SCIEPublish

Review

Bibliometric Review of Bio-Based Building Materials and a Comparison with Prior Research: Toward Probabilistic Approaches for Material Variability

Sara Darwish [1], Ariane Abou Chakra [1,*], Tamara Al Bittar [2] and Joseph Absi [3]

[1] LMDC, INSA Toulouse, Université de Toulouse, 31077 Toulouse, France; darwish@insa-toulouse.fr (S.D.)
[2] Faculty of Engineering, Lebanese University, President Michel Sleiman Campus, Tripoli 1300, Lebanon; tamara.albittar@gmail.com (T.A.B.)
[3] IRCER, UMR 7315, CNRS, University of Limoges, European Ceramics Center, 12 Rue Atlantis, 87068 Limoges, France; joseph.absi@unilim.fr (J.A.)

* Corresponding author. E-mail: abouchak@insa-toulouse.fr (A.A.C.)

Received: 18 September 2025; Revised: 12 November 2025; Accepted: 1 December 2025; Available online: 12 December 2025

ABSTRACT: The purpose of this article is to extend previous bibliometric research on bio-based building materials by conducting a comparative analysis. The objective is to expand the initial study by applying a broader and more inclusive set of search terms to evaluate the sensitivity of Web of Science to keyword variations. In parallel, a separate bibliometric analysis is performed on research related to probabilistic approaches, which are essential for managing the variability and uncertainty in building material properties. Finally, a third bibliometric analysis is carried out at the intersection of these two fields: bio-based building materials and probabilistic methods. This integrated analysis aims to highlight the existing gap in the literature. The findings reveal the limited application of probabilistic approaches in the study of bio-based building materials and underscore the need to incorporate uncertainty quantification and stochastic modeling to understand better and optimize these sustainable construction resources. Overall, the results highlight two main outcomes. First, they demonstrate the strong sensitivity of bibliometric outcomes to the choice of search terms and databases, emphasizing the need for transparent and consistent keyword strategies; second, they show that the overlap between probabilistic approaches and bio-based materials research remains extremely limited, underscoring the importance of fostering stronger integration between these areas.

Keywords: Bio-based building material; Probabilistic analysis; Comparative study; Web of Science; VOSviewer

1. Introduction

The construction industry remains a major contributor to global carbon emissions and resource depletion [1–5]. The extraction and processing of conventional construction materials, particularly cement and steel, account for significant greenhouse gas emissions and energy consumption [6]. In light of escalating environmental concerns and the urgent need for sustainable development, the search for environmentally friendly alternatives to traditional materials has become increasingly critical.

One promising solution is the integration of bio-based building materials into construction [7]. Ref. [8], particularly in concrete. Derived largely from agricultural waste and renewable resources, these materials offer multiple environmental advantages, including reductions in carbon emissions and waste generation, while promoting circular economy principles. Furthermore, bio-based building materials exhibit favorable insulation properties [8], enhancing building energy efficiency and reducing heating and cooling demands.

Despite these advantages, bio-based building materials are characterized by substantial variability [9]. Numerous factors, such as raw material type, origin, growth conditions, and processing methods, influence their properties. This variability affects key characteristics such as size, dimensions, orientation, as well as thermal and mechanical performance. Such variability poses significant challenges for standardization, structural reliability, and large-scale application.

https://doi.org/10.70322/ces.2025.10019

Addressing these challenges requires probabilistic approaches that systematically account for and quantify the inherent uncertainties in bio-based building materials. Unlike deterministic methods, which assume constant values for material properties, probabilistic frameworks capture uncertainties, identify influential parameters, and assess the propagation of uncertainties from inputs to structural performance [10,11]. It is important to note that probabilistic approaches have been widely and successfully applied across a broad range of construction materials such as concrete and steel, as well as within life cycle assessment, seismic analysis, and other areas of geotechnical and structural engineering [12–14]. However, most existing studies on bio-based building material performance rely on deterministic approaches, thereby leaving a critical gap in understanding their behavior under uncertainty. This study addresses two research questions: (1) How has research evolved in the domains of bio-based materials and probabilistic analysis? (2) To what extent do these fields intersect? Through comparative bibliometric analysis, the study highlights database sensitivity to keyword selection and reveals the underexplored integration of uncertainty modeling in sustainable-material research.

