BILINGUAL PUBLISHING GROUP

Research in Ecology
https://journals.bilpubgroup.com/index.php/re

REVIEW

Mapping the Research Progress of Soil Quality in Camellia oleifera Forests: A Bibliometric Analysis Using CiteSpace

Yan-kui Chen [1,2*], *Meng-Choung Chiong* [1], *Xue-fei Mo* [2,3], *Yu Zhang* [2,4]

[1] *Faculty of Engineering, Technology & Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2] *College of Geographical Science and Tourism, Jiaying University, Meizhou 514015, China*
[3] *Department of Land Management, Faculty of Agriculture, Universiti Putra Malaysia, Serdang 43400, Malaysia*
[4] *International College, Krirk University, Bangkok 10220, Thailand*

ABSTRACT

This study provides a comprehensive synthesis of the evolution, structural patterns, and emerging directions of research on soil quality in Camellia oleifera plantations. Relevant publications from 1995 to 2024 were retrieved from the China National Knowledge Infrastructure (CNKI) and the Web of Science Core Collection (WoSCC). A bibliometric and knowledge-mapping approach was employed using CiteSpace, with analyses focusing on publication trends, collaboration networks among authors and institutions, keyword co-occurrence and clustering, and burst detection. The results show that: (1) research in this field has evolved through three phases: initial exploration (1995–2006), steady expansion (2007–2020) and rapid advancement (2021–2024), a temporal analysis of keyword co-occurrence and thematic clustering reveals a thematic shift in research emphasis over time. Early studies centered on discrete soil components, while more recent work has increasingly focused on process-oriented management and integrative system-level regulation; (2) Chinese and international studies follow largely parallel but complementary trajectories, with Chinese research emphasizing applied topics such as soil fertility assessment, nutrient management, and yield improvement, while international research focuses more on underlying mechanisms, including soil acidification, microbial diversity, and aluminum-phosphorus interactions; these strands are increasingly converging; (3) current research hotspots include mechanisms and mitigation of soil acidification, construction of minimum data sets (MDS),

*CORRESPONDING AUTHOR:
Yan-kui Chen, Faculty of Engineering, Technology & Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia; College of Geographical Science and Tourism, Jiaying University, Meizhou 514015, China; Email: 1002475143@ucsiuniversity.edu.my or 519958394@qq.com

ARTICLE INFO
Received: 27 January 2026 | Revised: 26 March 2026 | Accepted: 1 April 2026 | Published Online: 14 July 2026
DOI: https://doi.org/10.30564/re.v8i4.13080

CITATION
Chen, Y.-k., Chiong, M.-C., Mo, X.-f., et al., 2026. Mapping the Research Progress of Soil Quality in Camellia oleifera Forests: A Bibliometric Analysis Using CiteSpace. Research in Ecology. 8(4): 53–76. DOI: https://doi.org/10.30564/re.v8i4.13080

and regulation of rhizosphere microecology, with methodological advances moving from conventional physicochemical analyses toward high-throughput sequencing and machine learning; and (4) future research should prioritize multi-scale integration and long-term observations, strengthen coupled soil plant microbe system studies, advance intelligent decision-support tools for soil health management, and develop integrated assessment frameworks for sustainable Camellia oleifera development.

Keywords: Camellia oleifera; Soil Quality; CiteSpace; Knowledge Mapping; Visualization

1. Introduction

Camellia oleifera Abel is a woody oilseed crop endemic to China, valued not only for its economic importance but also for its ecological functions. The oil extracted from its seeds is rich in nutrients and widely recognized as the Eastern olive oil [1]. In addition, Camellia oleifera plantations contribute to carbon sequestration, soil and water conservation, and biodiversity maintenance, representing a typical eco-friendly land-use system that does not compete directly with staple food crops for arable land. At present, the cultivation area in China exceeds 4.87 million hectares, with annual production accounting for more than 90% of global output [2]. As such, the industry plays a pivotal role in promoting rural development while supporting the coordinated advancement of ecological sustainability and food security. Soil serves as the fundamental medium for the growth and development of Camellia oleifera, and its quality directly influences plantation productivity, product quality, and long-term sustainability. However, long-standing management practices, particularly the excessive reliance on chemical fertilizers and insufficient return of organic matter have resulted in a range of soil-related problems. These include nutrient imbalances [3], accelerated soil acidification [4,5], and the deterioration of soil biological functions [6–8], all of which pose significant constraints on the sustainable development of the Camellia oleifera industry. A systematic understanding of soil quality dynamics in these plantations is therefore essential for optimizing management strategies and achieving a balance between ecological and economic objectives.

Soil quality is generally defined as the soil's capacity to sustain plant productivity, regulate environmental quality, and support biological health within a given ecosystem. While closely related, the concepts of soil nutrients, fertility, quality, and health differ in emphasis. Soil nutrients constitute the material basis of soil fertility [9,10], soil fertility represents a core component of soil quality [11], and soil quality, together with soil health, underpins overall ecosystem functioning. Notably, soil health places greater emphasis on biological activity and system resilience [12,13]. Despite increasing research attention, most existing studies remain focused on individual attributes or localized conditions, resulting in a fragmented understanding of soil quality in Camellia oleifera plantations. With advances in analytical techniques, molecular biology, and information science, research in this field has expanded beyond macroscopic assessments of agronomic traits to increasingly include integrative approaches that combine molecular-level mechanisms with ecosystem-scale processes. For example, field experiments have shown that the combined application of organic and inorganic fertilizers can enhance soil fertility [14], while microbial inoculation can mitigate the adverse effects of continuous cropping and improve nutrient availability [15]. In addition, studies have revealed dynamic interactions between soil health indicators, such as microbial biomass carbon, and fertility parameters, such as available potassium [15,16]. Nevertheless, despite the expansion in research scope, increasing analytical depth, and growing interdisciplinarity, a notable gap persists between practical advancements and the development of comprehensive theoretical frameworks. Moreover, existing findings are dispersed across different disciplines and publication outlets, lacking an integrated knowledge system that captures the evolution, internal linkages, and future trajectories of soil quality research in Camellia oleifera plantations. Scientific knowledge mapping, as an emerging bibliometric approach, provides an effective means of addressing this gap by visualizing complex knowledge structures and revealing patterns of scientific development. Among the available tools, CiteSpace, developed by Chen Chaomei, has become widely adopted due to its robust capabilities in network analysis and visualization, particularly in keyword co-occurrence, clustering, and burst detection [17]. It has been successfully applied in fields such as agroecology [18] and soil health [19], offering a systematic method to map the intellectual structure of studies on camellia oil plantation soils and uncover latent conceptual relationships within this domain. It is important to clarify that this bibliometric approach is descriptive and exploratory in nature; it maps the structural evolution and thematic landscape of the literature rather than establishing causal explanations for scientific change. The knowledge maps serve as a heuristic tool to identify patterns of co-occurrence and shifting emphasis, providing a macro-level overview to complement traditional, in-depth reviews.

Against this background, the present study employs CiteSpace and bibliometric methods to conduct a systematic quantitative analysis of core literature on soil quality in Camellia oleifera plantations, based on data retrieved from the China National Knowledge Infrastructure (CNKI) and the Web of Science Core Collection (WoSCC) for the period 1995–2024. The specific objectives are as follows: (1) to characterize the overall research landscape, including temporal trends, spatial distribution, and major contributing entities; (2) to elucidate the structural organization of research themes through keyword co-occurrence and clustering analysis; (3) to track the evolution of research hotspots using burst detection and timeline visualization; and (4) to identify current focal areas and emerging frontiers, including soil acidification, the development of minimum data sets (MDS), and rhizosphere microecological regulation. The findings are expected to provide a theoretical foundation and decision-support reference for optimizing soil quality management and promoting the sustainable development of the Camellia oleifera industry.

2. Data Sources and Research Methods

2.1. Data Sources and Retrieval Strategy

To systematically delineate the academic development of soil quality research in Camellia oleifera plantations, this study selected CNKI and WoSCC as primary data sources. CNKI provides comprehensive coverage of domestic scholarship on Camellia oleifera, particularly excelling in region-specific and application-oriented studies, whereas WoSCC encompasses high-impact international journals, enabling a global perspective on disciplinary developments. The combined use of these databases reduces the potential bias inherent in single-source analyses and enhances the robustness and validity of the findings. The temporal scope was defined from 1 January 1995 to 31 December 2024, thereby capturing the full trajectory of research alongside the development of China's Camellia oleifera industry. In terms of document types, CNKI indexes academic journals as well as master's and doctoral theses, whereas Web of Science (WoS) includes only peer-reviewed journal articles and review papers-ensuring methodological consistency and cross-database comparability.

For the CNKI database, a combined subject search strategy was employed using a Boolean query: (camellia oleifera OR camellia oleifera plantation OR camellia oleifera cultivation) AND (soil quality OR soil nutrients OR soil fertility OR Soil health OR soil degradation OR soil acidification OR soil properties), with additional synonyms and related terms incorporated to maximise retrieval coverage. Following the initial search, core journals and CSSCI-indexed publications were selected by source category, and irrelevant or duplicate records were manually excluded. In the Web of Science database, a topic search (TS) was conducted using the query: TS = (Camellia oleifera) AND TS = (soil quality OR soil fertility OR soil nutrient OR Soil health OR soil degradation OR soil acidification OR soil property), with wildcard operators applied to capture lexical variations. During the secondary screening, studies focusing exclusively on plant physiology or oil processing without reference to soil properties were excluded. Publications within relevant disciplinary categories, such as environmental science, soil science, agriculture, and forestry were prioritised to ensure thematic alignment.

The literature screening and data cleaning followed a standardized, reproducible procedure. First, two independent researchers conducted title-and-abstract screening to exclude studies clearly unrelated to the research scope. Search results were exported in both RefWorks and plain-text formats and subsequently imported into EndNote for integration and automated deduplication. Next, the preliminarily included records underwent manual title-and-abstract screening to exclude: (i) studies not focused on Camellia oleifera; (ii) those not addressing soil quality; (iii) duplicate publications; and (iv) those with incomplete or non-reportable data. Full-text assessment followed, with all inclusion/exclusion decisions independently verified by both reviewers. Discrepancies were resolved through consensus discussion; unresolved cases were adjudicated by a third senior researcher. Ultimately, 383 Chinese-language and 88 English-language articles met all eligibility criteria. **Figure 1** (PRISMA flowchart) presents the complete screening process and the number of records at each stage. The retained records were further standardised through author name disambiguation, institutional name harmonisation, and keyword unification. The resulting structured database was then imported into CiteSpace for subsequent visualisation and knowledge mapping analyses. The entire workflow was transparent and reproducible, providing a rigorous and reliable data foundation for subsequent analyses.

