Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Scientometric assessment of research trends in vegetation indices applied to the Cerrado

Lidiane Machado Dionizio
Goiás State University
Emilly Nogueira Alves
Goiás State University
Isadora Lopes Pires
Goiás State University
Jéssica Karina Mesquita Vieira
Goiás State University
Rosivaldo Machado da Silva Júnior
Goiás State University
Vagner Santiago do Vale
vagner.vale@ueg.br
Goiás State University

Research Article

Keywords: Remote sensing, NDVI, bibliometric analysis, environmental monitoring, scientific collaboration, land use change

Posted Date: July 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9829945/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study aims to analyze research trends and scientific production related to environmental monitoring in the Brazilian Cerrado, with an emphasis on the application of remote sensing and vegetation indices such as the Normalized Difference Vegetation Index (NDVI). Despite the increasing number of studies addressing environmental changes in tropical savannas, there is still a lack of integrated analyses that explore the evolution, collaboration networks, and thematic development of this research field. To address this gap, a bibliometric and scientometric analysis was conducted using data collected from international databases, applying mapping and visualization tools to identify publication patterns, key authors, institutions, and research hotspots. The results reveal a significant growth in scientific output over recent years, with strong international collaboration networks and a predominance of studies focused on vegetation dynamics, land use change, and climate variability. Additionally, NDVI-based approaches were identified as central tools for monitoring environmental changes in the Cerrado. The findings contribute to a better understanding of the scientific landscape and highlight emerging research directions, supporting future studies and decision-making processes related to environmental monitoring and conservation in tropical ecosystems.

Introduction

The Cerrado is widely recognized as the most biodiverse tropical savanna in the world and constitutes a strategic biome for Brazil's ecological balance and water security [1]. Despite its significance, the region has experienced a rapid loss of native vegetation due to agricultural expansion and land-use change [2]. In this context, vegetation monitoring is essential. Vegetation indices derived from remote sensing, including NDVI, EVI, LAI, and other multispectral indices, serve as fundamental tools for assessing seasonal dynamics, climate responses, degradation processes, and fire impacts [3].

The growing number of studies on the Cerrado and vegetation indices reflects both technological advances in orbital sensors and increased scientific interest in savanna ecosystem dynamics[4, 5]. However, comprehensive syntheses that elucidate the evolution of this field, identify leading countries and institutions, highlight emerging central themes, and clarify the organization of scientific structures over time remain scarce. Scientometrics provides a strategic approach to address these gaps by enabling the mapping of trends, collaborations, research gaps, and potential directions for future investigations [6].

In scientometric analysis, keywords constitute central elements within knowledge domains, as they synthesize the thematic content of articles and facilitate the identification of conceptual patterns, emerging trends, and relationships among subfields. Mapping and analyzing keyword co-occurrence reveals how ideas, methods, and research objects cluster, enabling the identification of thematic cores and established research lines [7]. Distinct keyword clusters often correspond to studies with similar characteristics that share theoretical or methodological frameworks, indicating potential co-citation, the formation of schools of thought, and the shaping of citation flows within a scientific field [8].

Similarly, the analysis of international collaborations is a fundamental component of scientometric studies, as international cooperation enhances information exchange and strengthens technical capacities. Publications resulting from international collaborations typically achieve greater impact, higher citation rates, and increased global visibility [9]. Evaluating these networks is essential for identifying centers of scientific leadership, asymmetries in knowledge production, structural dependencies, and opportunities to strengthen collaborative ties. This is particularly relevant in fields such as ecology, remote sensing, and biodiversity conservation, where global challenges require integrated efforts.

In complex ecosystems such as the Cerrado, scientometric analysis is particularly relevant because it demonstrates how technological advances, research agendas, funding policies, and global priorities—such as climate change, conservation, and environmental monitoring—influence scientific production over time. In this context, scientometric approaches have contributed to advancing research on vegetation indices in the Cerrado biome by revealing thematic structures and collaboration networks through tools such as VOSviewer [10]. The application of scientometrics in this study not only describes the current state of research but also elucidates its historical drivers, existing gaps, and future development perspectives.

In response to these demands, this study presents the first integrated analysis that combines classical scientometrics (keywords, countries, co-citations), multivariate analysis (ordination and clustering), and temporal statistical analysis (regression-based models). This approach enables a comprehensive understanding of not only the contributors and topics but also the field's structural evolution, the interactions among productivity variables, and the identification of periods marking significant inflection points in research growth.

Material and methods

Search strategies and data selection

Data were collected from the Web of Science (WoS) database, chosen for its extensive coverage and international significance in indexing scientific literature. The following search string was applied:

("cerrado" OR "savanna" OR "savannah") AND ("vegetation index") AND ("tree"). The objective was to identify studies addressing vegetation indices in savanna ecosystems, including the Cerrado biome, with emphasis on arboreal vegetation. The analysis included publications from 1977 to 2025, limited to scientific articles. This search yielded 567 publications.

Records were exported in both .txt (full records) and .xlsx (structured metadata) formats to facilitate bibliometric, scientometric, and network analyses. Metadata fields included title, authors, institutions, journal, year, DOI, ISSN, keywords, abstract, and citation metrics. This multimodal approach enabled flexible data processing across various software platforms.

Before conducting co-occurrence analysis in VOSviewer, all keywords were standardized. Synonymous or morphologically variant terms were consolidated into a single form (e.g., savannah to savanna; leaf area index, LAI to LAI; remote-sensing to remote sensing). After data cleaning and duplicate removal, only unique terms remained. No minimum frequency threshold was imposed, ensuring broad thematic coverage in the co-occurrence maps.

Analyses in vosviewer 1.6.18

Scientometric analysis was performed using VOSviewer (version 1.6.18), a tool developed by [10] and widely used for knowledge network visualization. VOSviewer efficiently processes large bibliographic datasets and offers features for analyzing collaboration networks among authors and countries, reference co-citation, and keyword co-occurrence. Its primary strength lies in generating interactive and intuitive scientific maps.

Data exported from Web of Science (WoS) in ".txt" format were imported into VOSviewer to generate maps representing:

1. international collaboration networks based on production and citations;
2. co-citation relationships among authors and reference works;
3. keyword co-occurrence patterns.

This methodological strategy aligns with established practices in contemporary scientometric research, which recommend the use of network analysis tools to investigate the structure and evolution of scientific fields [6]

Multivariate analysis of scientometric data

Scientometric data were subjected to multivariate analyses to identify research patterns on vegetation indices in Cerrado trees. The dataset was initially prepared in Excel, retaining only the variables relevant to the analysis: number of authors (NA), number of words in the title (NW), authors' keywords (DE), Keywords Plus (ID), number of words in the abstract (AB), number of cited references (NR), total citations across all WoS databases (Z9), and recent usage metrics (U1 and U2). All data were linked to the year of publication to facilitate temporal analyses.

To ensure equal contribution of all variables to the statistical analyses, data were standardized using the Min–Max normalization method. This approach rescaled values to a common 0 to 1 range, preventing distortions caused by variable magnitudes while preserving relative proportions. Subsequent analyses employed PAST software (version 5.2.1), utilizing Euclidean distance as the dissimilarity criterion and Ward's hierarchical clustering method [11].

Before the final PCA, correlation analysis was conducted to detect multicollinearity among variables. An integrated methodological approach was implemented to understand patterns in scientific publications comprehensively. Rigorous variable screening was conducted to eliminate redundancies and improve model robustness. The first step of the statistical procedure involved computing a Pearson correlation matrix at a 5% significance level to assess the degree of linear association among variables (Fig. 2). The Pearson correlation coefficient (r) ranges from − 1 to 1, with values near 0 indicating no linear correlation, values near 1 indicating strong positive correlation, and values near − 1 indicating strong negative correlation.

Three particularly strong correlations were identified in the analysis (Fig. 2): between number of pages (PG) and number of references (NR) with r = 0.81, indicating that longer articles typically include more citations; between citations in the WoS Core (TC) and citations in all databases (Z9) with r = 1.00, demonstrating a perfect linear relationship; and between combined keywords (ID + DE) and Keywords Plus (ID) with r = 0.94, indicating near redundancy between these variables.

Based on these results (Fig. 2), the variables PG, TC, and DE + ID were excluded from subsequent multivariate analyses. After this exclusion, ID and AB (number of words in the abstract) still exhibited overlap (r = 0.72). Consequently, ID was removed because AB reflects the authors' precise language regarding the Cerrado, while ID contains more generic and less relevant terms. Additionally, ID content is partially represented in DE (authors' keywords), resulting in informational redundancy. After removing highly collinear variables, NR, Z9, and AB showed acceptable levels of correlation and were maintained for subsequent multivariate analyses. This decision was based on the principle that highly correlated variables provide redundant information to the statistical model. Retaining only one variable preserved essential information and eliminated redundancies that could distort the final results.