The present study aims to conduct a comparative bibliometric analysis, building upon a previous work [1], using data extracted from the Web of Science (WoS) database. This approach aims to assess the sensitivity of bibliometric outcomes to variations in keyword selections. Despite its wide acceptance, WoS exhibits certain limitations, including disciplinary biases favoring the Natural Sciences, Engineering, and Biomedical fields, while underrepresenting the Social Sciences, Arts, and Humanities. Additionally, WoS's emphasis on English-language publications may restrict the visibility of research published in other languages, thereby introducing further biases. These limitations underscore the need for continued refinement of bibliometric methodologies and the development of more inclusive and representative research metrics.

Before advancing methodological frameworks, however, it is essential to understand how research on bio-based building materials and probabilistic methods has evolved and where the two fields intersect. Bibliometric analysis provides a valuable quantitative tool. By applying statistical methods to map research trends, co-authorship networks, and the scientific impact of publications within a particular field, bibliometrics provides a comprehensive overview of the structure and evolution of the research domain. This approach helps to identify influential publications, leading authors, key institutions, emerging themes, and active countries over time, while also uncovering gaps in knowledge.

Prior to 2004, most bibliometric studies relied on data retrieved from the Institute of Scientific Information (ISI), a pioneering organization based in the United States. ISI played a central role in indexing scholarly publications and tracking citations. Today, these databases are integrated into the Web of Science (WoS) Core Collection [15], one of the most widely used and trusted sources for bibliometric data across various scientific disciplines.

To conduct the bibliometric mapping and visualization, this study utilizes VOSviewer version 1.6.20, an open-access software widely adopted for constructing and analyzing bibliometric networks. VOSviewer version 1.6.20 enables the identification of key research clusters, co-authorship patterns, co-citation, and thematic trends. As an open-access tool, the software is freely accessible to researchers with limited funding. This accessibility does not imply any technical limitations; on the contrary, comparative studies using multidimensional scaling methods indicate that VOSviewer version 1.6.20 may offer certain advantages [16]. This study contributes to the literature by (1) examining the sensitivity of bibliometric results to keyword variation, (2) identifying publication trends and key contributions in both bio-based building material and probabilistic research, and (3) revealing the significant gap that persists in their integration. While both domains have independently witnessed increasing attention in recent years, bio-based building materials as a response to the demand for sustainable alternatives in construction and manufacturing, and probabilistic methods as powerful tools for managing uncertainty in complex systems, their integration remains significantly underexplored in the scientific literature. By jointly analyzing these two fields, this study provides a novel interdisciplinary perspective that not only addresses methodological sensitivity but also promotes the integration of probabilistic reasoning into bio-based materials research. Bridging these aspects not only offers a methodological innovation but also promises to enhance the understanding of the variability in bio-based building materials, leading to more robust performance prediction models and optimization strategies for bio-based materials, supporting a more sustainable construction industry.

2. Methodological Tools

This study combines data retrieval, processing, and visualization techniques to conduct a comprehensive bibliometric analysis. The following subsections outline the key platforms and software used to gather, process, and analyze the relevant bibliographic data.

2.1. Web of Science

In this study, the Web of Science Core Collection was selected as the primary data source because of its rigorous indexing standards and multidisciplinary coverage. The following subsections describe the scope and structure of the database, followed by the specific search strategies applied to ensure accurate and inclusive retrieval of relevant publications.

2.1.1. Database Scope and Structure

The Web of Science (WoS) database is among the most widely used sources for bibliometric studies [17]. However, the choice of any bibliographic database inevitably introduces inherent biases, as no single platform covers all scientific publications pertaining to a specific research area. Each database applies its own scope, indexing policies, data structure, and coverage criteria, which leads to notable discrepancies in literature representation. A transparent search process and careful interpretation of results are therefore essential to minimize these biases.

To ensure comprehensive and reliable coverage of the literature, several major bibliometric databases were considered, including Web of Science (WoS), Scopus, Dimensions, and Google Scholar. Each database offers unique strengths and limitations. The comparison Table 1 summarizes the main features of these databases and provides a justification for the choice of WoS as the primary data source.

Table 1. Comparison of major bibliometric databases.