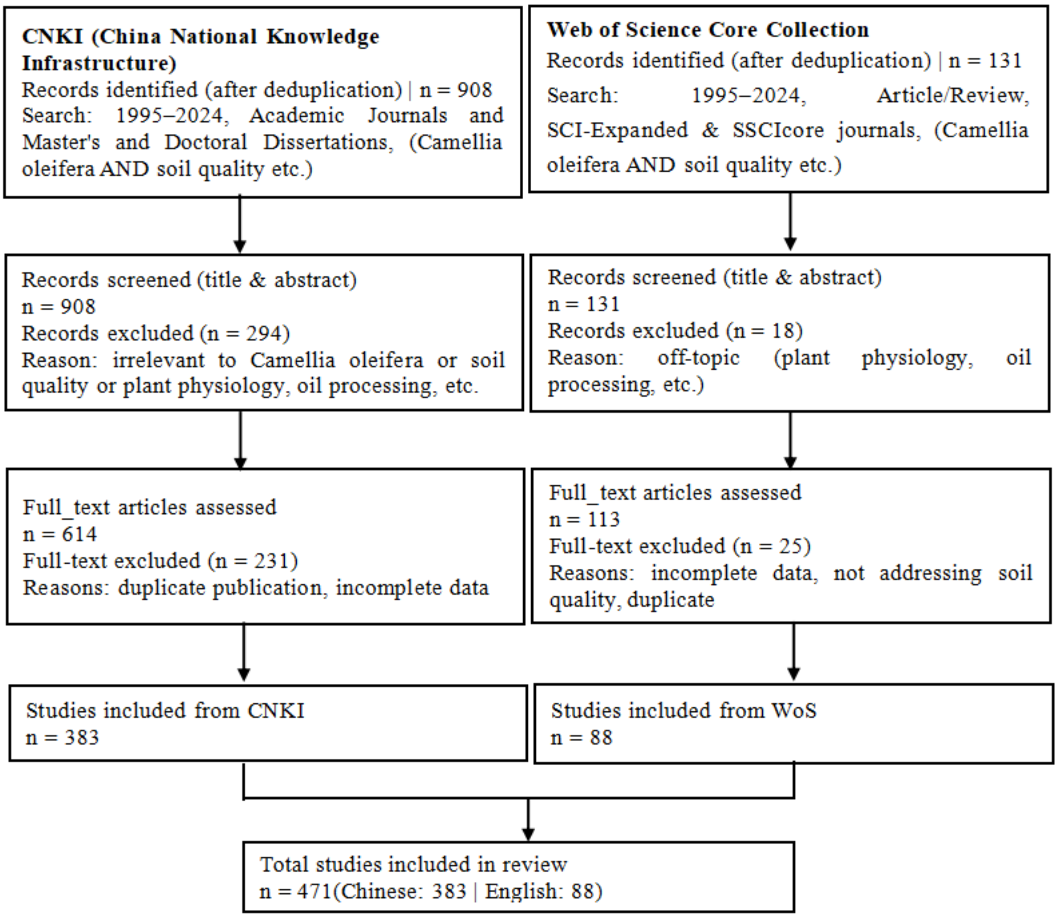

Figure 1. RISMA flowchart.

2.2. Scope and Limitations of the Study

This study conducts a bibliometric analysis of literature on soil quality in camellia oil plantations from 1995 to 2024, based on the CNKI and WoSCC databases. The following limitations exist in terms of data sources and analytical methods:

Firstly, the scope of data sources. This study selected CNKI and WoSCC as data sources, representing Chinese and English literature repositories respectively. Other databases such as Wanfang, VIP, and Scopus were not included, and the coverage of the literature is limited to the content indexed on these two platforms.

Second, language bias and its analytical consequences. Given that Camellia oleifera is endemic to China, the dataset is heavily skewed toward Chinese-language publications (approx. 79% in this study). While this reflects the reality of the field, it introduces a specific structural bias into the CiteSpace networks. Clusters derived from CNKI data are more likely to emphasize local management practices and region-specific terminology (e.g., Zijin soil), whereas the smaller WoS dataset may over-represent internationally generalizable mechanisms (e.g., Al-P interaction). Consequently, the observed divergence between applied (CNKI) and mechanistic (WoS) clusters is partially a function of database indexing bias and linguistic scope, rather than a pure reflection of differing scientific philosophies. This limitation cautions against interpreting the parallel trajectories as mutually exclusive; they are complementary views shaped by publication venue and audience.

Third, the boundaries of the analytical methods. CiteSpace knowledge graph analysis is based on keyword co-occurrence, clustering and outlier detection; its results are influenced by the selection of search terms, the standardisation of keywords and the configuration of algorithmic parameters. The generation of cluster labels is based on the log-likelihood ratio algorithm.

Fourth, the handling of cross-database comparisons. CNKI and WoS differ in terms of indexing standards, subject classification and keyword indexing methods. This study adopted a strategy of separate analysis and comparative analysis of paradigms to balance data comparability with the specificity of the research.

The above boundary conditions establish a foundational frame of reference for interpreting this study's findings. Specifically, although the search strategies-including keyword combinations (such as soil quality) were iteratively refined and validated, they inevitably shaped the composition of the final literature corpus and, consequently, the bibliometric outputs. Moreover, while this study strives to generate an objective knowledge map, certain interpretive labels introduced during initial analysis (such as paradigm shift) extend beyond what the CiteSpace-derived evidence directly supports. To ensure methodological fidelity, the authors have revised these formulations to reflect only what is empirically warranted; future studies may extend such interpretations by integrating complementary data sources and analytical approaches.

2.3. Research Tools and Analytical Methods

This study employed CiteSpace (version 6.4.R1) to perform visual analytics and construct knowledge maps based on the screened literature dataset [20]. To ensure analytical consistency and comparability, the time span was defined as 1995–2024, with a one-year interval (Years Per Slice = 1), allowing for a fine-grained representation of the temporal evolution of research in this field. The node types were specified as Author, Institution, and Keyword, corresponding to analyses of collaboration networks, institutional distribution patterns, and the structural configuration of research hotspots, respectively. In terms of node selection, the g-index algorithm (k = 25) was adopted. Compared with conventional Top N approaches, this parameterisation balances the representation of high-frequency nodes with the inclusion of medium-frequency contributions, making it particularly suitable for datasets of moderate size and improving the overall integrity of the network structure.

For network optimisation and visualisation, a combined pruning strategy incorporating Pathfinder and Pruning Sliced Networks was applied to eliminate redundant links, emphasise salient structural relationships, and enhance the interpretability of the resulting knowledge maps. Cluster View and Timeline View were used for visual representation, enabling the identification of thematic clustering patterns and their temporal evolution, respectively. Cluster labelling was conducted using the Log-Likelihood Ratio (LLR) algorithm, which improves the specificity and interpretability of cluster descriptors. In addition, the robustness and reliability of the network structure and clustering outcomes were quantitatively assessed using key metrics, including betweenness centrality, modularity Q, and silhouette S.

3. Research Findings and Analysis

3.1. Overall Trend Analysis of Research Status

3.1.1. Temporal Evolution: Characteristics of the Interannual Distribution of Research Literature

An examination of overall publication trends indicates that research in this field has progressed through three distinct phases: initial exploration, steady growth, and accelerated development, as illustrated in **Figure 2**. A comparison between Chinese-language and English-language outputs shows that publications indexed in CNKI account for 79.48% of the total, representing approximately four times the volume of those in the Web of Science database. This distribution underscores the dominant role of Chinese scholarship and confirms China as the primary centre of research on Camellia oleifera soils. Notably, Chinese-language research emerged considerably earlier than its English-language counterpart (which began around 2004) and has maintained a clear quantitative advantage throughout the entire period. The recent surge in English-language publications has been concentrated within the past five years, reflecting both the rapid advancement of China's research capacity and a shift towards broader, mechanism-oriented ecological inquiries with greater international relevance.

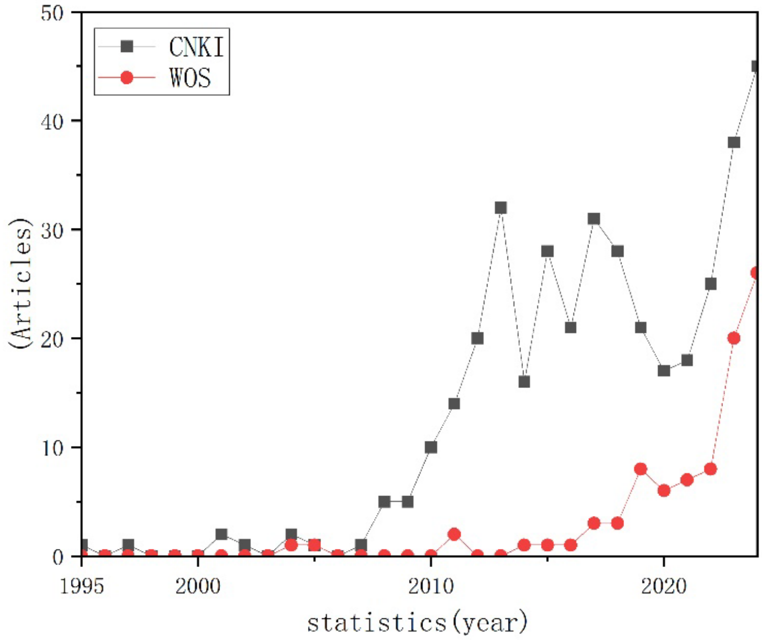

Figure 2. Statistical Chart of Research Literature Volume.

The initial exploration phase (1995–2006) was characterised by a limited number of publications in both CNKI and WoS, with an average annual output of fewer than four articles. Research topics during this period were relatively dispersed, encompassing agronomic practices such as mulching and intercropping, monitoring of soil fertility dynamics [21], and the identification of nutrient-limiting factors [22]. These studies were primarily aimed at addressing practical production challenges, with limited theoretical integration and minimal exploration of underlying mechanisms. As such, this phase represents the foundational stage of knowledge accumulation in the field.

The steady growth phase (2007–2020) coincided with increased national policy support for the Camellia oleifera industry, exemplified by the National Camellia Oil Industry Development Plan (2009–2020). During this period, research output exhibited consistent growth, with annual publications in CNKI rising from single-digit figures to several dozen per year. However, the rate of increase slowed notably after 2018, likely due to a combination of factors, including the longer timeframes required for in-depth basic research, rising standards of academic quality, and disruptions to experimental activities during the COVID-19 pandemic. English-language publications also began to emerge in the WoS database during this stage, although overall output remained modest, averaging fewer than ten articles annually. This phase marked the consolidation of Camellia oleifera soil research as a more clearly defined field of inquiry. Thematically, research became more focused, with increased attention to the effects of management practices and stand age on soil physicochemical properties. At the same time, soil acidification [23,24] emerged as a prominent research topic and gained increasing scholarly attention.

The accelerated development phase (2021–2024) is characterised by a rapid expansion in publication output, signalling a transition into a period of intensified research activity. This growth has been driven in part by the integration of advanced methodologies, including high-throughput sequencing in molecular biology and machine learning approaches in data analysis. In particular, following the implementation of the Three-Year Action Plan for Accelerating the Development of the Camellia oleifera Industry (2023–2025), research priorities have broadened to include more comprehensive investigations of biological and ecological processes, complementing traditional soil chemical analyses towards more comprehensive investigations of biological and ecological processes. Key research themes during this phase include soil ecological functions, carbon cycling [25], microbial diversity [26,27], root exudation and rhizosphere processes [27], and integrated soil quality assessment frameworks [28], especially the development and application of minimum data sets (MDS). In parallel, both the extent of international collaboration and the volume of English-language publications have increased substantially, indicating a growing alignment of this research field with global scientific frontiers.

3.1.2. Core Strength Analysis: Author and Institutional Collaboration Networks

Co-occurrence analysis of publishing institutions and authors using CiteSpace provides insight into the core research actors and their collaboration patterns in the field of soil quality in Camellia oleifera plantations. The institutional collaboration network is characterised by a pattern of overall dispersion accompanied by localized clustering. Prominent nodes include South-Central University of Forestry and Technology, the Institute of Subtropical Forestry of the Chinese Academy of Forestry, Jiangxi Agricultural University, and the Hunan Academy of Forestry Sciences. Situated within China's principal Camellia oleifera production regions, these institutions benefit from strong geographical advantages and extensive research experience. They not only contribute substantially to publication output and exhibit high betweenness centrality, but also function as key hubs for knowledge generation and network connectivity. Collaborative relationships tend to form geographically defined sub-clusters, such as the Hunan cluster centred on South-Central University of Forestry and Technology and the Jiangxi cluster centred on Jiangxi Agricultural University. Nevertheless, cross-regional and international collaboration remains limited, and the overall network density is relatively low. Although several institutes affiliated with the Chinese Academy of Sciences participate in this field, they generally occupy peripheral positions, primarily contributing theoretical and methodological support.