The analytical process integrated three complementary techniques: hierarchical clustering analysis, which used dendrograms to reveal temporal relationships among publication years and scientometric variables; principal component analysis (PCA), which reduced data dimensionality and identified variable combinations explaining variance [11] and regression models for variables with R^2 g > 0.60, which detected significant relationships between periods and scientific indicators.

Simple regressions were also conducted between the annual number of published articles and the number of citations since 2013 across all databases. The same analysis was repeated, excluding the years 2025, 2024, and 2023 to mitigate citation-delay effects associated with recently published articles.

Results

International Collaboration Networks: VOSviewer Software

The cluster generated in *VOSviewer* (Fig. 3) maps the international collaboration network among countries with at least ten scientific publications on the analyzed topic.

Countries such as the United States, Brazil, South Africa, and China are prominently represented and are recurrent contributors to this type of scientific production. The cluster further indicates that proximity between countries reflects the frequency of collaboration in scientific publications, while node size represents the number of publications or the relative weight of each country's participation within the network.

The United States serves as the principal *hub* of international collaboration, exhibiting strong connections with countries in South America (notably Brazil), Europe, Asia, and Africa. Brazil also demonstrates extensive connections, particularly with the United States, the United Kingdom (Scotland), Portugal, and Spain. In Africa, South Africa and Zimbabwe form a significant regional nucleus, maintaining connections with European and American countries. In Asia, China, India, and Nigeria are grouped within the same cluster (represented in blue). European countries such as France, Germany, the Netherlands, Belgium, and Sweden are strongly interconnected, forming a cohesive cluster (green/yellow).

The visualization of international collaboration networks reveals both the geographic distribution of scientific production and the axes of dependence and leadership within the field. The centrality of the United States, which connects it to multiple countries, including Brazil, underscores its role as a global scientific leader. Although Brazil is a leading producer of knowledge on the Cerrado, it remains closely linked to institutions in the Global North, suggesting a degree of academic dependence. This pattern highlights the importance of strengthening South–South alliances, particularly with African and Latin American countries that share similar ecological contexts. Fostering such connections could diversify the epistemic base and promote more context-sensitive solutions to tropical environmental challenges, thereby enhancing regional scientific leadership. The network structure demonstrates that scientific collaboration is shaped by geographic and cultural affinities, as well as institutional and strategic alliances in global knowledge production.

Keyword Analysis: Voyant Tools and VOSviewer

The word cloud analysis using Voyant Tools (Fig. 4) visualizes the most recurrent terms in author keywords related to vegetation indices in the Cerrado. The most frequently mentioned terms include vegetation (83), canopies (77), tree (57), plants (49), savanna (48), index (48), soils (41), forests (41), species (34), and land (31). The recurrence of these terms underscores the predominance of research on vegetation structure and plant cover monitoring in tropical environments, particularly savanna regions. These findings corroborate the co-occurrence analyses performed in VOSviewer, indicating that scientific production is primarily directed toward studies on biodiversity and forest ecosystems.

A total of 3,252 keywords were identified. After filtering duplicates and standardizing synonyms, 1,729 unique keywords remained and were processed in VOSviewer. The resulting co-occurrence map (Fig. 5) represents the thematic structure of scientific production in the analyzed field, grouping terms into three main clusters differentiated by color.

The green cluster, represented by terms such as vegetation, biodiversity, species, plants, tropical forests, and tree canopies, is strongly associated with ecological and biological themes. This cluster indicates a focus on plant diversity, vegetation structure, and tropical forest dynamics.

The red cluster includes terms such as savanna, soils, carbon fluxes, grasses, leaf area index, and water availability. This group primarily focuses on studies of carbon dynamics, ecological modeling, and the physical and structural characteristics of savanna vegetation, with an emphasis on edaphic and physiological aspects. The presence of the term Africa within this cluster suggests recurring research focused on tropical regions of the African continent, where this domain is predominant.

The blue cluster includes terms such as fire cycle, cover changes, indexes, climatic changes, and dynamic analysis. This cluster refers to studies addressing environmental change, remote monitoring, and spatial and temporal analyses of landscape transformations. It is strongly linked to methodological approaches for detecting land-use change and disturbance regimes such as fire and climate change.

Overall, the clustering reveals a diversified and interconnected thematic structure, with interactions among different clusters. This pattern reflects the interdisciplinary nature of scientific production in environmental studies, particularly in research on vegetation, environmental change, and ecosystem dynamics in tropical regions.

Scientific Production: Dendrograms, Principal Component Analysis, and Regression

The dendrogram analysis (Fig. 6) identified three distinct temporal groups in scientific production on vegetation indices in the Cerrado. The first group (1990–1996) represents the initial research phase, characterized by low association with productivity variables. The second group (1997–2002) exhibits an intermediate and transitional pattern. The third and largest group (2003–2024) demonstrates strong internal similarity and greater scientific consolidation. The cutoff point at Euclidean distance 3.0 (red line) clearly delineates these groupings, indicating structural changes in publication patterns over time.

The dendrogram constructed using Euclidean distance (Fig. 7) revealed two distinct groups of variables. The substantial distance (3.0) between these groups confirms a clear division among the analyzed variables.

Principal Component Analysis (PCA) indicated that the first two axes explained 66.28% of the total variance, with 44.92% attributed to Component 1 and 21.36% to Component 2 (Fig. 8). These results reveal distinct patterns in scientific production over time.

Component 1 (x-axis) grouped the years 2003 to 2025 (excluding 2008), which are mainly associated with productivity variables (U2, AB, DE, NW, NR, U1, and NA). In contrast, the years 1990 to 1996 are positioned far from these variables, indicating lower scientific productivity during that period.

Component 2 (y-axis) highlights the years 1997 to 2002 and 2008, which are mainly associated with the variable Z9 (citations in all databases).

The results indicate that recent years have maintained consistent productivity patterns, whereas earlier periods exhibited weaker associations with productivity variables. This finding highlights distinct temporal trends in the characteristics of scientific publications on vegetation indices in the Cerrado.

Regression Analysis Between Years and Variables

Analysis of the evolution in the number of authors per publication (Fig. 9a) reveals a clear upward trend over time. The linear regression model ($y = 0.1297x - 256.13$) indicates an average increase of approximately 0.13 authors per article per year, reflecting steady growth in scientific collaboration. The coefficient of determination ($R^2 = 0.7427$; 74.27%) indicates that the linear model satisfactorily explains the observed variation, accounting for 74.27% of the data's variability.

The analysis also demonstrates a significant upward trend in the number of references cited per publication over time (Fig. 9b). The linear regression model ($y = 1.5271x - 3013.7$) indicates an average increase of approximately 1.53 references per article per year, reflecting a substantial rise in citation volume. The coefficient of determination ($R^2 = 0.7133$; 71.33%) indicates that the linear model adequately explains the observed variation, accounting for over 70% of the dataset's variability. This result suggests consistent growth in the number of references throughout the analyzed period.

Significant growth was observed in both the number of authors over time ($R^2 = 0.74$) and the number of references per article ($R^2 = 0.71$). The relationship between the number of articles published per year and annual citations is positive, but is affected by citation delays in recent years. Excluding the most recent years further strengthens this positive trend, confirming the field's expansion and increasing scientific relevance.

The number of citations per year generally increases as more articles are published (Fig. 10a). However, fewer citations were observed in the three most recent years (2025, 2024, and 2023) compared to previous years, because recently published articles typically accumulate fewer citations. This phenomenon, known as citation delay, occurs when newly published articles have not yet had sufficient time to be read and cited. Excluding the three most recent years reduces this effect and strengthens the relationship between the number of published articles and annual citations (Fig. 10b).

Discussion

Collaborative Networks of Brazil

The international collaboration network demonstrates that Brazil maintains particularly strong scientific partnerships with the United States and England, consistent with studies identifying these countries as

principal historical partners of the Brazilian scientific system [12]. The robustness of these connections suggests that Brazilian scientific output on savannas and vegetation indices is integrated within established global research networks, primarily comprising nations with longstanding traditions in savanna ecology, remote sensing, and environmental modeling, which serve as central hubs in international science [13]. Collaboration with the United States is especially prominent, reflecting both the historical influence of North American science in this field and the country's structural role as the principal hub for global collaboration and knowledge dissemination[14, 15].