Database	Coverage Strengths	Limitations
Web of Science	Among the main sources for citation data [18]. Accurate assessment of authors' impact [19]. Suitable multidisciplinary and international bibliometric analyses [18]. Annual growth [20]. Covers publications dated back to 1900 [20]. It has the most selective journal coverage [21].	Countries and languages are underrepresented [18]. English-language journals are overrepresented [21].
Scopus	Among the main sources for citation data [18]. Suitable for multidisciplinary and international bibliometric analyses [18]. Annual Growth [20]. Covers publications dated back to 1970 [20].	Countries and languages are underrepresented [18]. Coverage mainly focuses on journals less on other forms of scientific knowledge dissemination (e.g., books, proceedings, and reports) [18]. English-language journals are overrepresented [21].
Dimensions	Broad coverage for publications and citation [21]. It has an article-level subject classification system [21].	Limited historical depth [21]. Have better coverage in Social Sciences and Arts and Humanities [21].
Google Scholar	Very broad coverage [22]. Accurate assessment of authors' impact [19]. Useful for self-assessment [23].	Low data quality [18]. Lack of transparency of the coverage [23]. Indexation of non-existing journals [24]. Sporadic coverage of non-English literature [24]. Less commonly used for research assessment [21]. Sometimes, one paper counts multiple times [22].

WoS was ultimately selected for this study due to its broad coverage and widespread use in bibliometric analyses, despite some limitations in representing certain disciplines and languages. Moreover, maintaining consistency with the previous study [1] ensures a valid comparative analysis and enables assessment of the sensitivity of WoS to keyword variations.

2.1.2. Search Strategies Used in This Study

Because WoS is highly sensitive to search terms, careful formulation of queries was critical to ensure comprehensive and representative data retrieval. To ensure a fair representation of global research output, it is essential to consider the use of synonyms, alternative spellings, and variations in terminology that may exist across different regions and countries. Without this consideration, valuable contributions from certain authors, institutions, or countries could be underrepresented or completely missed, leading to biased or incomplete results. To capture a broad and representative spectrum of the relevant literature on bio-based building materials and probabilistic methods, this study employed a comprehensive set of search terms designed to reflect the diverse vocabulary commonly used in these research fields.

Table 2 presents the number of publications retrieved from the Web of Science database since 1900 using different terms and synonyms related to the study. The data clearly show that relying on a single search term is not sufficient to capture the full scope of relevant literature. When only one term was used to search for publications related to bio-based building materials, 2576 results were found. However, by including multiple synonyms and related expressions, the number of publications increased significantly to 5631. This demonstrates the importance of developing a comprehensive search strategy that accounts for variations in terminology to ensure a more accurate and complete bibliometric analysis. The keyword sets presented in Table 2 were selected through a structured process that combined both deductive and inductive strategies. Initially, the baseline term "bio-based building material" was chosen due to its frequent use in the previous bibliometric study [1]. To expand the scope, a synonym-expansion approach was applied, which involved reviewing terminology used in highly cited publications, consulting controlled vocabulary headings in the Web of Science database, and examining keyword lists from related reviews. From this process, alternative phrases such as "Bio Sourced Building material, Bio Renewable Building Material, and Plant Based Building Material" were identified and incorporated into the final keyword sets.

A similar iterative process was adopted for identifying publications related to probabilistic analysis. An initial search using the term "probabilistic analysis in buildings" yielded 3574 publications. However, by expanding the search with additional related terms, the number of relevant publications increased significantly to 50,620. Given the vast number of publications in this broader field, and to better focus the analysis on content relevant to this study, the term *material* was added to the search query. This refinement effectively narrowed the scope of results, yielding a more targeted subset of 5485 publications directly relevant to the probabilistic approaches applied to building materials.

Table 2. Number of publications based on various terminologies.

Topic	Year Published	Number of Publications
Bio-based building material	1900–2025	2576
Bio-based building material Bio Sourced Building material Bio Renewable Building Material Plant Based Building Material	1900–2025	5631
Probabilistic Analysis	1900–2025	75,950
Probabilistic analysis in buildings	1900–2025	3574
Probabilistic Analysis in Buildings Stochastic Analysis in Buildings Statistical Analysis in Buildings Probabilistic Analysis in Construction Stochastic Analysis in Construction Statistical Analysis in Construction	1900–2025	50,620
Probabilistic Analysis in Building material Stochastic Analysis in Building material Statistical Analysis in Building material Probabilistic Analysis in Construction material Stochastic Analysis in Construction material Statistical Analysis in Construction material	1900–2025	5485

In order to identify how probabilistic analysis intersects with research on bio-based construction materials, pairs of the keywords mentioned in Table 2 from the two fields were combined two at a time, and the relevant records were retrieved. These publications were subsequently screened and filtered through a structured selection workflow that included manual removal of duplicates and relevance evaluation. The overall identification and selection steps were conducted in accordance with the PRISMA flow diagram presented in Figure 1.