The author collaboration network exhibits a structure broadly similar to that of institutional collaborations, characterised by widespread dispersion interspersed with small, cohesive clusters. Within the Chinese research landscape, scholars affiliated with institutions such as South China University of Forestry and Technology, Jiangxi Agricultural University, Jiangxi Academy of Forestry Sciences, and Fujian Agriculture and Forestry University form an extensive authorship network. High-output researchers, including Hu Dongnan, Guo Xiaomin, Wang Yujuan, and Chen Yongzhong, serve as central figures, around whom several closely connected collaborative groups have emerged. These groups have made foundational contributions to key research areas, including plantation productivity, nutrient cycling, and soil fertility assessment. However, inter-group collaboration remains relatively weak, resulting in a fragmented structure characterised by several largely independent research teams. In contrast, the author network in English-language publications is more dispersed, with fewer nodes and weaker interconnections. This pattern is likely attributable to the later development of English-language research outputs and the broader geographic distribution of contributing scholars. Nonetheless, Chinese researchers constitute the dominant group within the English-language literature, reflecting the increasing international dissemination of high-quality domestic research and the growing global visibility of this field.

Overall, research on soil quality in Camellia oleifera plantations has led to the formation of a core group of researchers closely aligned with major production regions. However, the overall collaboration network remains relatively sparse, and large-scale, cross-institutional and cross-regional cooperation is still limited. From a descriptive standpoint, the network topology suggests that the field may benefit from strengthened collaborative frameworks and reduced geographical barriers to enhance knowledge integration. However, this interpretation remains speculative and extends beyond the purely descriptive capacity of the bibliometric data.

3.1.3. Regional Distribution Characteristics of Research

A comparison between the frequency of soil quality studies in camellia oil-producing provinces and actual cultivation areas (**Figure 3**) reveals a general positive correlation between research intensity and industrial scale, although significant regional imbalances persist. The three major production regions, Hunan, Jiangxi, and Guangxi, demonstrate strong alignment between research output and planting area, suggesting a relatively rational distribution of research resources. At the national level, central government funding supports camellia oil industry development through demonstration and subsidy programs administered by the Forestry and Grassland Reform and Development Fund. At the provincial level, Hunan, Jiangxi, and Guangxi have established dedicated special funds to promote research initiatives. In contrast, Fujian's research frequency (9.02%) substantially exceeds its share of national planting area (3.52%), indicating robust academic capacity but relatively underdeveloped industrial scaling. Conversely, provinces such as Hubei and Yunnan exhibit research investment that lags behind production scale. Hubei, with 6.40% of total planting area, accounts for only 1.95% of published studies, revealing a pronounced deficit in research support. Hainan also shows a severe mismatch between research activity and cultivation scale. In response, both national and provincial authorities are advancing efforts to establish dynamic mechanisms for reallocating research funding. These initiatives aim to facilitate cross-provincial transfer of scientific achievements, improve the industrial application rate of research outcomes, and implement performance evaluation systems based on planting area and output value thereby guiding research resources to more effectively meet practical production demands.

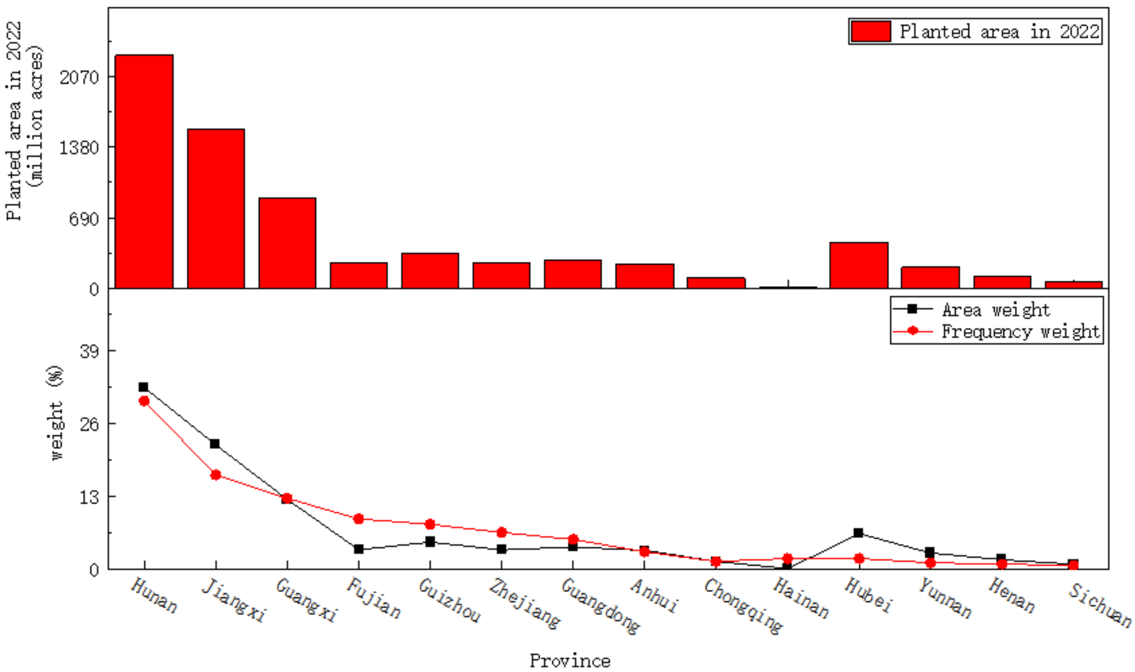

Figure 3. Analysis of Research Frequency and Planting Area by Province.

Source: On January 5, 2023, the National Forestry and Grassland Administration, the National Development and Reform Commission, and the Ministry of Finance jointly issued the Three-Year Action Plan for Accelerating the Development of the Camellia Oil Industry (2023–2025).

3.2. Research Topic Clustering and Evolutionary Path Analysis

3.2.1. Keyword Co-Occurrence Network and Cluster Analysis

Based on bibliometric analyses of the CNKI and

WoS databases, the key parameters of the knowledge graph networks derived in this study are presented in **Table 1**.

Table 1. Knowledge graph network parameter settings for Citespace.

Parameter	CNKI	WoS
Nodes	410	251
Edges	888	859
Density	0.0106	0.0274
Modularity Q	0.6445	0.6942
Weighted Mean Silhouette S	0.8776	0.8735

As shown in **Table 1**, the network modularity coefficient (Q) exceeds 0.3, indicating a well-defined community (cluster) structure. The weighted mean silhouette coefficient (S) is greater than 0.7, suggesting high internal consistency and strong separation among clusters, thereby confirming the reliability of the clustering results. Taken together, these metrics demonstrate that research on soil quality in Camellia oleifera plantations has developed into several distinct thematic domains with clear structural boundaries, as illustrated in **Figures 4–7**. The most frequently occurring keywords further reflect the core knowledge structure of the field. In the Chinese-language literature, dominant terms include Camellia oleifera, soil fertility, soil nutrients, Camellia oleifera plantations, soil, intercropping, and mixed cropping. In the English-language literature, high-frequency keywords include diversity, nitrogen, carbon, growth, and phosphorus. Collectively, these terms define the principal research framework of soil studies in Camellia oleifera plantations.

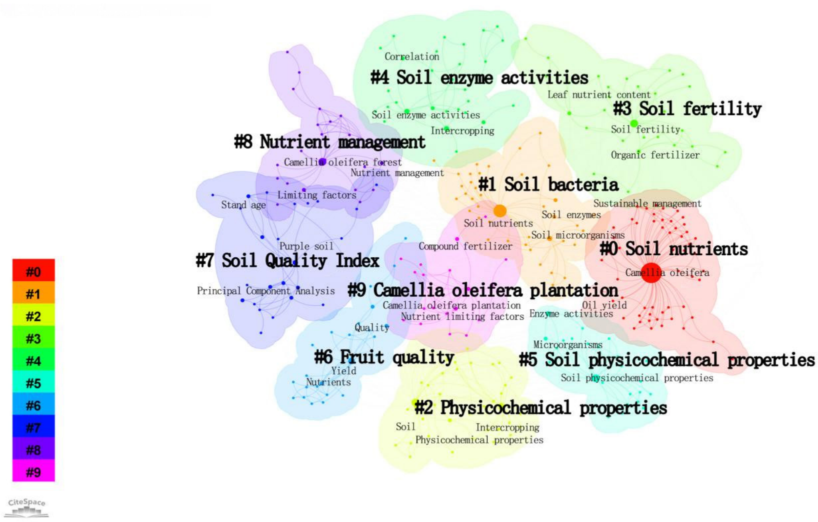

Figure 4. CNKI keyword clustering analysis.

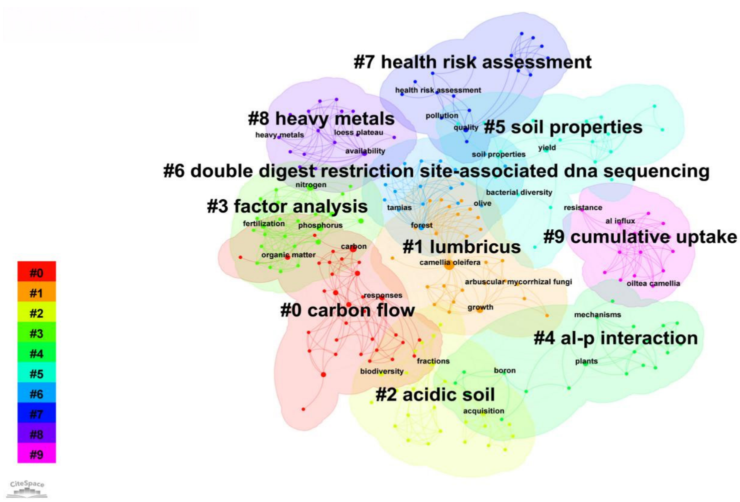

Figure 5. WOS keyword clustering analysis.

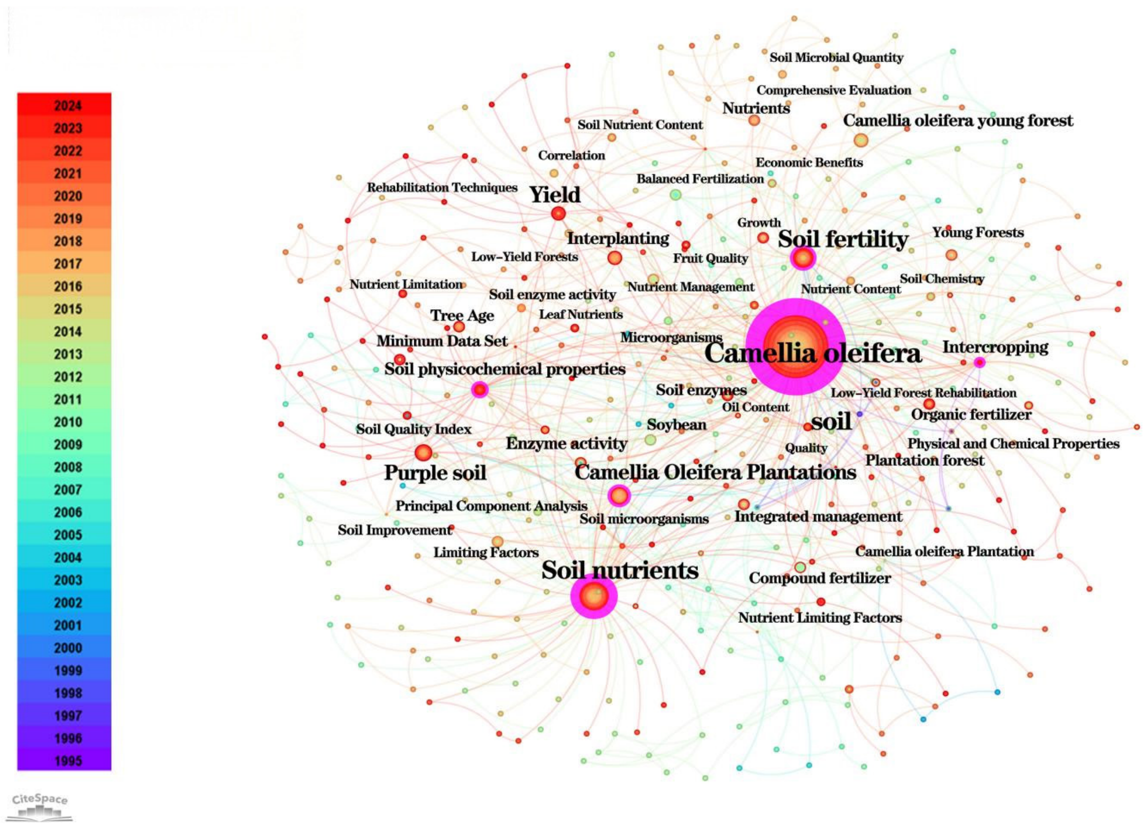

Figure 6. CNKI Keyword Contribution Network Distribution Map.