The red cluster represents a collaboration axis characterized by historical proximity, linguistic affinities, and strong academic connections between Latin European countries and Brazil, alongside the dominant role of the United States as the primary global hub of scientific collaboration. Research indicates that cultural proximity, shared languages, and similar linguistic roots significantly enhance the likelihood of international collaboration, particularly among countries with Latin heritage[16, 17]. Accordingly, the association among Brazil, Portugal, Spain, and Italy can be attributed to shared scientific traditions, linguistic similarities, and cultural patterns that facilitate communication, academic mobility, and researcher exchange. These factors are recognized as structural determinants of co-authorship networks [14], reinforcing the tendency of Latin countries to maintain closer collaborative relationships. Additionally, these ties are strengthened by the activity of European research groups with established cooperation in Mediterranean ecology, vegetation dynamics, and open biome studies, which methodologically align with research conducted in the Brazilian Cerrado.

Keywords and Clusters

The qualitative analyses conducted with Voyant Tools and VOSviewer reveal that scientific production is concentrated on key themes such as “vegetation,” “biodiversity,” and “climate change,” reflecting relevant global concerns [18]. The diversified thematic structure evidenced by the generated clusters demonstrates the interconnection among research areas and the importance of interdisciplinary approaches in ecosystem studies. This finding aligns with the growing need for approaches that integrate scientific knowledge applicable to contemporary social and environmental problems, corroborating the idea presented by [19] regarding the relevance of scientometrics in visualizing research gaps and informing appropriate policy formulation.

The predominance of terms such as vegetation, canopies, forest, plants, species, and tree in the word cloud reflects the essential focus of studies on savannas and the Cerrado, where vegetation structure and the arboreal component are fundamental to understanding ecological dynamics [20]. In savanna ecosystems, the tree canopy plays a crucial role in modulating microclimate, nutrient cycling, and water resource partitioning, and is frequently analyzed using vegetation indices such as NDVI, EVI, and LAI derived from remote sensing [21, 22]. Therefore, terms related to vegetation composition and architecture recur in research aimed at characterizing the phenology, productivity, and spatial patterns of arboreal vegetation in these formations.

Furthermore, terms such as savanna, canopies, biomass, fluxes, variability, and water availability underscore the significance of ecophysiological and biogeochemical processes that are central to the functioning of tropical savannas [23]. The pronounced hydrological seasonality of the Cerrado and its sensitivity to disturbances, particularly fire, necessitate metrics that capture temporal and spatial fluctuations in vegetation, accounting for the frequent appearance of terms related to variability, season, cycle, and fluxes. Additionally, the prevalence of modeling and environmental analysis terms such as analysis, spatial, surface, changes, and variability reflects the growing adoption of quantitative and geospatial methods, supported by orbital data, to monitor arboreal vegetation dynamics

The concurrent use of ecology-related terms (such as biodiversity, ecosystems, habitats) and remote sensing-related terms (such as images, surfaces, indices) highlights the interdisciplinary nature of this research field. Investigations of savannas and the Cerrado using vegetation indices operate at the intersection of ecology, climatology, geoprocessing, and environmental analysis, resulting in a convergence of terminology that encompasses both biological components and technological methodologies [24]. Consequently, the word cloud encapsulates the field's principal conceptual axes: arboreal vegetation structure, ecological dynamics, ecophysiological processes, and the application of spectral metrics to interpret environmental changes in savanna ecosystems.

Scientific Production (Multivariate Analyses and Regressions)

The scientometric analysis conducted in this study seeks to elucidate the evolution of scientific production in specific fields by employing quantitative methods to assess relationships among variables over time. The results highlight correlations between the number of pages and the number of references, and the application of statistical techniques, such as the Pearson correlation matrix and principal component analysis, which revealed significant trends in scientific productivity.

Preliminary correlation analysis revealed that variables such as number of pages (PG) and number of references (NR) are positively correlated, indicating that longer publications typically include more references. This observation is consistent with studies documenting an increase in reference counts over time, reflecting the expansion and reconstruction of scientific knowledge, where intertextuality is pivotal [19, 25]. Additionally, the exclusion of redundant variables constituted a key methodological approach, as supported by [26], who emphasize the importance of clarity and efficiency in scientific data analysis.

The principal component analysis (PCA) and dendrogram construction allowed not only mapping of the temporal evolution of scientific production but also categorization of publications into distinct periods, revealing substantial growth in research activity. The isolated clustering of the years 1990 to 1996 (except 1994) in the dendrogram reflects an embryonic phase of research on vegetation indices in the Cerrado.

This pattern indicates limited scientific production during that period, potentially attributable to the minimal inclusion of the topic in the international research agenda, technological constraints in remote sensing studies, and insufficient targeted funding. The lack of association with productivity variables supports this hypothesis and underscores the contrast with later years, which exhibit more robust and consolidated publication patterns. This growth may reflect heightened interest and increased funding for environmental and ecological research, particularly in critical regions such as the Cerrado, which has attracted growing attention for its ecological and economic significance [25]

PCA further revealed distinct temporal patterns. Component 1, responsible for 44.92% of the variance, grouped the years from 2003 to 2025 (except 2008), strongly associated with scientific productivity variables: number of authors (NA), words in the title (NW), authors' keywords (DE), words in the abstract (AB), number of references (NR), and usage metrics (U1 and U2).

This grouping suggests that recent scientific production is more consolidated, collaborative, and technically advanced. In contrast, the years from 1990 to 1996 are positioned further from these vectors, confirming the lower sophistication and maturity of research during the early phase of the historical series. Component 2, accounting for 21.36% of the variance, highlights the years 1997 to 2002 and 2008, which are strongly influenced by the variable Z9 (citations across all databases), indicating a period of greater academic impact, likely associated with the consolidation of theoretical frameworks and vegetation-monitoring methodologies.

The regression analysis reveals expanding scientific collaboration, evidenced by an average increase in co-authors per publication. This trend is consistent with global patterns of academic collaboration, where multidisciplinary teamwork is increasingly prevalent [27, 28]. The quadratic growth model for citations indicates that, despite increases in production and referencing, a potential citation saturation phenomenon emerged after 2010, underscoring the need to examine publication dynamics and shifting research interests [29].

The post-2010 inflection in citation patterns may be attributed to thematic saturation or increasing fragmentation within the field. As publication volume rises, individual articles face greater challenges in achieving high citation counts, especially when research topics are divided into narrowly defined niches. Furthermore, recent studies may emphasize localized or methodological applications with limited global relevance. These trends underscore the need to foster research that integrates originality with international significance and highlight the importance of editorial policies and funding strategies that prioritize scientific quality and impact.

In summary, this study's findings illustrate the evolution of research on vegetation indices in the Cerrado and offer a critical perspective on the fundamental role of scientometrics in modeling and interpreting scientific data. This approach fosters a deeper understanding of collaboration dynamics and the evolution of knowledge within academia.

Conclusion

This scientometric study examined trends and developments in research on vegetation indices in the Cerrado, emphasizing the temporal evolution of scientific production, international collaborations, and principal thematic areas. A marked increase in publications occurred after the 1990s, especially between 2000 and 2025, characterized by higher productivity and enhanced collaboration among authors.

Collaboration network analysis revealed the central roles of the United States and Brazil in knowledge production on this topic, although some academic dependence on the Global North persists. The interdisciplinary character of the research is evidenced by the co-occurrence of terms such as vegetation, biodiversity, savanna, climate, and changes, reflecting integration across ecology, conservation, and environmental monitoring.

Regression models identified key trends, including increases in the number of authors per article and citation volumes, indicating stronger scientific collaboration and greater theoretical depth in publications. However, the quadratic citation pattern observed after 2010 suggests potential thematic saturation, underscoring the need for innovative research with broader global relevance.

This study is limited by the exclusive use of the Web of Science database, which may not encompass all relevant scientific production related to vegetation indices in the Cerrado. Additionally, the scientometric approach emphasizes quantitative patterns and may not fully capture the qualitative dimensions of scientific contributions. Future studies should integrate additional databases and complementary analytical approaches to provide a broader understanding of research trends in this field.

Declarations

Competing interests

The authors have no competing interests to declare.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

Funding

This study was funded by the Call for Proposals PRÓ-PROJETOS – Institutional Strategic Project – No. 20/2023 (funding ID 135 in the Pró-Pesquisa System, UEG; SEI: 202300020023365) and the Call for Proposals Institutional Platform for Research and Innovation in Water Security PrP No. 04/2024.

Author Contribution

L.M.D. and V.S.V. conceived the study. L.M.D., E.N.A., and V.S.V. developed the methodology. L.M.D. performed the formal analyses and wrote the original draft of the manuscript. L.M.D., I.L.P., and J.K.M.V. organized and curated the data. L.M.D. and R.M.S.J. prepared the figures and visualizations. E.N.A., I.L.P., J.K.M.V., R.M.S.J., and V.S.V. reviewed and edited the manuscript. V.S.V. supervised the study. All authors reviewed and approved the final manuscript.