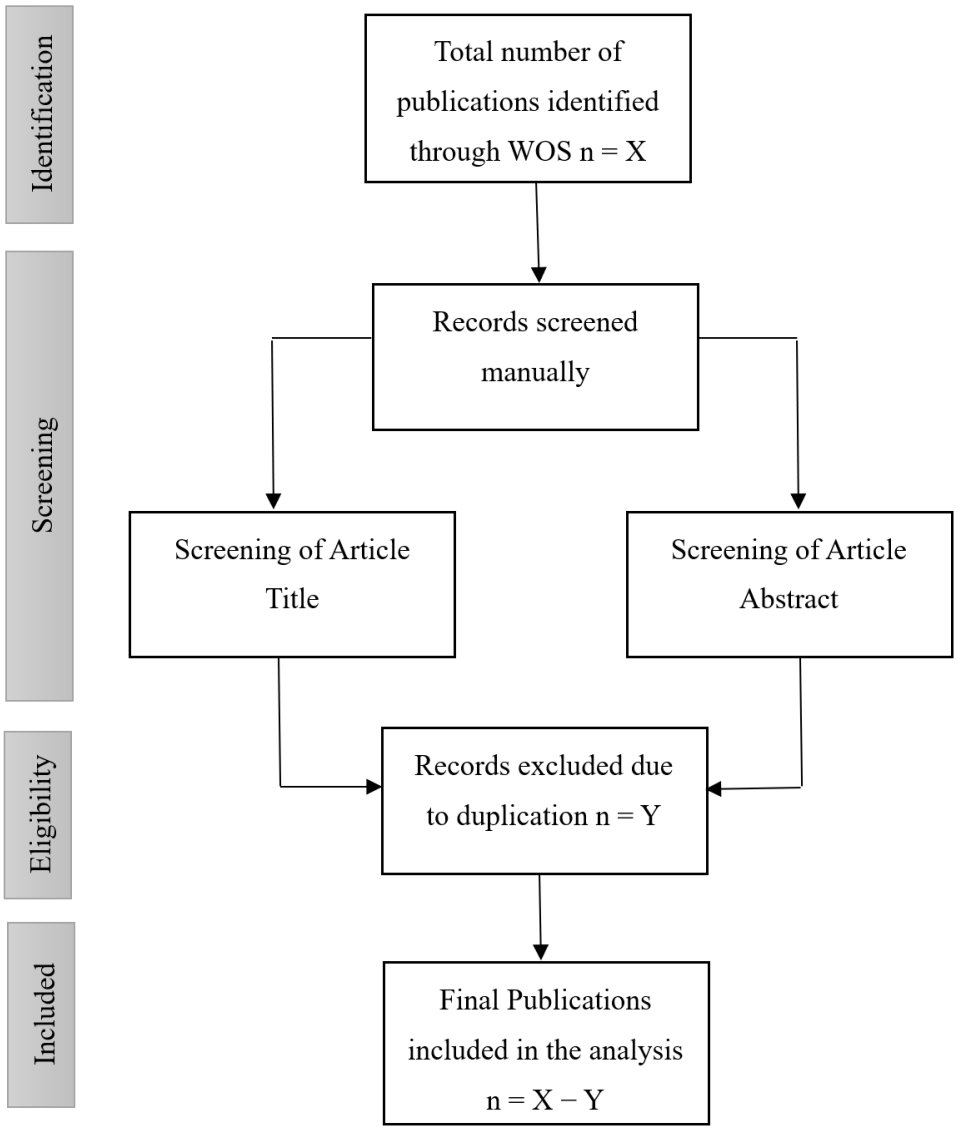

Figure 1. PRISMA-style flowchart illustrating the selection and screening process for bibliometric data.

2.2. VOSviewer Version 1.6.20

VOSviewer version 1.6.20 was selected for this study. The following subsections highlight the importance of using this software and outline the types of bibliometric networks it generates.

2.2.1. Visualization Approach

VOSviewer version 1.6.20 is an open access software tool widely used in bibliometric studies due to its simplicity, flexibility, and accessibility for new users [25]. In the present analysis, VOSviewer version 1.6.20 was selected for its seamless compatibility with the Web of Science database, its advanced graphical visualization capabilities, and its specialization in mapping co-authorship, co-citation, and keyword co-occurrence networks. The software provides a user-friendly interface and employs an optimized layout algorithm capable of handling large bibliometric networks with minimal parameter adjustment. It utilizes a distance-based mapping technique and a clustering method that produce clear and interpretable visualizations, even when analyzing dense datasets. Although various map configurations can be derived from the same database, in this study only the most relevant results are presented. Deep analysis and interpretation are performed on these results in order to highlight significant directions for future research.