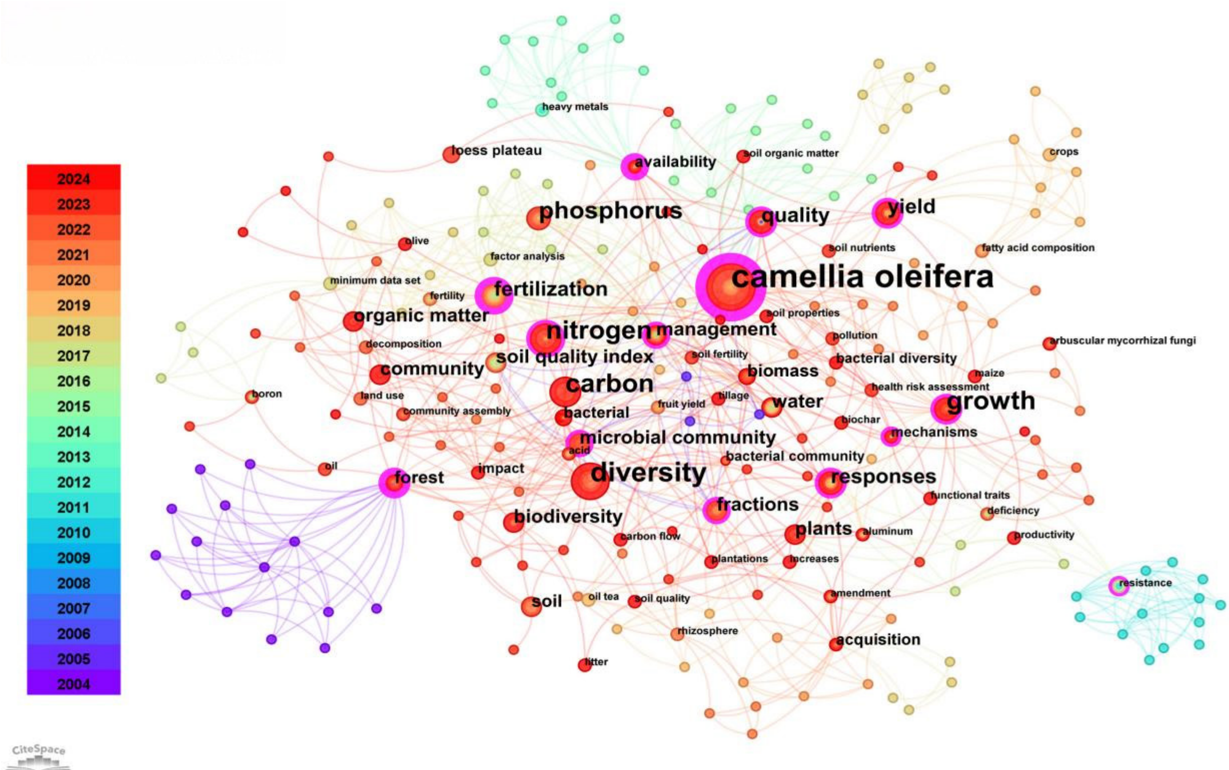

Figure 7. WOS Keyword Contribution Network Distribution Map.

An in-depth comparison of cluster labels and keyword contributions (**Figures 4–7** and **Table 2**) reveals several salient characteristics:

(1) Divergence in research perspectives

The CNKI clusters exhibit a strong application-oriented and management-driven perspective. Clusters such as #3 (soil fertility), #4 (soil enzyme activity), and #5 (soil physicochemical properties) are organised around keywords including nutrient management, organic fertiliser, intercropping, yield, and quality. This indicates that research is closely aligned with the practical needs of the Camellia oleifera industry, with a primary focus on enhancing plantation productivity and fruit quality through management interventions. The content is highly context-specific, as reflected by keywords such as camellia forest land, compound fertiliser, and Zijin soil, which highlight a strong regional orientation and applied focus.

By contrast, the WoS clusters demonstrate a pronounced emphasis on mechanistic understanding and fundamental science. Clusters including #2 (acidic soil), #4 (Al-P interaction), #5 (soil properties), and #6 (double-digest restriction site-associated DNA sequencing) focus on soil chemical processes, nutrient transformations, and molecular biological approaches, indicating a deeper exploration of underlying mechanisms. Frequently occurring keywords such as bacterial diversity, acquisition, mechanisms, and resistance further suggest an emphasis on soil microbial community structure, nutrient uptake processes, and plant adaptive strategies.

(2) Divergence in research focus

The CNKI clusters are primarily centred on the relationship between soil fertility and crop performance. High-frequency keywords such as yield, quality, oil yield, and fruit quality recur consistently, and strong linkages between clusters #3 (soil fertility) and #5 (soil physicochemical properties) indicate that improving soil conditions to enhance productivity and product quality constitutes the central research objective. Although terms such as soil microorganisms and soil enzyme activity are present, they are generally treated as indicators of soil fertility rather than as focal points of independent ecological process research.

In contrast, the WoS clusters reveal a broader and more integrative research focus. Clusters such as #0 (carbon flow), #7 (health risk assessment), and #8 (heavy metals) indicate that international studies place greater emphasis on ecological functions and environmental risks associated with soil systems. This reflects a more comprehensive perspective on soil sustainability and environmental health, consistent with the increasing prominence of the soil health paradigm in global soil science.

(3) Differences in methodological approaches and application domains

The CNKI clusters are largely dominated by conventional analytical methods, including soil physicochemical measurements, principal component analysis, and correlation analysis. These approaches are well-established and stable, with a clear problem-oriented focus, primarily addressing specific management practices such as intercropping and fertilisation regimes.

In contrast, the WoS clusters incorporate more advanced and diverse methodologies. The presence of clusters related to high-throughput sequencing (#6), along with topics such as arbuscular mycorrhizal fungi and micronutrients (e.g., boron), indicates a stronger reliance on cutting-edge techniques and a greater emphasis on microscale processes. This reflects a distinctly method-driven research paradigm with higher levels of technical sophistication.

(4) Degree of integration within the knowledge structure

The WoS clusters exhibit strong interconnections across thematic domains. For instance, cluster #2 (acidic soil) shows close associations with #4 (Al-P interaction), while #5 (soil properties) overlaps conceptually with #8 (heavy metals). Such linkages suggest a highly integrated knowledge structure, characterised by multidimensional synthesis driven by specific research questions.

By comparison, the CNKI clusters are more concentrated and clearly delineated, with relatively distinct thematic boundaries. They form a central research trajectory centred on soil fertility-physicochemical properties-enzyme activity-yield and quality, reflecting a more linear and application-oriented knowledge framework with comparatively limited systemic integration.

Based on clustering and temporal evolution analyses (**Figures 4–9**), domestic and international studies on Camellia oleifera soils-though convergent in their shared emphasis on yield-related outcomes-exhibit markedly divergent research paradigms.

Table 2. Comparative Analysis of Keyword Contributions in CNKI and WoS.

Comparison Criteria	CNKI Network Nodes	WoS Network Nodes
Research Focus	Focusing on the industrial applications of the tea oil tree (Camellia oleifera), with an emphasis on yield, quality, the rehabilitation of low-yielding plantations, and soil improvement	Focusing on ecology and mechanisms, with an emphasis on nutrient cycling, microbial communities, functional traits, and the carbon cycle
Key Terms	Yield, Soil Nutrient Content, Growth, Low-Yield Forests, Organic fertilizer, Fruit Quality, Nutrient Limitation	yield, quality, soil nutrients, phosphorus, organic matter, microbial community, biodiversity, carbon, functional traits
Research Scale	Focusing on the stand level, with an emphasis on the relationship between silvicultural practices (such as intercropping and fertilisation) and yield and quality	Focusing on the ecosystem scale, with an emphasis on ecological processes such as biodiversity, carbon fluxes and functional traits
Time Span	1995–2024: Significant clustering structure (Q = 0.6445), with research themes relatively concentrated	2004–2024: higher network density (0.0274) and closer keyword associations

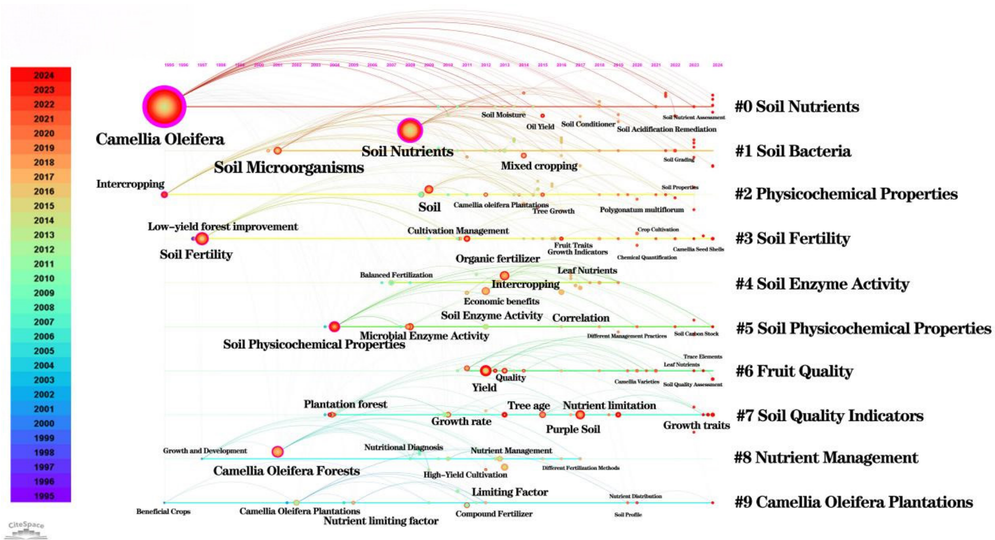

Figure 8. Temporal Evolution of Camellia oleifera Soil Research in the CNKI Database.

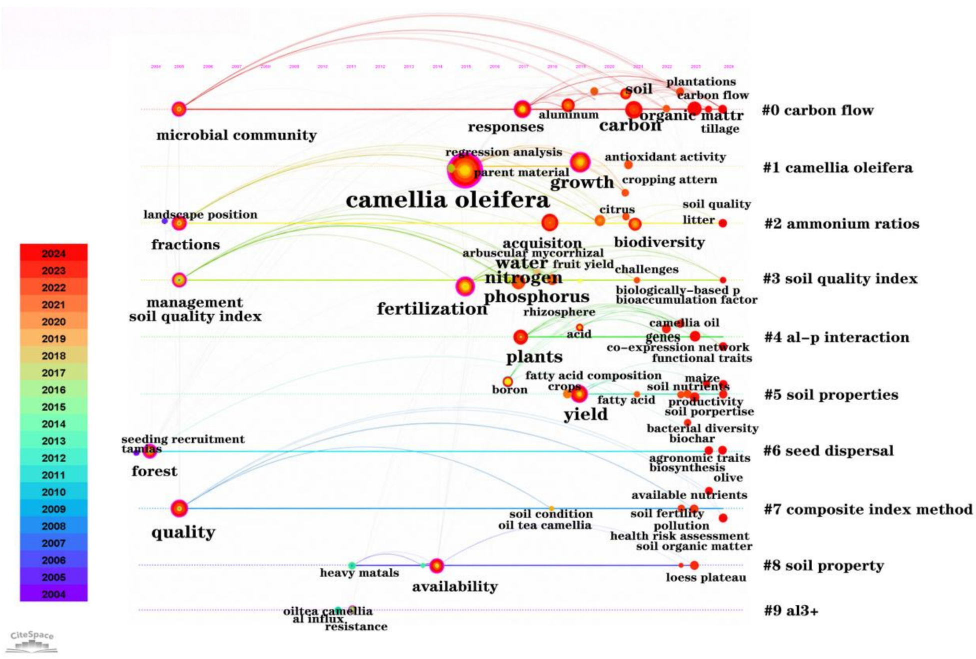

Figure 9. Temporal Evolution of Camellia oleifera Soil Research in the WOS Database.