Acknowledgement

The authors acknowledge the financial support provided by the Call for Proposals PRÓ-PROJETOS – Institutional Strategic Project – No. 20/2023 (funding ID 135 in the Pró-Pesquisa System, UEG; SEI: 202300020023365) and the Call for Proposals Institutional Platform for Research and Innovation in Water Security PrP No. 04/2024.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Strassburg BBN, Brooks T, Feltran-Barbieri R, et al. Moment of truth for the Cerrado hotspot. Nat Ecol Evol. 2017;1:0099. https://doi.org/10.1038/s41559-017-0099.
2. Sano EE, Rosa R, Brito JLS, Ferreira LG. Land cover mapping of the tropical savanna region in Brazil. Environ Monit Assess. 2010;166:113–24. https://doi.org/10.1007/s10661-009-0988-4.
3. Silva AGP, Galvão LS, Ferreira Júnior LG, et al. Discrimination of Degraded Pastures in the Brazilian Cerrado Using the PlanetScope SuperDove Satellite Constellation. Remote Sens. 2024;16:2256. https://doi.org/10.3390/rs16132256.
4. Shiroma GM, Gamarra RM, Guirra APM, et al. Comparison Between Vegetation Indexes from Different Sensors to Identify Vegetation Types from Tropical Savanna. Anuário IGEO UFRJ. 2020;43:18–32. https://doi.org/10.11137/2020_1_18_32.
5. Guimarães MJM, Reis IDD, Barros JRA, et al. Identification of Vegetation Areas Affected by Wildfires Using RGB Images Obtained by UAV: A Case Study in the Brazilian Cerrado. Geomatics. 2025;5:13. https://doi.org/10.3390/geomatics5010013.

6. Donthu N, Kumar S, Mukherjee D, et al. How to conduct a bibliometric analysis: An overview and guidelines. J Bus Res. 2021;133:285–96. https://doi.org/10.1016/j.jbusres.2021.04.070.
7. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. SciMAT: A new science mapping analysis software tool. J Am Soc Inf Sci Tec. 2012;63:1609–30. https://doi.org/10.1002/asi.22688.
8. Zupic I, Čater T. Bibliometric Methods in Management and Organization. Organizational Res Methods. 2015;18:429–72. https://doi.org/10.1177/1094428114562629.
9. Leydesdorff L, Bornmann L, Wagner CS. The Relative Influences of Government Funding and International Collaboration on Citation Impact. Asso Info Sci Tech. 2019;70:198–201. https://doi.org/10.1002/asi.24109.
10. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84:523–38. https://doi.org/10.1007/s11192-009-0146-3.
11. Gotelli NJ, Ellison AM. Princípios de Estatística em Ecologia. Artmed Editora; 2016.
12. Sidone OJG, Haddad EA, Mena-Chalco JP. A ciência nas regiões brasileiras: evolução da produção e das redes de colaboração científica. Transinformação. 2016;28:15–32. https://doi.org/10.1590/2318-08892016002800002.
13. Glänzel W. National characteristics in international scientific co-authorship relations. Scientometrics. 2001;51:69–115. https://doi.org/10.1023/A:1010512628145.
14. Wagner CS, Leydesdorff L. Network structure, self-organization, and the growth of international collaboration in science. Res Policy. 2005;34:1608–18. https://doi.org/10.1016/j.respol.2005.08.002.
15. Leydesdorff L, Wagner CS. International collaboration in science and the formation of a core group. J Informetrics. 2008;2:317–25. https://doi.org/10.1016/j.joi.2008.07.003.
16. Hoekman J, Frenken K, Van Oort F. The geography of collaborative knowledge production in Europe. Ann Reg Sci. 2009;43:721–38. https://doi.org/10.1007/s00168-008-0252-9.
17. Pan RK, Kaski K, Fortunato S. World citation and collaboration networks: uncovering the role of geography in science. Sci Rep. 2012;2:902. https://doi.org/10.1038/srep00902.
18. Dias T, Silva VS, Silva JVD et al. (2023) BrCris como ferramenta para acesso e avaliação da ciência. RevBib 9:. https://doi.org/10.21680/2447-7842.2023v9n2ID33838
19. Parra MR, Coutinho RX, Pessano EFC, TENDÊNCIAS E SUA CONTRIBUIÇÃO PARA O ENSINO DE CIÊNCIAS. UM BREVE OLHAR SOBRE A CIENCIOMETRIA: ORIGEM, EVOLUÇÃO. Contexto. 2019;34:126–41. https://doi.org/10.21527/2179-1309.2019.107.126-141.
20. Ratter J. The Brazilian Cerrado Vegetation and Threats to its Biodiversity. Ann Botany. 1997;80:223–30. https://doi.org/10.1006/anbo.1997.0469.
21. Huete A, Didan K, Miura T, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens Environ. 2002;83:195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

22. Figueiredo EO, d'Oliveira MVN, Braz EM, et al. LIDAR-based estimation of bole biomass for precision management of an Amazonian forest: Comparisons of ground-based and remotely sensed estimates. Remote Sens Environ. 2016;187:281–93. https://doi.org/10.1016/j.rse.2016.10.026.
23. Grace J, José JS, Meir P, et al. Productivity and carbon fluxes of tropical savannas. J Biogeogr. 2006;33:387–400. https://doi.org/10.1111/j.1365-2699.2005.01448.x.
24. Bond WJ. Ancient grasslands at risk. Science. 2016;351:120–2. https://doi.org/10.1126/science.aad5132.
25. Chaebo RFP, Chaebo G. Análise cienciométrica em zoneamento ecologico-econômico: tendências e contribuições. BJD. 2019;5:13292–310. https://doi.org/10.34117/bjdv5n8-136.
26. Moresi EAD, Pinho I. Proposta de abordagem para refinamento de pesquisa bibliográfica. New Trends Qualitative Res (NTQR). 2021;9:11–20. https://doi.org/10.36367/ntqr.9.2021.11-20.
27. De Camargo LS, Barbosa RR, CIENCIOMETRIA E UM POSSÍVEL CAMINHO PARA A CONSTRUÇÃO, DE INDICADORES E MAPAS DA PRODUÇÃO CIENTÍFICAca. RPA 12:109. https://doi.org/10.9771/rpa.v12i3.28408.
28. Silva Filho JND, Damaceno C. Gestão escolar e Cienciometria: estudo sobre o volume de publicações científicas na base Scielo Brasil. ER. 2021;22:121–36. https://doi.org/10.36311/2236-5192.2021.v22esp2.p121.
29. Cremasco Receputi C, Weiss A, Vogel M, ESTADO DA ARTE DAS PUBLICAÇÕES DOS DISCENTES DO CURSO DE LICENCIATURA EM QUÍMICA. Rev Elet Deb Educ Cient Tecnol. 2019;7:26–45. https://doi.org/10.36524/dect.v7i03.205.