2.2.2. Network Types

The algorithm used in VOSviewer version 1.6.20, developed in JavaScript, is specifically designed to extract, process, and visualize bibliometric data from structured scientific databases such as Web of Science [1]. Its core lies in generating network maps that illustrate relationships between various elements within the research, such as publications,

authors, keywords, or countries, based on different types of bibliometric networks. These elements are graphically represented as nodes on the map, with the size of each node proportional to its weight in the dataset, often measured by the number of citations or occurrences. Larger nodes indicate a higher frequency of citations or mentions, suggesting greater influence or relevance within the field.

The nodes are connected with lines that represent the strength of the relationship between them. The thickness of each line denotes the intensity or frequency of interaction, such as co-authorship frequency, keyword co-occurrence, or citation patterns. VOSviewer version 1.6.20 supports several types of bibliometric links, each offering unique insights into the structure and dynamics of scientific research. Some of these links are presented below:

- Co-authorship links reveal collaborative relationships among authors, organizations, or countries, highlighting key contributors and research networks.
- Co-occurrence links are based on the simultaneous appearance of keywords within publications, allowing identification of prevailing themes.
- Co-citation links track how frequently two items are cited together in later works, which helps identify conceptual linkages between the highly cited publications.

It is important to note that the size and complexity of the database can influence the clarity and usability of the visualized network. Large datasets may produce dense and complex visualizations, necessitating the application of filters or threshold adjustments to ensure that the resulting maps remain coherent and analytically meaningful.

The bibliometric analysis was conducted in three stages:

- Co-authorship analysis was performed to identify leading contributors and examine the collaborative dynamics within the intersection of bio-based building materials and probabilistic approaches.
- Co-citation analysis was conducted to uncover the seminal works and most influential publications that have shaped both fields.
- Co-occurrence analysis of keywords was carried out to map prevailing research themes and trends, as well as to identify underexplored areas warranting further investigation.

Together, these analyses provide a comprehensive overview of research development at the intersection of bio-based building materials and probabilistic methods, guiding future work toward underexplored areas.

3. Comparative Bibliometric Analysis—Bio-Based Building Materials

To strengthen the validity of this research, a comparative bibliometric analysis was conducted. The following subsections present a comparison with a previous study [1] and highlight the key findings.

3.1. Comparative Bibliometric Analysis with a Previous Study

This study builds upon previous work [1] that conducted a bibliometric analysis on bio-based building materials. While this work provides valuable foundational insights into the development of this research area, the present study extends the analysis by incorporating a comparative bibliometric approach with a particular emphasis on search strategy sensitivity within the Web of Science (WoS) database. The aim is to evaluate how variations in keyword selection influence bibliometric outcomes, including publication counts, network structure, and thematic clustering. By analyzing and comparing different keyword sets and their impact on the structure of the resulting bibliometric networks, this study aims to highlight how terminology choices can introduce biases, affecting the representativeness and completeness of bibliometric findings.

3.2. Results from Web of Science

The analysis of the Web of Science database provides insights into different dimensions of research activity. The following subsections present the publication trends, geographic distribution, and leading authors contributing to the field.

3.2.1. Publication Trends

According to the previous study [1], a total of 1778 publications related to bio-based building materials were identified in the Web of Science database between 1900 and 2022. In contrast, the current study identified 5631 publications over the same period. This increase is directly attributed to the use of a broader and more inclusive set of search terms, incorporating multiple synonyms and alternative expressions such as bio-based building material, bio-

sourced building material, bio-renewable building material, and plant-based building material. The search terms used in this study are detailed in Table 3.

Table 3. Number of publications based on various terminology related to bio-based building materials.

Topic	Year Published	Number of Publications
"Bio-based building material" or "Bio Sourced Building material" or "Bio Renewable Building Material" or "Plant Based Building Material" or	1900–2025	5631

In order to perform a year-by-year comparison of the data retrieved from Web of Science, Figure 2 was generated in this study based on the data collected using the expanded set of four terminologies outlined in Table 3.