Chinese-language literature (indexed in CNKI) formed 10 coherent thematic clusters, following a tightly coupled, linear evolutionary trajectory. In the early phase (2004–2010), research centered on Cluster #2 (soil physicochemical properties) and Cluster #3 (soil fertility), establishing foundational knowledge for fertility-based assessment frameworks. During the middle phase (2010–2015), Clusters #4 (soil enzyme activity) and #1 (soil bacterial communities) gained prominence, signaling an expansion from chemical to biological indicators of soil health. In the later phase (2015–2023), Clusters #6 (fruit quality) and #8 (nutrient management) emerged, strongly associated with practice-oriented keywords such as intercropping and organic fertilizer. This temporal pattern highlights an evolution of analytical scope. While earlier work was largely confined to descriptive characterization of soil status, recent studies have expanded to encompass process-oriented regulation of soil functions, aiming to simultaneously enhance yield and quality.

In contrast, English-language literature (indexed in Web of Science) also yielded 10 clusters, but displayed a pattern of diversified, parallel development, with pronounced emphasis on ecological processes and environmental impact assessment. Clusters #0 (carbon flow), #1 (Camellia oleifera), #2 (ammonium ratios), #3 (soil quality index), and #4 (Al-P interaction) emerged nearly concurrently during 2004–2006 and remained active through 2020, indicating that multiple, coexisting thematic lines were established early in the international research agenda. Notably, Clusters #4 (Al-P interaction) and #9 (Al^{3+} dynamics), key mechanisms governing phosphorus availability and plant tolerance in acidic soils [29] persisted across the entire study period, underscoring sustained attention to soil chemical processes and plant-soil adaptation mechanisms [30]. The more recent emergence of Cluster #7 (composite index methodology) reflects growing efforts to systematize quantitative soil quality assessment, marking an expansion of the analytical framework to include integrative, multi-dimensional evaluation alongside single-factor analysis. Collectively, this complementary technology-mechanism framework not only mirrors region-specific agronomic and environmental priorities but also illustrates the current landscape for cross-scale, integrated soil-plant system research.

3.2.2. Evolutionary Pathways of Research Hotspots: Keyword Emergence Analysis

Emerging terms-defined as concepts that exhibit a statistically significant and sustained increase in co-occurrence frequency within a given time window-serve as robust indicators of emerging research frontiers and evolving disciplinary priorities. Comparative analysis of emergence intensity, temporal persistence, and thematic succession across the two datasets (CNKI vs. WOS) reveals systematic differences in focal emphases and developmental trajectories between Chinese and English-language scholarship on Camellia oleifera soils. These findings are presented in **Figures 10** and **11**.

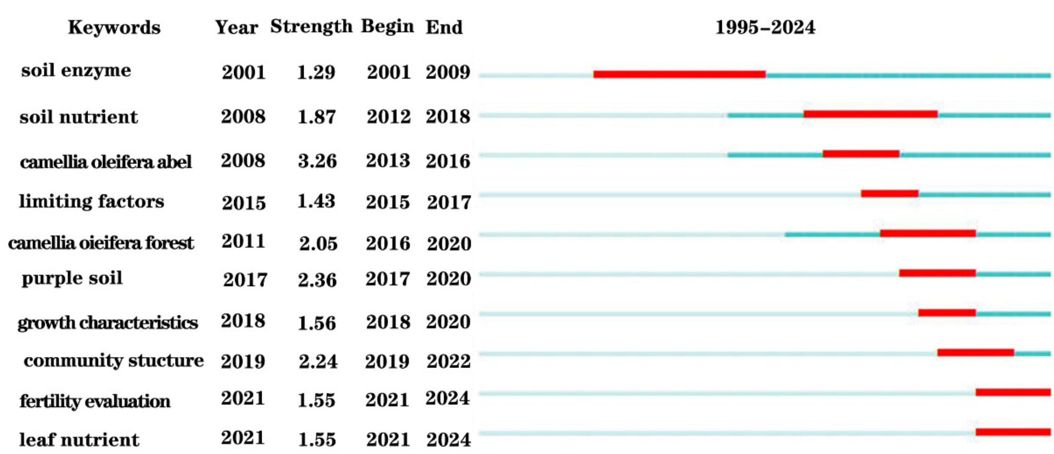

Figure 10. Top 10 Highlighted Keywords with the Strongest Citation Bursts (CNKI).

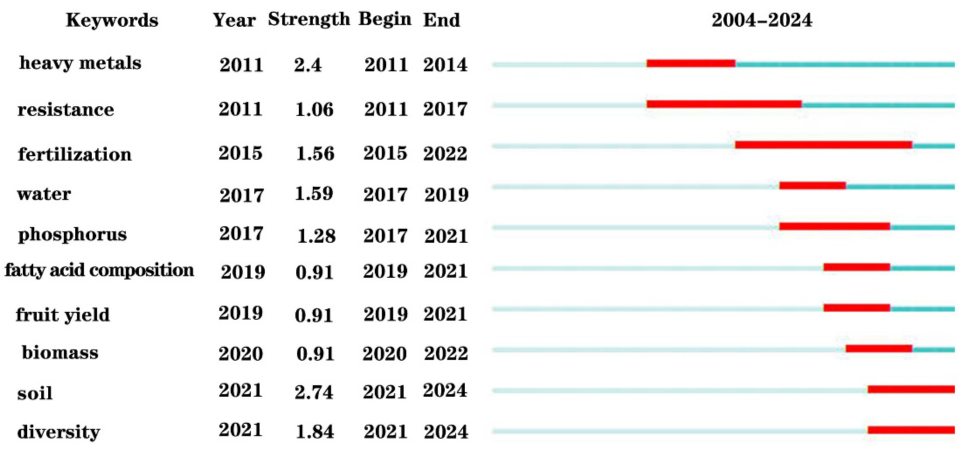

Figure 11. Top 10 Highlighted Keywords (WOS).

(1) Temporal dynamics of emerging terms: Chinese research exhibits phased hotspot transitions, whereas English-language research demonstrates thematic persistence

The temporal distribution of emerging terms in the CNKI database reveals three distinct, non-overlapping phases of thematic succession. From 2001 to 2009, soil enzyme emerged as the earliest indicator term-signalling an initial emphasis on soil biochemical properties as proxies for soil health. During 2012–2020, a cluster of terms-including soil nutrient, Camellia oleifera forest, purple soil, and growth characteristics gained prominence, reflecting a deliberate pivot toward applied questions in forest nutrient management and the characterization of regionally dominant soil types (e.g., purple soils in Southwest China). Most recently (2019–2024), terms such as community structure, fertility evaluation, and leaf nutrient emerged sequentially, indicating a methodological deepening into microbial ecology, integrative soil fertility assessment, and plant-based nutrient diagnostics. Collectively, these shifts occur at regular intervals of approximately 5-8 years, forming a rhythmically structured evolution tightly coupled with national industrial development milestones (e.g., the Camellia oleifera Industry Development Plan) and evolving policy priorities (e.g., ecological forestry initiatives).

In contrast, emerging terms in the Web of Science (WoS) database display marked thematic continuity. Resistance remained emergent from 2011 to 2017; fertilisation persisted from 2015 to 2022; and phosphorus exhibited sustained emergence between 2017 and 2021, each spanning 5–8 years. Critically, the concurrent emergence of soil and diversity during 2021–2024 does not represent a departure from prior themes but rather an integrative extension of long-standing commitments to fundamental soil biogeochemical processes and ecosystem-level functional attributes. This enduring focus underscores a research culture prioritizing mechanistic depth and cross-system generalizability over rapid topical turnover.

(2) Thematic orientation: Chinese scholarship is application-driven; English-language scholarship is mechanism-driven

CNKI emerging terms consistently reflect an applied, management-oriented agenda. Terms such as Camellia oleifera Abel (synonym for tea-oil tree), Camellia oleifera forest, and purple soil anchor research to concrete agroforestry systems and geographically specific edaphic conditions-highlighting strong regional applicability and direct alignment with domestic industrial needs. Further, terms including growth characteristics, leaf nutrient, and fertility evaluation explicitly target operational objectives: rehabilitating low-yield stands, optimizing site-specific fertilisation, and monitoring soil fertility trajectories. Notably, limiting factors emerged with high intensity (1.43) during 2015–2017, revealing a time-limited, problem-driven effort to identify soil constraints on Camellia oleifera productivity-yet this line of inquiry did not consolidate into a sustained thematic trajectory.

By contrast, WoS emerging terms coalesce around universal biophysical mechanisms. Heavy metals, resistance, phosphorus, and diversity collectively orient research toward foundational questions in soil contamination ecology, plant stress physiology, nutrient biogeochemistry, and microbial community assembly. The emergence of fatty acid composition and fruit yield further illustrates how English-language studies extend soil-focused inquiry to downstream physiological and biochemical mechanisms governing oil quality and reproductive output without shifting away from mechanistic explanation. Even the transient emergence of water (2017–2019), though modest in intensity and duration, was framed within mechanistic contexts such as rhizosphere water retention dynamics and drought-induced metabolic adjustments, rather than irrigation scheduling or field-level water conservation practices.

(3) Evolution of analytical depth: Chinese research advances from indicators to integrated soil–plant systems; English-language research progresses from isolated factors to holistic system integration

CNKI emerging terms trace a clear trajectory from discrete biochemical markers toward integrative, multi-compartment soil–plant system analysis. Early emergence of soil enzyme (2001–2009) reflects reliance on single-process biochemical indicators; the subsequent rise of soil nutrient (2012–2020) broadened scope to include major macronutrient dynamics. Most recently (2019–2024), the concurrent emergence of community structure, fertility evaluation, and leaf nutrient signals a paradigmatic consolidation shifting from fragmented indicator measurement to unified diagnostic frameworks that jointly assess microbial composition, soil functional capacity, and plant nutritional status. This progression substantiates the construction of a China-specific research architecture centred on fertility evaluation and explicitly spanning biogeochemical interfaces among soil, rhizosphere microbiota, and host physiology.

WoS emerging terms follow a complementary trajectory from reductionist factor analysis toward synthesis across physical, chemical, and biological domains. Initial emphasis on heavy metals and resistance (2011–2017) addressed discrete stressors and organismal responses; mid-phase expansion into fertilisation, phosphorus, and water (2015–2021) incorporated resource availability and management dimensions. Critically, the simultaneous emergence of soil and diversity (2021–2024) marks a conceptual pivot: soil is no longer treated as a passive medium but as an active, multi-dimensional system whose integrity depends on the interplay of abiotic properties and biotic complexity. With an emergence intensity of 2.74, the highest observed across all time windows soil now serves as the unifying anchor for integrative soil science in English-language literature, signifying maturity in systemic thinking.

(4) Patterns of interdisciplinary integration and adoption of frontier methodologies

In CNKI literature, community structure emerged during 2019–2022, coinciding with the methodological uptake of high-throughput sequencing to characterize Camellia oleifera rhizosphere microbiomes. While adoption of molecular tools has been gradual, its accelerating integration signals a decisive transition from descriptive microbiology toward mechanism-informed soil management.

In WoS literature, diversity emerged robustly during 2021–2024 (intensity = 1.84), reflecting international consensus on microbial diversity as a functional proxy not merely a taxonomic descriptor for soil health and ecosystem service delivery. Concurrently, the sustained seven-year emergence of resistance (2011–2017) underscores persistent engagement with plant stress adaptation mechanisms particularly aluminium toxicity and heavy metal exclusion in acidic soils, a theme deeply embedded in long-standing biogeochemical research agendas on nutrient bioavailability and rhizosphere chemistry.