Figures

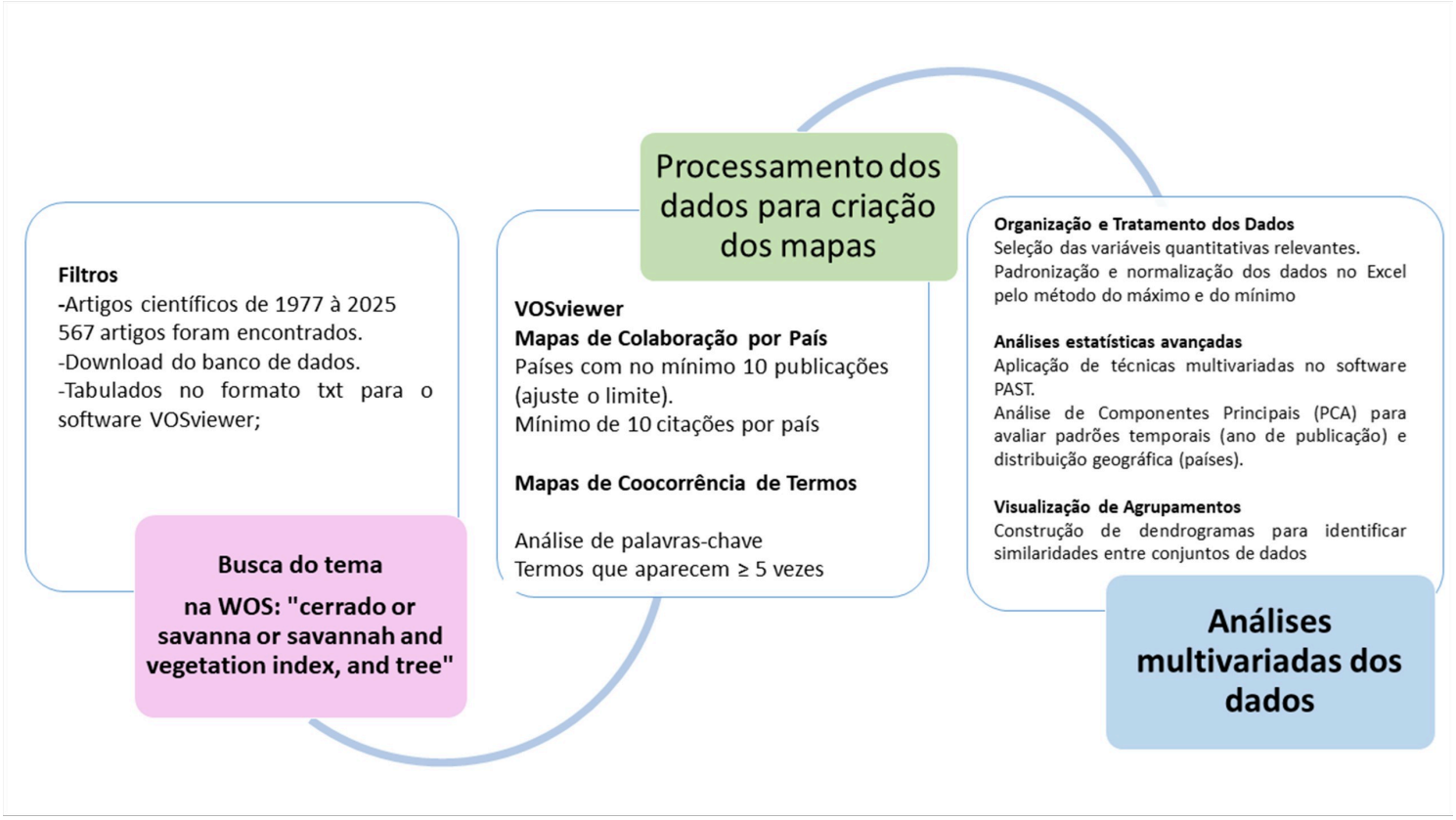


Figure 1

Schematic representation of the methodological steps, illustrating the processes of data search, treatment, and analysis

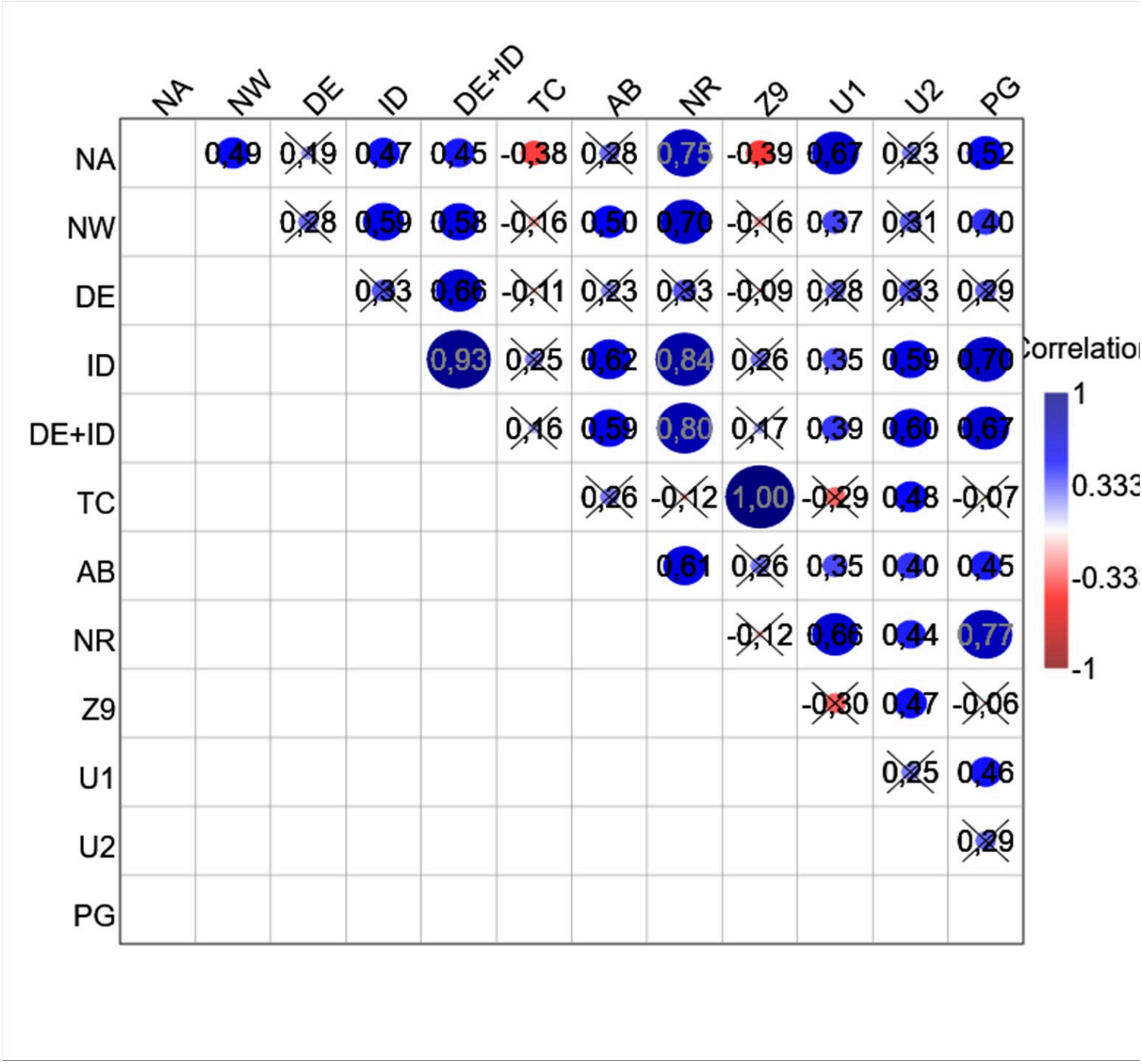


Figure 2

Pearson correlation matrix ($p < 0.05$) highlighting variables excluded due to high similarity. Colored circles indicate the strength and direction of correlations, with size proportional to the magnitude of the coefficient (r). Circles marked with an “x” indicate non-significant correlations

