaluations. **Journal of the American Society for Information Science and Technology**, v. 59, n. 2, p. 278–287, 2008. doi: 10.1002/asi.20743. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/asi.20743>. Acesso em: 01 abr. 2026.

MACHADO JUNIOR, C.; SOUZA, M. T. S. de; PARISOTTO, I. R. dos S.; PALMISANO, A. As leis da bibliometria em diferentes bases de dados científicos. **Revista de Ciências da Administração**, [S. l.], v. 18, n. 44, p. 111–123, 2016. DOI: 10.5007/2175-8077.2016v18n44p111. Disponível em: <https://periodicos.ufsc.br/index.php/adm/article/view/2175-8077.2016v18n44p111>. Acesso em: 30 abr. 2026.

PERSSON, O.; DANELL, R.; SCHNEIDER, J. Wiborg. How to use Bibexcel for various types of bibliometric analysis. In: GLÄNZEL, Wolfgang; SCHUBERT, András; SCHMOCH, Ulrich (ed.). **Celebrating scholarly communication studies: a Festschrift for Olle Persson at his 60th birthday**. Leuven: International Society for Scientometrics and Informetrics, 2009. p. 9–24. Disponível em: <https://lucris.lub.lu.se/ws/portalfiles/portal/5902071/1458992.pdf>. Acesso em: 28 de abr. 2026.



10° EBBC

Encontro Brasileiro de
Bibliometria e Cientometria



PRICE, D. J. de S. **Little science, big science**. New York: Columbia University Press, 1963.