4. Discussion

4.1. Thematic Reconfiguration and the Expansion of Analytical Scope

The temporal distribution of publications and the clustering timeline (**Figures 8** and **9**) provide a structural basis for understanding how research priorities have been reconfigured over time. Analysis of the Timeline View reveals a notable pattern in cluster persistence and connectivity. Early clusters, such as #2 (physicochemical properties) in CNKI, exhibit long, continuous time spans (2004–2020) with relatively sparse inter-cluster connections. This structural feature suggests that early research was characterized by isolated, parallel investigations of discrete soil properties, with limited cross-theme integration.

In contrast, clusters emerging in the later period (post-2018), such as #1 (soil bacterial) and #6 (fruit quality), show shorter individual durations but denser interconnections with other clusters (as indicated by overlapping timelines and increased cross-cluster citation links). This topological shift in the knowledge map from sparse, persistent silos to dense, interconnected modules supports the interpretation that the field has moved from fragmented characterization toward integrative, system-level inquiry.

This structural evolution aligns with the three broad phases identified by publication volume trends (Section 3.1.1): an initial phase of establishing foundational indicators (sparse network), a growth phase of expanding indicator categories (increasing node diversity), and a recent phase of network densification, reflecting the convergence of soil chemistry, microbiology, and plant performance studies.

4.2. Structural Divergence and Complementarity between Chinese and International Scholarship

Beyond differences in cluster labels, comparative analysis of network topology between CNKI and WoS datasets reveals fundamental structural distinctions that underpin the observed thematic divergence. As shown in **Table 1**, the CNKI keyword network exhibits a higher overall density and tighter intra-cluster cohesion, but with fewer bridging nodes of high betweenness centrality. In practical terms, this means CNKI research clusters are internally well-integrated but relatively isolated from one another, forming distinct, application-focused silos (e.g., fertilization, yield, soil properties).

Conversely, the WoS network, while smaller in node count, displays a higher proportion of nodes with elevated betweenness centrality particularly terms like diversity, phosphorus, and carbon. These terms function as conceptual bridges, structurally linking clusters related to soil chemistry, microbial ecology, and plant physiology. This topological feature explains why WoS literature appears more mechanism-driven: the network architecture itself facilitates cross-disciplinary integration, positioning key biogeochemical concepts as central hubs through which diverse research threads pass.

These structural differences are not merely incidental, they reflect the distinct publication ecosystems. The dense but siloed CNKI structure mirrors a research landscape organized around solving region-specific production problems, whereas the high-centrality WoS structure reflects an emphasis on generalizable, transferable mechanisms. Recognizing this topological basis helps avoid over-interpreting thematic differences as purely philosophical, they are partially encoded in the network architecture shaped by database scope and audience.

4.3. In-Depth Analysis of Research Focus: From CiteSpace Clustering to the Deepening of Scientific Questions

4.3.1. Mechanisms of Soil Development and Soil Acidification: From Mechanistic Analysis to Synergistic Remediation

Before examining specific empirical findings, the structural properties of the burst detection results (**Figures 10** and **11**) provide critical framing for the discussion. Analysis of burst strength and temporal persistence reveals three key patterns:

(1) CNKI Burst Profile: The term limiting factors exhibited a short, high-intensity burst (Strength = 1.43, 2015–2017), indicating a concentrated but transient focus on diagnosing acidification constraints. In contrast, community structure shows a more recent and sustained burst (2019–2024), suggesting a methodological pivot toward microbial ecology as a lens for understanding acidification impacts.

(2) WoS Burst Profile: The term Al (aluminum) demonstrates remarkable burst persistence spanning nearly the entire study period (2004–2020), with consistently high strength values. This indicates that aluminum toxicity, a direct consequence of acidification, has been a structurally embedded, long-term research priority in international literature, not a passing trend.

(3) Cross-Database Comparison: The asynchronous burst patterns between CNKI and WoS where CNKI bursts reflect applied diagnostic interests and WoS bursts reflect persistent mechanistic inquiry topologically reinforce the complementary division of labor

discussed in Section 4.2.

Grounded in these bibliometric signatures, the following review of empirical acidification research is organized to reflect both the transient diagnostic focus in Chinese literature and the sustained mechanistic focus internationally.

Soil acidification in Camellia oleifera plantations driven primarily by prolonged nitrogen fertilization is escalating as the predominant biogeochemical constraint on industrial sustainability, widely characterised in domestic policy and technical guidelines as the primary limiting factor for yield stability and quality consistency. Bibliometric analysis of emerging terms (CiteSpace) documents a pronounced surge in scholarly attention to acidification, aluminium toxicity, and microbial diversity since 2018. Empirical research has evolved systematically: from initial spatial surveys quantifying pH decline and exchangeable Al^{3+} accumulation, to mechanistic investigations into how acidification restructures soil microbial communities, suppresses expression of key functional genes (e.g., nirK, amoA), and inhibits extracellular enzyme activities (e.g., phosphatase, urease) [24–26]. Concurrently, the efficacy of soil remediation using amendments such as lime, biochar and organic fertilisers, as well as intercropping and relay cropping, and the underlying microbiological mechanisms have become a research focus [31–34].

The empirical studies catalogued below are not a generic collection of acidification research; rather, they were selected because they exemplify the thematic clusters and high-centrality keywords identified in the CiteSpace networks:

(1) Cluster #3 (Soil Fertility) and Keyword Organic Fertilizer (CNKI): The composite microbial fertilizer study by Huang et al. [31], which reported a 20.7% increase in soil organic carbon, directly instantiates the applied focus of this cluster. In the CNKI network, organic fertilizer exhibits strong co-occurrence links to both yield and soil acidification, positioning it as a key intervention node.

(2) Cluster #2 (Acidic Soil) and Keyword Al-P Interaction (WoS): The research on legume intercropping [32,33] exemplifies the mechanistic depth of this WoS cluster. The high betweenness centrality of phosphorus in the WoS network (**Figure 7**) explains why studies investigating synergistic alleviation of Al toxicity and P deficiency are structurally central to the international research agenda.

By mapping these empirical examples back onto the network topology, we observe a coherent pattern: CNKI studies tend to cluster around application-oriented intervention nodes, while WoS studies gravitate toward mechanistic interaction hubs. This structural mapping provides a principled, data-driven way to organize the literature review, rather than relying on arbitrary thematic categorization. Collectively, these interventions converge on a central scientific challenge: How can productivity be sustainably enhanced while concurrently strengthening the soil's intrinsic acid-buffering capacity? Addressing this requires a paradigm shift from empirical screening of individual amendments toward systematic investigation of synergistic organic–inorganic microbial co-regulation. Two critical knowledge gaps impede progress: First, the quantitative attribution of acidification drivers remains unresolved. Current studies rely predominantly on qualitative observations or single-factor controlled experiments, lacking robust field-based quantification of the relative contributions of long-term nitrogen fertilisation, atmospheric acid deposition, rhizosphere proton extrusion, and mineral weathering to cumulative soil acidity. This attribution uncertainty hampers cost–benefit optimisation of mitigation strategies. Future work must integrate stable isotope tracing, multi-decadal fixed-point monitoring networks, and process-based biogeochemical modelling to disentangle and quantify source-specific acidification fluxes. Second, the bidirectional feedback between acidification and soil biological functions is incompletely characterised. Specifically, how does acidification-induced restructuring of microbial communities such as suppression of diazotrophic bacteria and enrichment of acid-tolerant fungi alter functional gene expression, disrupt carbon and nitrogen transformation rates, and ultimately erode or reconstitute soil acid-base buffering capacity? Elucidating this microbially mediated biogeochemical feedback loop is essential for designing next-generation, microbiome-informed acidification remediation strategies.

4.3.2. Minimum Data Set (MDS) Construction: From Indicator Screening to Dynamic Diagnosis

Clustering and temporal evolution analyses identify soil quality assessment and minimum data set (MDS) as the dominant thematic clusters (**Figure 4**, Cluster #7), notably, the keyword co-occurrence network reveals a strong link and high betweenness centrality for Principal Component Analysis (PCA) [35], acting as a methodological bridge between Soil Nutrients and Minimum Data Set [36,37]. This structural feature suggests that PCA is not merely one of many statistical tools but a pivotal methodological pathway through which fragmented soil property data are integrated into cohesive evaluation frameworks. Furthermore, the time-sliced visualization shows that MDS emerges as a prominent node only after 2018, coinciding with a decline in burst strength for singular indicators like available phosphorus. This shift in the knowledge map supports the narrative that the field is moving from single-factor monitoring toward multivariate index construction. However, the low silhouette score within this cluster indicates considerable heterogeneity in indicator selection, which directly aligns with the empirical challenge discussed below: the tension between regional specificity and universal applicability.

However, the clustering metrics also reveal an inherent tension. While the modularity (Q) is high, the Silhouette score for the MDS cluster is relatively moderate ($S \approx 0.75$), suggesting that while this research area is distinct, its internal homogeneity is not absolute. This bibliometric ambiguity mirrors the empirical challenge documented in the literature: the severe regional heterogeneity that undermines indicator universality. During bibliometric and statistical synthesis, the research team systematically catalogued drivers of soil development and quality degradation in Camellia oleifera cultivation zones. These drivers operate across two interdependent domains: the internal environment [38] encompassing soil physical structure, chemical composition, and biological activity, all jointly governing soil functional integrity; and the external environment [39] including climate, topography, native vegetation, and anthropogenic management practices. Grounded in this dual-domain framework, a scientifically robust MDS for Camellia oleifera soils must prioritise high-impact, context-specific stressors evidenced by bibliometric prominence: red soil acidification, nutrient imbalances, structural degradation and heavy metal accumulation. The high-frequency co-occurrence of MDS and PCA in the network diagram reflects the mainstream approach currently adopted in research regarding indicator selection. Based on this, researchers have developed five commonly used evaluation models [40–44], each integrating multi-dimensional physical, chemical, and increasingly biological indicators while balancing scientific rigour with regional applicability. Most established MDSs comprise 5–9 core indicators; consistently prioritised metrics include soil organic carbon, available potassium, and available phosphorus. Some regionally calibrated models further incorporate micronutrients (Fe, B, Zn, Cu) or biological proxies, The differences in the number of indicators across these models (ranging from 5 to 9) and the universality of core indicators such as organic matter and readily available potassium serve to highlight the inherent tension between the concepts of universality and specificity in the framework; these challenges are also reflected within the framework itself. Two fundamental challenges constrain MDS advancement: First, severe regional heterogeneity undermines indicator universality. MDS configurations are inherently site-specific no single standard applies across agroecological zones. For instance, in north-western Guangxi, available zinc (AZn) and available boron (AB) were designated as essential MDS components due to widespread deficiency, whereas these elements show negligible limitation in acidic red soils of southern Hunan. A scalable solution requires a tiered MDS architecture: a universal core layer supplemented by modular, context-activated tiers. Second, persistent physicochemical bias limits biological integration. Although biological indicators including microbial alpha-diversity, abundance of functional genes, and extracellular enzyme activities offer superior sensitivity to early-stage soil degradation and functional resilience, their adoption remains constrained by analytical cost, methodological variability, and lack of standardised protocols. Future research may draw inspiration from the emergence of keywords such as machine learning in the field of soil mapping, and attempt to construct a 'dynamic MDS model' capable of reflecting the spatiotemporal dynamics of soil quality.