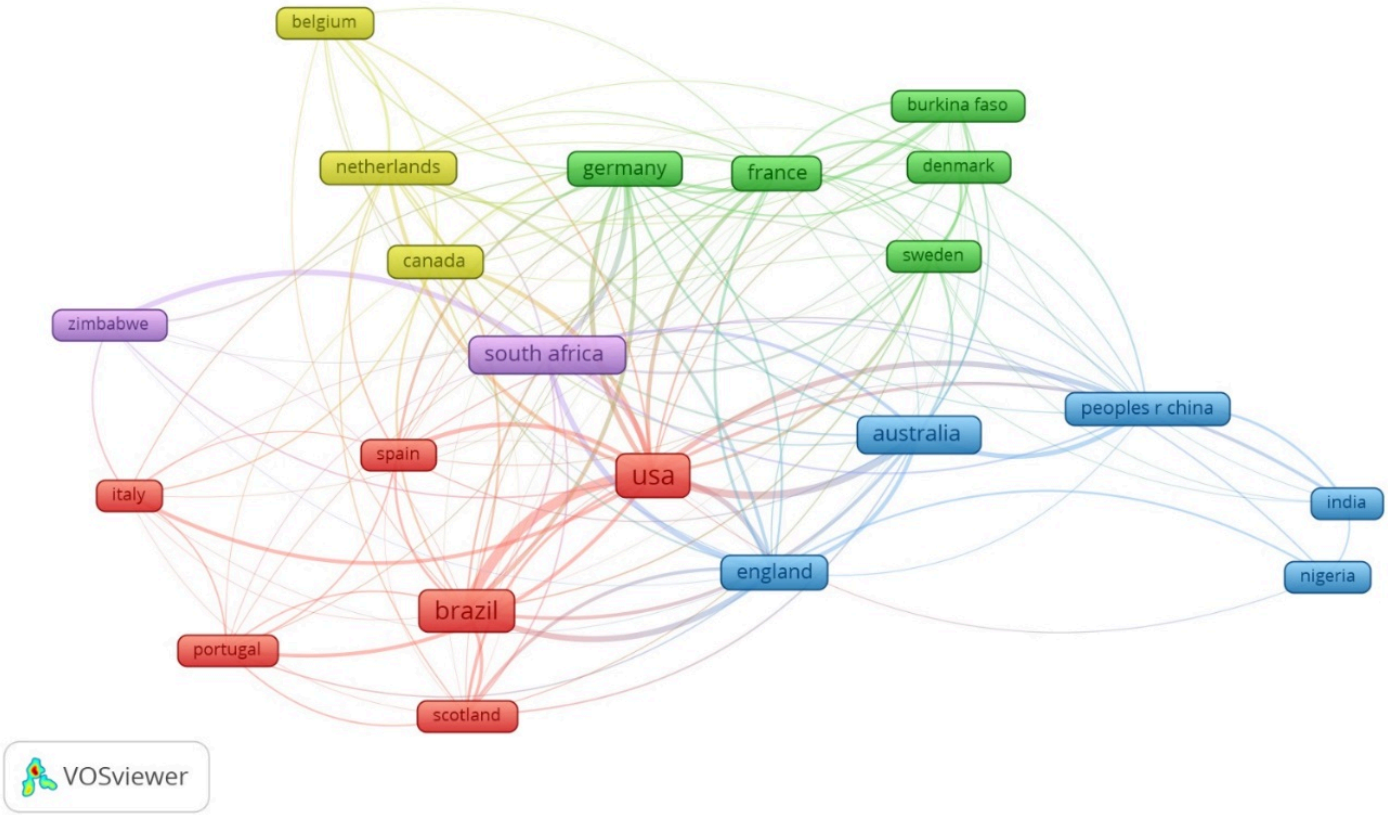


Figure 3

Countries with studies related to the terms "cerrado," "savanna," "savannah," "vegetation index," and "tree." Only countries with at least 10 articles were included. Source: Authors



Figure 4

Word cloud generated by the Voyant Tools platform based on author keywords extracted from Web of Science

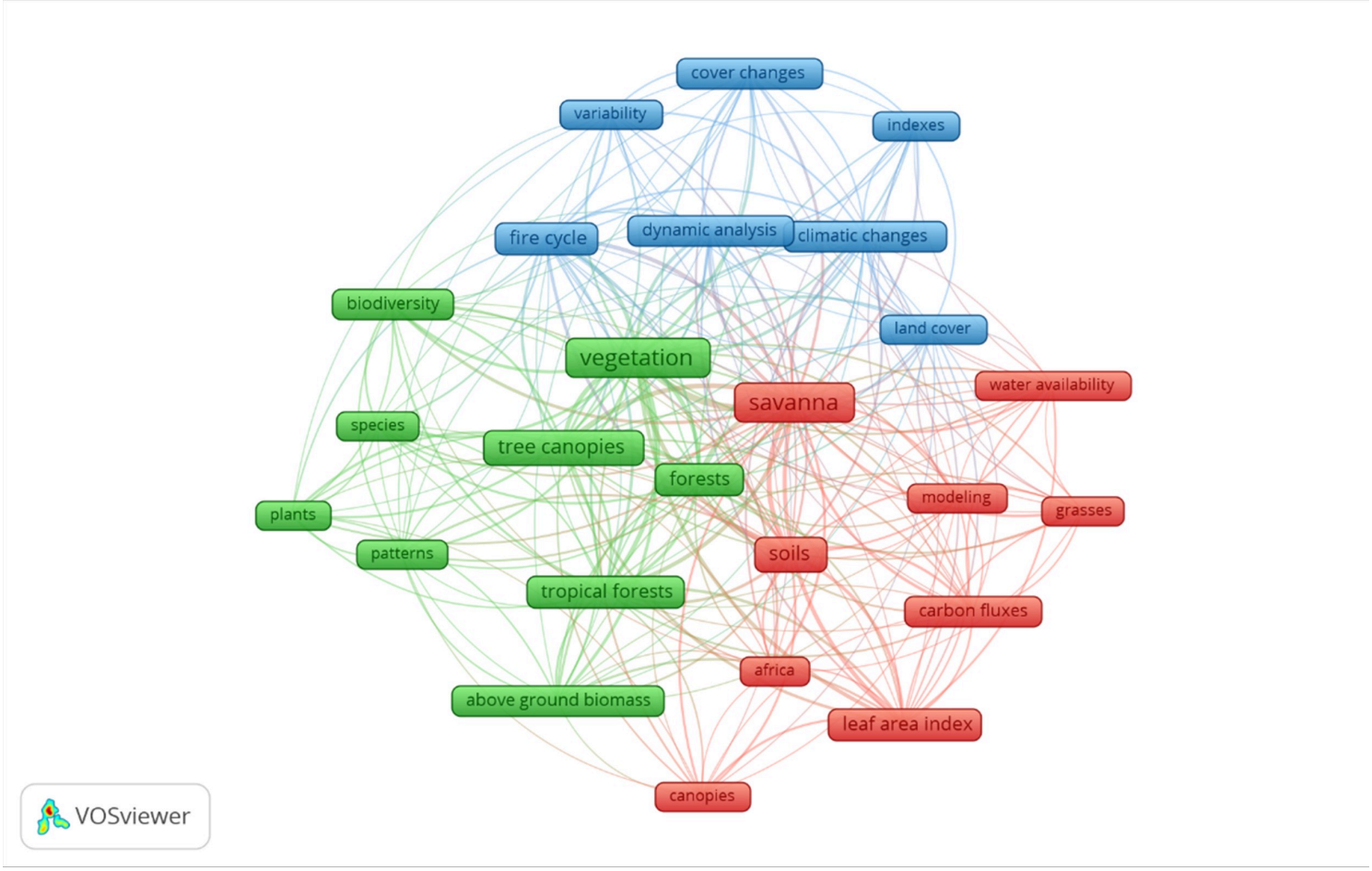


Figure 5

Most frequently mentioned keywords related to the terms "cerrado," "savanna," "savannah," "vegetation index," and "tree" (only unique keywords with more than five occurrences were included). Larger circles represent higher keyword frequency.

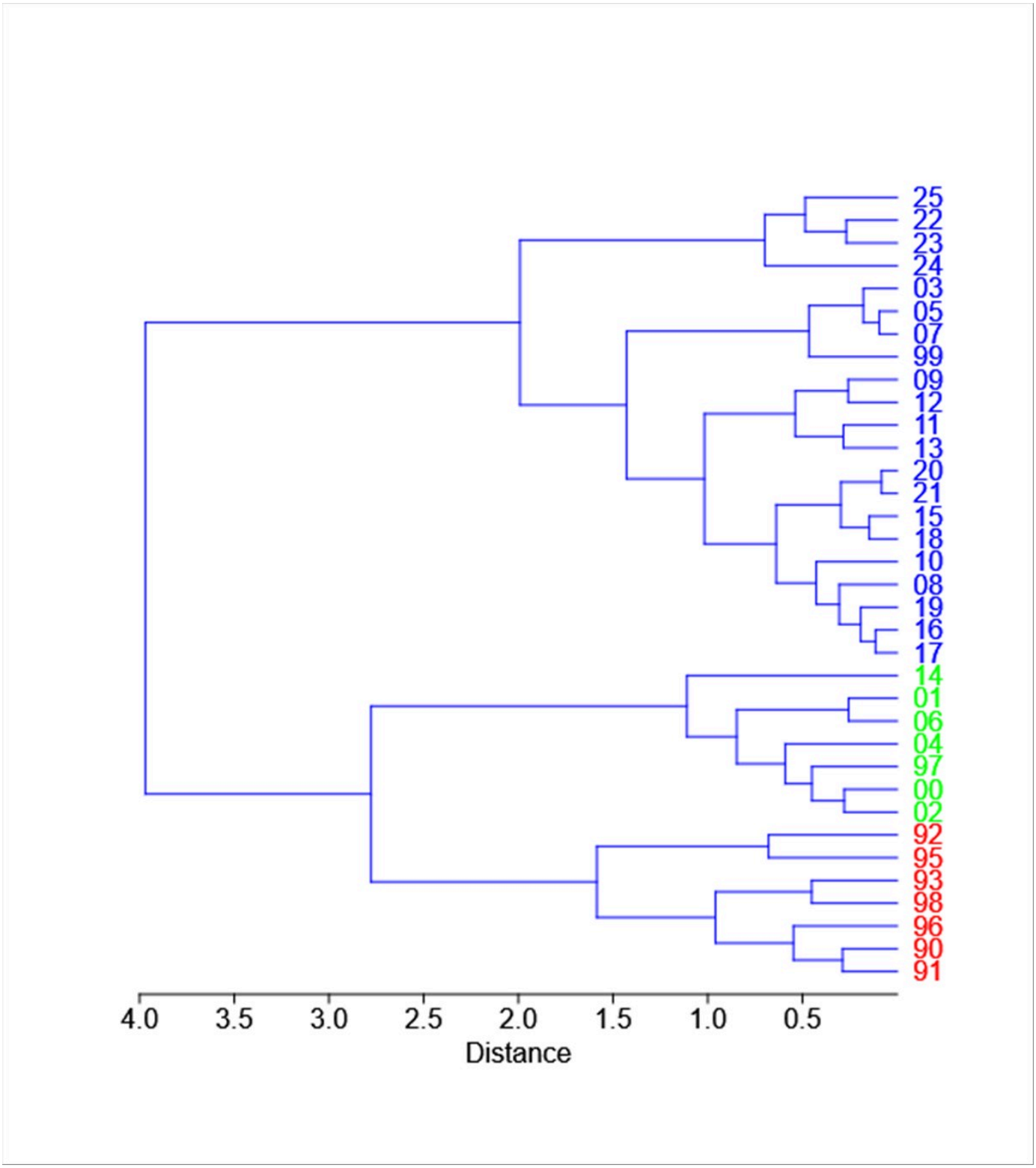


Figure 6

Hierarchical clustering dendrogram of the analyzed publication years, constructed based on Euclidean distance and Ward's method, revealing temporal similarity patterns in scientific production on vegetation indices in the Cerrado. Cutoff point at 3.0 (red line)

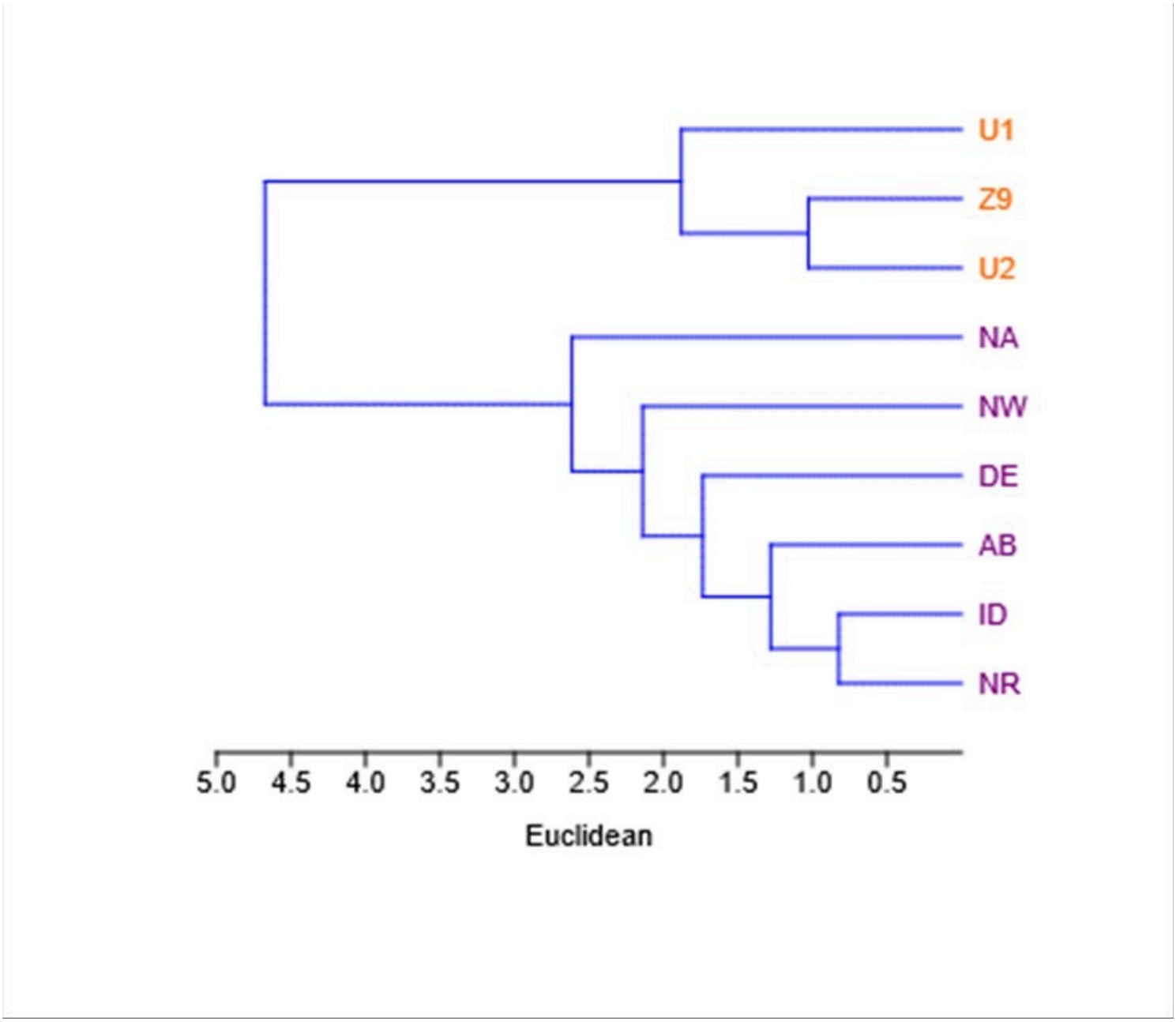


Figure 7

Hierarchical clustering dendrogram showing similarity patterns among scientific production variables on vegetation indices in the Cerrado. Constructed using Euclidean distance and Ward's method. Cutoff point at 3.0 (red line)

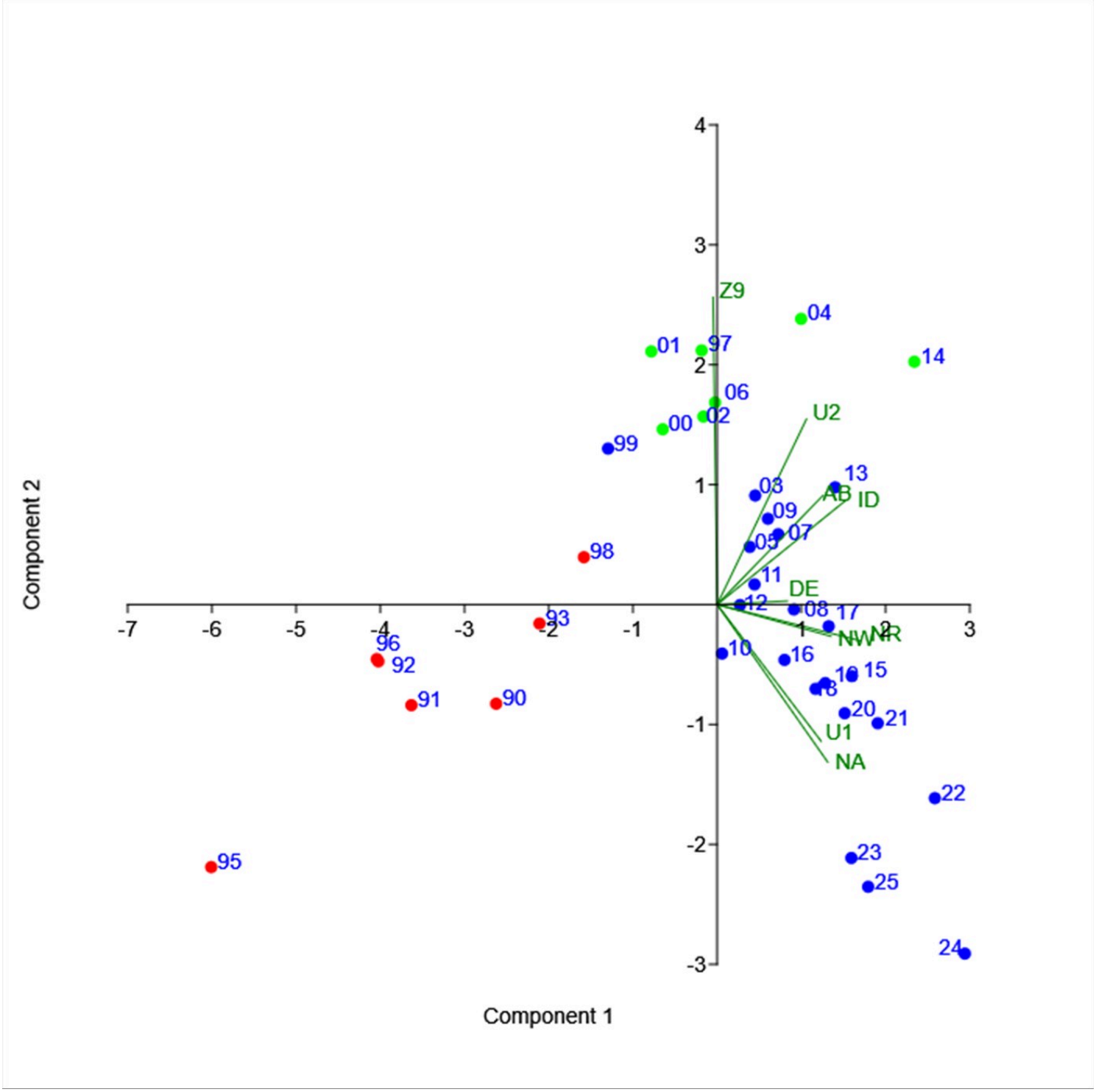


Figure 8

Principal Component Analysis (PCA) results applied to scientometric data, showing the distribution of publication years (1990–2025) across productivity variables