4.3.3. Soil Microecosystem Regulation: Advancing from Taxonomic Description to Functionally Targeted Intervention

In Camellia oleifera cultivation systems, soil quality governs root architecture and seed yield through tightly coupled physicochemical–microbial feedback loops [25]. Bibliometric analyses including thematic clustering, temporal evolution, and keyword burst detection reveal a pronounced scholarly pivot toward soil microecology since 2018, with high-impact terms encompassing soil enzymes, microbial genomics, community assembly, soil C/N cycling, and high-throughput sequencing. Empirical evidence confirms that aluminium toxicity in acidic soils directly impairs root meristem activity and cortical cell elongation; however, phosphate fertilisation mitigates this inhibition via Al-P co-precipitation, thereby restoring root growth capacity and hydraulic conductivity [45]. Nutrient availability (N, P, K) further modulates root phenotypic plasticity including lateral root density, root hair length, and exudation profiles which collectively determine nutrient foraging efficiency. Critically, functional microbial guilds such as phosphate-solubilising bacteria and diazotrophic bacteria enhance this plasticity by secreting organic acids, siderophores, and phytohormones, thereby facilitating symbiotic nutrient acquisition and systemic stress resilience. Concurrently, microbial biofertilisers improve soil structural stability by stimulating aggregation through increased glomalin-related soil protein (GRSP) production and extracellular polymeric substance (EPS) secretion, ultimately enhancing root drought tolerance [46].

Persistent knowledge gaps impede translational progress: First, the principles governing microbial community assembly and functional redundancy in Camellia oleifera rhizospheres remain inadequately resolved. Within these complex networks, which taxa serve as keystone species exerting disproportionate influence on nutrient transformation rates, enzyme kinetics, or pathogen suppression? How do community composition, interaction topology, and functional gene expression collectively determine system-level stability and recovery after disturbance? Without mechanistic clarity on these foundational questions, precise, predictive microecological regulation remains elusive. Second, field-scale efficacy of functional microbial inoculants suffers from poor ecological fidelity. Laboratory-isolated strains despite high in vitro activity frequently exhibit low rhizosphere colonisation success, rapid population decline, and context-dependent functional attenuation under field conditions. Emerging solutions centre on Synthetic Microbial Communities (SynComs): rationally designed consortia integrating complementary functional traits with enhanced environmental robustness validated through gnotobiotic rhizosphere assays and multi-year field trials [47,48].

5. Conclusions and Forward-Looking Research Agenda

5.1. Synthesised Conclusions

This bibliometric mapping descriptively synthesizes the thematic evolution of soil quality research in Camellia oleifera plantations over the past three decades (1995–2024), revealing a coherent epistemic trajectory: from elemental characterization to process-oriented management and then to system-integrated regulation. Specifically: First, the Resource Baseline Phase (1995–2010) prioritised descriptive soil fertility assessments quantifying pH, organic carbon, available nutrients, and texture to establish foundational inventories for regional industrial planning. Research remained largely indicator-centric, with minimal integration across biogeochemical compartments. Second, the Process Optimisation Phase (2010–2018) shifted focus toward agronomic intervention efficacy: evaluating how fertilisation regimes, tillage practices, and intercropping systems modulate soil functional capacity. This phase advanced beyond single-indicator reporting to multi-dimensional evaluation frameworks incorporating chemical, physical, and early biological metrics. Third, the System-Integrated Regulation Phase (2019–2024) is characterized by a thematic focus on soil as a dynamic, adaptive biogeochemical system evaluated through integrated diagnostics and regulated via feedback-informed, precision strategies targeting yield, quality, and ecological resilience simultaneously. Critically, Chinese (CNKI) and English-language (WoS) scholarship follow structurally complementary trajectories constituting a globally coherent yet contextually grounded knowledge architecture. CNKI research advances an application-driven, industry-aligned framework: progressing from site-specific indicator validation to integrated soil–plant–microbe diagnostics, with increasing emphasis on delivering scalable farm-to-fork technical protocols. WoS research pursues a mechanism-driven, principle-oriented framework: advancing from reductionist factor analysis to multi-scale causal modelling of universal biogeochemical processes thereby providing theoretical foundations that validate, refine, and generalise applied solutions. The parallel development and increasing integration of these two bodies of literature constitute the current published knowledge base on Camellia oleifera soil science, as reflected in the CNKI and WoS databases.

Integrating bibliometric mapping with empirical synthesis, we identify four priority scientific questions:

Scientific Question 1: Multi-scale mechanisms of acidification-driven biogeochemical disruption. While research has progressed from pH mapping to microbial community shifts, the coupled regulatory dynamics of acidification–Al^{3+} toxicity–nutrient stoichiometry imbalance particularly their non-linear interactions across rhizosphere, root, and microbial scales remain unresolved.

Scientific Question 2: Dynamic, tiered soil health assessment. Although MDS frameworks have matured, two interdependent challenges persist: (i) reconciling universal core indicators with modular, region-activated tiers; and (ii) transitioning from static snapshot evaluation to spatiotemporally resolved diagnosis via long-term monitoring and predictive modelling.

Scientific Question 3: Rhizosphere-to-phenotype cascade effects. Empirical evidence confirms that soil pH, K availability, and camellia genotype shape microbial community composition but the mechanistic cascade linking soil physicochemistry, functional guild assembly, root exudate signalling and seed oil biosynthesis pathways remains uncharacterised at molecular and physiological levels.

Scientific Question 4: Long-term ecological resilience under intensified management. Cross-sectional studies cannot capture ecosystem hysteresis or threshold responses. Establishing coordinated, multi-decadal fixed-site observatories coupled with factorial field experiments isolating climate, management, and soil type effects is essential to move from correlative inference to causal, process-based understanding.

5.2. Forward-Looking Research Priorities

Building on the bibliometric knowledge mapping and synthesis of emergent scientific frontiers, this study proposes four priority research directions to advance Camellia oleifera soil science toward greater depth, scalability, and intelligence while directly strengthening translational impact for industrial sustainability.

5.2.1. Multi-Scale Integration and Long-Term Ecological Observatories

The timeline visualization of CNKI clusters (**Figure 8**) reveals that while clusters such as #5 (Soil Physicochemical Properties) have persisted for over a decade, newer clusters related to biological indicators (e.g., #1 Soil Bacterial) exhibit relatively short active durations and fragmented temporal continuity. This pattern suggests that research on dynamic biological processes is often project-based and short-term, limiting the ability to discern long-term ecological trajectories.

Addressing this temporal fragmentation requires infrastructure that transcends typical funding cycles. Therefore, we propose the establishment of a Long-Term Ecological Research (LTER) network for Camellia oleifera agroforestry systems. Additionally, the sparse cross-regional collaboration revealed by the institutional network analysis further underscores the need for a coordinated network to harmonize protocols and enable multi-site comparisons, which is a prerequisite for validating any long-term observatory data.

5.2.2. Rhizosphere-Mediated Soil–Plant–Microbe Co-Regulation

Co-occurrence of keywords: Terms such as root exudates, rhizosphere and mycorrhizal fungi frequently co-occur with soil nutrients and stress tolerance, suggesting that the research community has recognised the pivotal role of the rhizosphere microbiome in linking soil, plants and microorganisms. Research must shift from soil-centric analysis to mechanistic investigation of tripartite crosstalk: (i) genotype-specific root exudate profiles across Camellia oleifera cultivars; (ii) microbial sensing and transcriptional responses; and (iii) downstream effects on nutrient acquisition efficiency, abiotic stress resilience, and soil structural stability. This integrated framework will inform both the breeding of rhizosphere-competent cultivars and the design of microbiome-targeted soil amendments transforming passive soil management into active biological regulation.

5.2.3. Toward Predictive Soil Health Diagnostics: Integrating Mature Assessment Frameworks with Emerging Computational Methods

Keyword emergence analysis indicates that machine learning has become a cutting-edge hot topic in English-language literature (2021–2024), indicating that these computational tools are being specifically applied to soil assessment problems, not merely appearing as isolated methodological imports, whilst the emergence of clustering methods such as the composite index method (Cluster #7) also reflects a trend towards the systematisation of quantitative evaluation.

The structural juxtaposition of a mature, static assessment framework (MDS) with emerging, dynamic computational methods (ML) creates a technological readiness condition, the field possesses a validated set of what to measure (MDS indicators) and is now acquiring the tools to transition from how to measure retrospectively to how to forecast prospectively.

To transcend the static limitations of conventional MDS frameworks, we propose an intelligent soil health system integrating: (i) dynamic MDS diagnostics; (ii) high-resolution digital soil maps [41]; (iii) real-time IoT data; (iv) gridded meteorological forecasts; and (v) farm-level management histories. Leveraging machine learning algorithms [44,49], this system will deliver spatially explicit, time-resolved soil health scores; project 6–12-month quality trajectories.

5.2.4. Broadening the Analytical Lens: The Case for Socio-Ecological Integration

Analysis of the keyword co-occurrence networks (**Figures 6** and **7**) reveals a conspicuous structural absence: terms related to socio-economic dimensions, such as farmer adoption, cost-benefit analysis, policy incentives, or livelihood impacts, those are entirely missing from the high-centrality core and appear rarely, if at all, in cluster labels. The knowledge map is structurally dominated by biophysical and agronomic nodes. Drawing on established frameworks from social-ecological systems (SES) research and sustainability transitions literature rather than solely on CiteSpace outputs, we suggest that future research on Camellia oleifera soil quality would benefit from expanding its analytical scope to include:

Economic viability assessments of recommended soil management practices (e.g., cost-effectiveness of biochar application vs. conventional fertilization). Adoption barrier analysis among smallholder farmers, including labor constraints, access to organic amendments, and land tenure security. We advocate establishing a National Camellia Agroecosystem Observatory, modelled on China's Ecosystem Research Network (CERN) [50]. This observatory will conduct comparative life-cycle assessments of key strategies including organic fertiliser substitution, cover cropping, and biochar integration quantifying their net environmental benefits, and socio-economic outcomes. The goal is to co-develop context-sensitive, incentive-aligned soil stewardship models that deliver synergistic ecological, economic, and social returns.

Author Contributions

Conceptualization, Y.-k.C. and M.-C.C.; Investigation, Y.-k.C. and X.-f.M.; Data curation, Y.-k.C. and Y.Z.; Validation, X.-f.M.; Writing—original draft preparation, Y.-k.C.; Writing—review and editing, M.-C.C., X.-f.M. and Y.Z.; Visualization, Y.-k.C.; Supervision, M.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by Jiaying University's University-level Teaching Quality and Teaching Reform Project grant number [ZLGC2023102].

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The literature data used in this study were retrieved from relevant publications between 1995 and 2024 from China National Knowledge Infrastructure (CNKI) and the Web of Science Core Collection (WoSCC). These are publicly available literature datasets, and no new data were generated.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

[1] Zhao, Z., Fan, J., Fu, M., et al., 2020. A Review of the Main Constituents and Pharmacological Studies of Camellia Oil. Traditional Chinese Medicine. 9(4), 368–373. DOI: https://doi.org/10.12677/TCM.2020.94055 (in Chinese)

[2] National Forestry and Grassland Administration, National Park Administration, 2024. In the First Three Quarters, 3.71 Million Mu of New Oil-Tea Camellia Forests Were Newly Planted Nationwide. Available from: https://www.forestry.gov.cn/search/590504 (cited 22 October 2024). (in Chinese)

[3] Liu, L., Zhang, L., Pan, J., et al., 2020. Soil C-N-P pools and stoichiometry as affected by intensive management of camellia oleifera plantations. PLOS ONE. 15(9), e0238227. DOI: https://doi.org/10.1371/journal.pone.0238227

[4] Ye, J., Wang, Y., Kang, J., et al., 2022. Effects of Long-Term Use of Organic Fertilizer with Different Dosages on Soil Improvement, Nitrogen Transformation, Tea Yield and Quality in Acidified Tea Plantations. Plants. 12(1), 122. DOI: https://doi.org/10.3390/plants12010122

[5] Meng, C., Tian, D., Zeng, H., et al., 2019. Global soil acidification impacts on belowground processes. Environmental Research Letters. 14(7), 074003. DOI: https://doi.org/10.1088/1748-9326/ab239c

[6] Huang, Y., Lin, Y., Zhang, L., et al., 2023. Effects of Interaction between Claroideogolmus etuicatum and Bacillus aryabhattai on the Utilization of Organic Phosphorus in Camellia oleifera Abel. Journal of Fungi. 9(10), 977. DOI: https://doi.org/10.3390/jof9100977