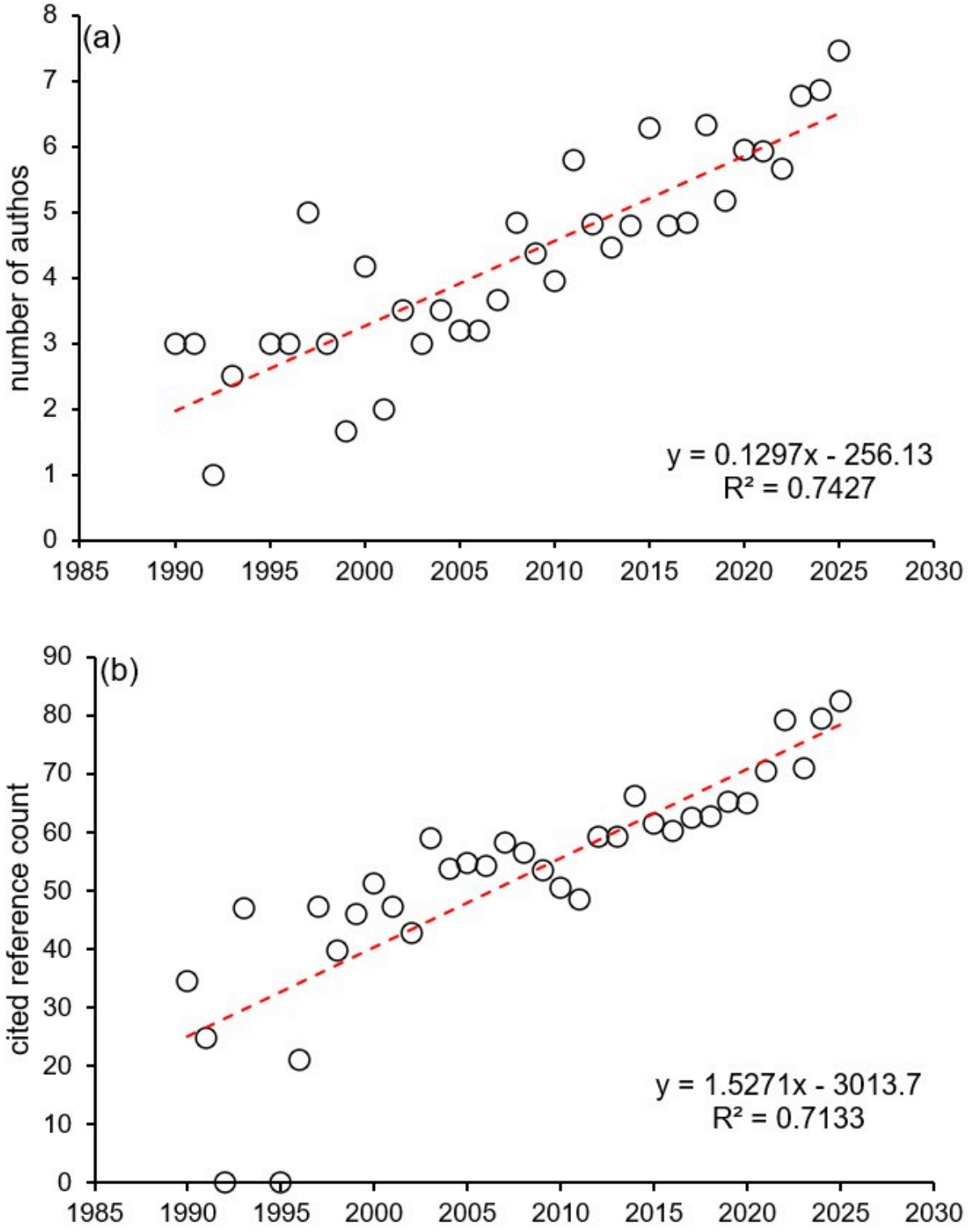


Figure 9

Regression analysis of temporal trends in scientific production: (a) linear growth in the number of authors; (b) linear increase in cited references.

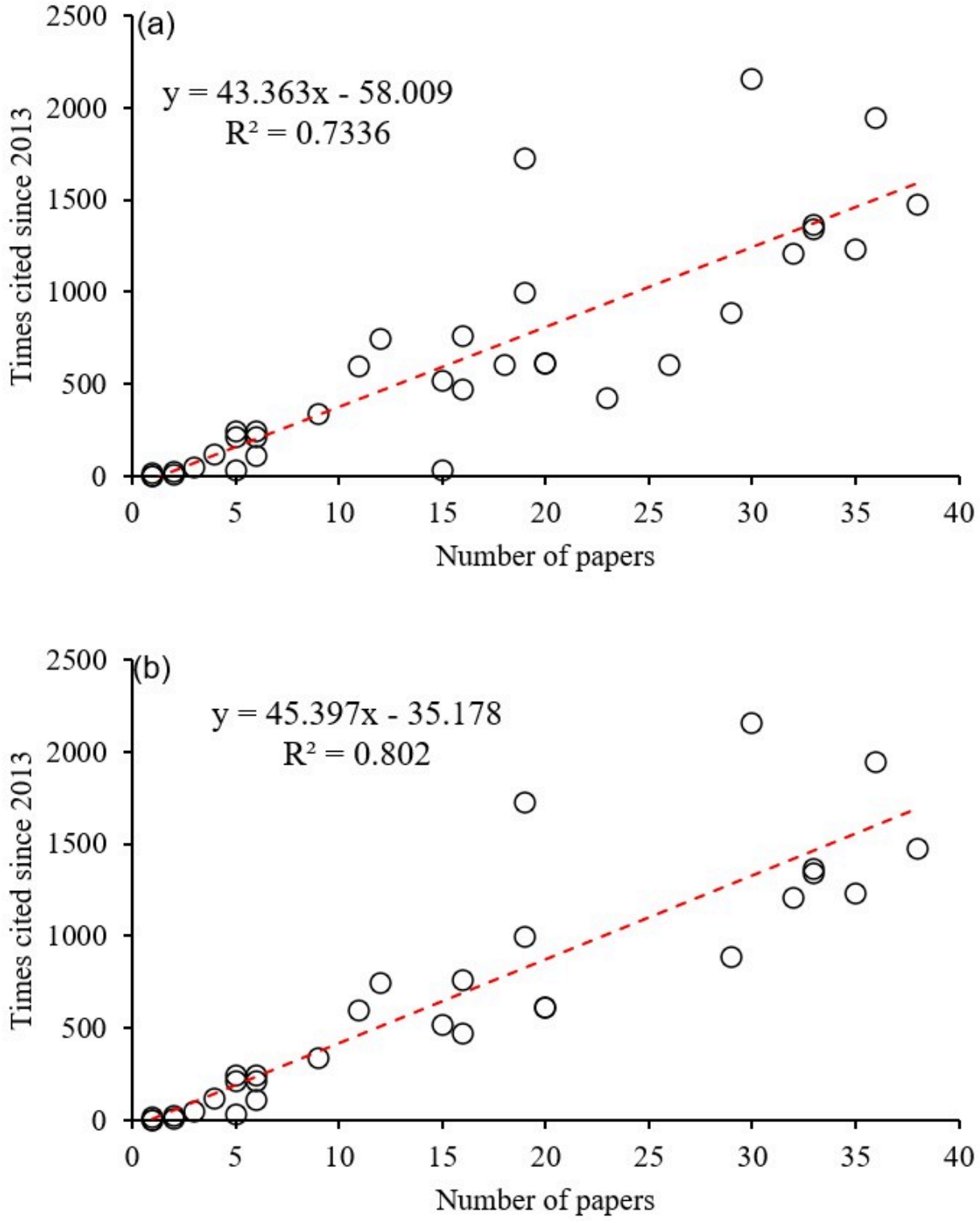


Figure 10

Simple regression of citations since 2013 against the number of published articles for all years (a) and excluding the years 2025, 2024, and 2023 (b).