[7] Zhang, M., Wang, W., Bai, S.H., et al., 2021. Intensive management of phosphorus fertilization in Camellia oleifera Abel. to minimize phosphorus losses to the environment. Industrial Crops and Products. 170, 113824. DOI: https://doi.org/10.1016/j.indcrop.2021.113824

[8] Yuan, J., Huang, L., Zhou, N., et al., 2017. Fractionation of Inorganic Phosphorus and Aluminum in Red Acidic Soil and the Growth of Camellia oleifera. HortScience. 52(9), 1293–1297. DOI: https://doi.org/10.21273/HORTSCI12189-17

[9] Pei, X., Han, X., Qian, J., et al., 2020. Soil fertility assessment indicators from the perspective of natural resources comprehensive observation. Resource Science. 42(10), 1953–1964. DOI: https://doi.org/10.18402/resci.2020.10.12

[10] Osman, K.T., 2013. Plant Nutrients and Soil Fertility Management. In Soils. Springer: Dordrecht, The Netherlands. pp. 129–159. DOI: https://doi.org/10.1007/978-94-007-5663-2_10

[11] Bünemann, E.K., Bongiorno, G., Bai, Z., et al., 2018. Soil quality – A critical review. Soil Biology and Biochemistry. 120, 105–125. DOI: https://doi.org/10.1016/j.soilbio.2018.01.030

[12] Lehmann, J., Bossio, D.A., Kögel-Knabner, I., et al., 2020. The concept and future prospects of soil health. Nature Reviews Earth & Environment. 1(10), 544–553. DOI: https://doi.org/10.1038/s43017-020-0080-8

[13] Doran, J.W., Coleman, D.C., Bezdicek, D.F., et al. (eds.), 1994. Defining Soil Quality for a Sustainable Environment, 1st ed. Wiley: Hoboken, NJ, USA. DOI: https://doi.org/10.2136/sssaspecpub35

[14] Zhao, B., Gao, R., Zhang, X., et al., 2024. Comparison of soil quality assessment methods for different vegetation eco-restoration techniques at engineering disturbed areas. PeerJ. 12, e18033. DOI: https://doi.org/10.7717/peerj.18033

[15] Luo, M., Pan, B., Wang, H.L., et al., 2024. Soil Quality Evaluation of Camellia oleifera Plantations Based on Minimum Data Set. Guangxi Forestry Science. 53(4), 438–446. DOI: https://doi.org/10.19692/j.issn.1006-1126.20240404 (in Chinese)

[16] Bajiu, A., Gao, K., Zeng, G., et al., 2024. Impact of Intercropping Five Medicinal Plants on Soil Nutrients, Enzyme Activity, and Microbial Community Structure in Camellia oleifera Plantations. Microorganisms. 12(8), 1616. DOI: https://doi.org/10.3390/

microorganisms12081616
[17] Chen, C., Hu, Z., Liu, S., et al., 2012. Emerging trends in regenerative medicine: a scientometric analysis in *CiteSpace*. Expert Opinion on Biological Therapy. 12(5), 593–608. DOI: https://doi.org/10.1517/14712598.2012.674507
[18] Zhang, J., Liu, J., Chen, Y., et al., 2021. Knowledge Mapping of Machine Learning Approaches Applied in Agricultural Management—A Scientometric Review with CiteSpace. Sustainability. 13(14), 7662. DOI: https://doi.org/10.3390/su13147662
[19] Wu, X., Cai, J., Lei, J., et al., 2023. Visualization Analysis of Research Progress on Soil Health Evaluation Based on Bibliometric Analysis. Soils. 55(3), 647–657. (in Chinese)
[20] Chen, C., 2017. Science Mapping: A Systematic Review of the Literature. Journal of Data and Information Science. 2(2), 1–40. DOI: https://doi.org/10.1515/jdis-2017-0006
[21] Tang, G., Tang, S., Li, Y., 2002. Study on Soil Characters of Camellia oleifera Forest Land. Jiangxi Forestry Science and Technology. 30(1), 20–24. (in Chinese)
[22] Hu, D., Sun, K., Fan, F., et al., 2005. Study on the Nutrient Limiting Factors in the Young Oil-Tea Soil. Soil and Fertilizer Sciences in China. (6), 20–22. (in Chinese)
[23] Song, Q., He, Y., Wu, Y., et al., 2020. Biochar Impacts on Acidic Soil from Camellia Oleifera Plantation: A Short-Term Soil Incubation Study. Agronomy. 10(9), 1446. DOI: https://doi.org/10.3390/agronomy10091446
[24] Qiao, H., Chen, L., Hu, Y., et al., 2021. Soil Microbial Resource Limitations and Community Assembly Along a Camellia oleifera Plantation Chronosequence. Frontiers in Microbiology. 12, 736165. DOI: https://doi.org/10.3389/fmicb.2021.736165
[25] Lv, J., Huo, C., Zhang, J., et al., 2024. Host genotype and age shape the microbial community in the rhizosphere soils of Camellia forests. Frontiers in Microbiology. 15, 1440255. DOI: https://doi.org/10.3389/fmicb.2024.1440255
[26] Zhang, P., Cui, Z., Guo, M., et al., 2020. Characteristics of the soil microbial community in the forestland of *Camellia oleifera*. PeerJ. 8, e9117. DOI: https://doi.org/10.7717/peerj.9117
[27] He, B., Li, Q., Zou, S., et al., 2024. Dynamic Changes of Soil Microbial Communities During the Afforestation of Pinus Armandii in a Karst Region of Southwest China. Microbial Ecology. 87(1), 36. DOI: https://doi.org/10.1007/s00248-024-02345-8
[28] Liu, J., Wu, L., Chen, D., et al., 2017. Soil quality assessment of different Camellia oleifera stands in mid-subtropical China. Applied Soil Ecology. 113, 29–35. DOI: https://doi.org/10.1016/j.apsoil.2017.01.010
[29] Chen, J., Han, X., Ye, S., et al., 2022. Integration of small RNA, degradome, and transcriptome sequencing data illustrates the mechanism of low phosphorus adaptation in Camellia oleifera. Frontiers in Plant Science. 13, 932926. DOI: https://doi.org/10.3389/fpls.2022.932926
[30] He, H., Shi, L., Yang, G., et al., 2020. Ecological Risk Assessment of Soil Heavy Metals and Pesticide Residues in Tea Plantations. Agriculture. 10(2), 47. DOI: https://doi.org/10.3390/agriculture10020047
[31] Huang, A., Wang, Z., Yang, D., et al., 2023. Effects of tea oil camellia (Camellia oleifera Abel.) shell-based organic fertilizers on the physicochemical property and microbial community structure of the rhizosphere soil. Frontiers in Microbiology. 14, 1231978. DOI: https://doi.org/10.3389/fmicb.2023.1231978
[32] Zha, Q., Yang, Y., Huang, Y., 2015. The Effects of Two Kinds of Inter-Planting on Soil Nutrition of Camellia oleifera Forest. Subtropical Plant Science. 44(2), 130–134. (in Chinese)
[33] Ding, Y., Lou, X., Fu, S., et al., 2018. Effects of Intercropping with Vulpia myuros on Soil Nutrients and Enzyme Activity of Camellia oleifera Forest. Forest Research. 31(2), 170–175. (in Chinese)
[34] Song, Y., Sun, L., Wang, H., et al., 2023. Enzymatic fermentation of rapeseed cake significantly improved the soil environment of tea rhizosphere. BMC Microbiology. 23(1), 250. DOI: https://doi.org/10.1186/s12866-023-02995-7
[35] Xiang, W., Zhou, W., Yu, L., et al., 2013. Establishment of Minimum Data Set for Evaluating Soil Fertility Quality of Oil Camellia Forest in Low Mountain Areas. Hunan Agricultural Sciences. (3), 44–46. DOI: https://doi.org/10.16498/j.cnki.hnnykx.2013.03.017 (in Chinese)
[36] Chen, Y., Wu, L., Tao, Y., et al., 2024. Soil Quality Evaluation of Camellia oleifera Abel. Forests in Guizhong Region Based on Minimum Data Set. Journal of Green Science and Technology. 26(8), 33–40. (in Chinese)
[37] Maćkiewicz, A., Ratajczak, W., 1993. Principal components analysis (PCA). Computers & Geosciences. 19(3), 303–342. DOI: https://doi.org/10.1016/0098-3004(93)90090-R
[38] Zhang, S., Kong, J., Chen, L., et al., 2022. Increased Tea Saponin Content Influences the Diversity and Function of Plantation Soil Microbiomes. Microbi-

ology Spectrum. 10(1), e02324-21. DOI: https://doi.org/10.1128/spectrum.02324-21

[39] Wang, Y., Xie, P., She, J., et al., 2023. Forest type Dominant the soil Bacterial Community Composition Rather than Soil Depth or Slope Position. Journal of Soil Science and Plant Nutrition. 23(3), 4558–4571. DOI: https://doi.org/10.1007/s42729-023-01372-7

[40] Tu, J., Chen, J., Zhou, J., et al., 2019. Plantation quality assessment of Camellia oleifera in mid-subtropical China. Soil and Tillage Research. 186, 249–258. DOI: https://doi.org/10.1016/j.still.2018.10.023

[41] Yang, G., Li, Z., Xiao, X., 2023. Determination of Soil Fertility Characteristics and Heavy Metal Health Risks Using the Camellia oleifera Planting Base in Guizhou Province, China. Water. 15(21), 3856. DOI: https://doi.org/10.3390/w15213856

[42] Lu, J., Li, J., Peng, P., et al., 2025. Effects of Different Varieties of Camellia oleifera on Soil Microbial Community Structure and Stability. Journal of Soil Science and Plant Nutrition. 25(2), 4183–4196. DOI: https://doi.org/10.1007/s42729-025-02391-2

[43] Xiang, T., Qiang, F., Liu, G., et al., 2023. Soil Quality Evaluation and Dominant Factor Analysis of Economic Forest in Loess Area of Northern Shaanxi. Forests. 14(6), 1179. DOI: https://doi.org/10.3390/f14061179

[44] Liu, C., Chen, L., Tang, W., et al., 2018. Predicting Potential Distribution and Evaluating Suitable Soil Condition of Oil Tea Camellia in China. Forests. 9(8), 487. DOI: https://doi.org/10.3390/f9080487

[45] Zhou, J., Ai, Z., Wang, H., et al., 2019. Phosphorus Alleviates Aluminum Toxicity in Camellia oleifera Seedlings. International Journal of Agriculture and Biology. 21(1), 237–243.

[46] Zheng, H., Wang, X., Wu, J., et al., 2023. Long-term impacts of extensive terracing on soil aggregates and associated C–N–P in the Camellia oleifera orchard of southern China. CATENA. 233, 107512. DOI: https://doi.org/10.1016/j.catena.2023.107512

[47] Chen, Y., Zheng, J., Yang, Z., et al., 2023. Role of soil nutrient elements transport on Camellia oleifera yield under different soil types. BMC Plant Biology. 23(1), 378. DOI: https://doi.org/10.1186/s12870-023-04352-2

[48] Yang, J., Chen, B., Manan, S., et al., 2022. Critical metabolic pathways and SAD/FADs, WRI1s, and DGATs cooperate for high-oleic acid oil production in developing oil tea (*Camellia oleifera*) seeds. Horticulture Research. 9, uhac087. DOI: https://doi.org/10.1093/hr/uhac087

[49] Tang, J., Jiang, F., Long, Y., et al., 2022. Identification of the Yield of Camellia oleifera Based on Color Space by the Optimized Mean Shift Clustering Algorithm Using Terrestrial Laser Scanning. Remote Sensing. 14(3), 642. DOI: https://doi.org/10.3390/rs14030642

[50] Dong, R., Ren, X.L., Gai, A.H., et al., 2020. Use of CERN for estimating the soil conservation capability of typical forest ecosystems in China. Acta Ecologica Sinica. 40(7), 2310–2320. (in Chinese)