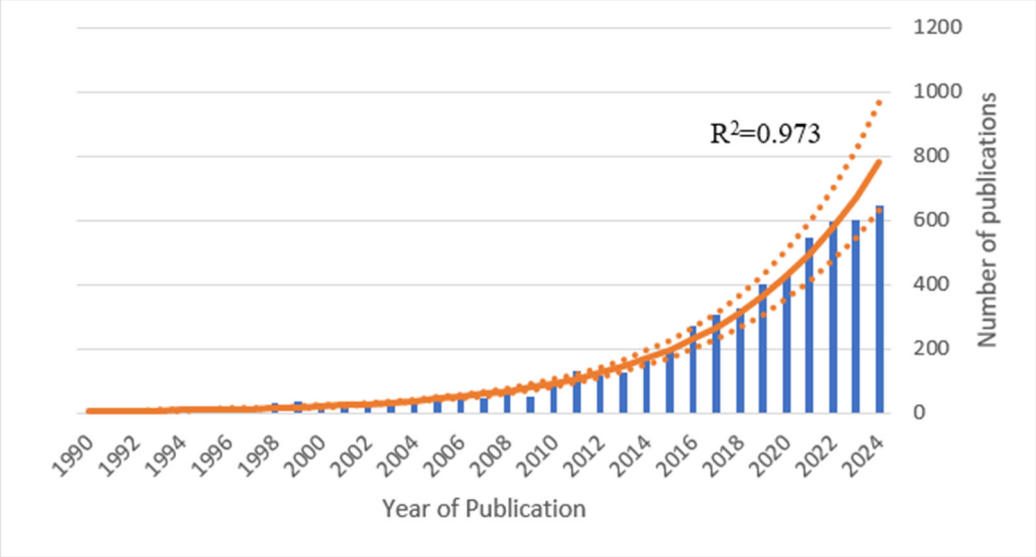

Figure 2. Number of publications over time using WOS based on Table 3 terminologies, with fitted exponential trend and 95% confidence interval.

The analysis conducted in this study and the previous study [1] highlight the significant increase in the number of publications over the years. This growth indicates a rising research interest in the field of bio-based building materials in the late 2000s. This surge in academic output corresponds with heightened global awareness of environmental challenges and the urgent need to identify sustainable alternatives to conventional building materials. Consequently, researchers and institutions have increasingly focused on the development and promotion of low-impact, bio-based solutions, which are now recognized as critical components of sustainable construction practices.

To further support this observation, an exponential trend line was fitted to the annual publication data, and a 95% confidence interval was calculated around the fitted curve. The trend yields an R^2 value of 0.975, which confirms a strong positive correlation between publication year and research output. This result demonstrates that the observed growth is not random but instead follows a statistically significant upward trend, indicating a sustained increase in scientific attention to this topic over time.

While the plots presented by the previous study [1] and Figure 2 demonstrate a similar upward trajectory, it is important to note that the number of publications in Figure 2 is approximately three times higher than that in the previous article [1]. This substantial discrepancy highlights the sensitivity of bibliometric search results within the Web of Science database to the specific terminology and keywords used during the data retrieval process.

To address this issue and ensure the validity and comprehensiveness of bibliometric findings, it is essential to adopt a carefully constructed, inclusive set of search terms that captures regional, linguistic, and disciplinary variations in terminology. Such a comprehensive approach minimizes potential biases, improves the representativeness of the dataset, and provides a more accurate reflection of global research efforts across diverse countries and scientific communities.

3.2.2. Geographic Distribution

A key difference between the present study (Table 4) and the earlier work [1] lies in the geographic distribution of publications by country. While both analyses rely on the Web of Science database, the current study employed a broader set of keywords for the term bio-based building material. As a result, the number of identified publications in each country increased significantly. Although China remains the leading country in terms of publication output, the positions of France and the United States have been switched, with the USA now ranking second, ahead of France. This change may be attributed to the fact that certain keywords are more commonly used in some countries than in others. More critically, China's dominance could also be associated with national research incentives and funding priorities that strongly promote sustainability-related studies, while France's lower ranking may reflect a terminological bias in English-language databases that underrepresents research published in French or with different keyword conventions. This implies the sensitivity of bibliometric results to keyword selection and highlights the importance of a comprehensive search strategy in capturing the full scope of research activity.

Table 4. Top 10 countries contributing to research using WoS (based on Table 3 terminologies).

Countries	Publications	Percentage of Publications %
China	991	23
USA	728	17
France	502	11
Italy	430	10
Germany	407	9
India	348	8
England	311	7
Spain	282	6
Canada	201	5
The Netherlands	156	4

Figure 3 presents the percentage distribution of publications among the top 10 countries contributing to the field in this study. Although European countries are not among the top two countries, they are well recognized with a strong presence, especially with France occupying the third position. As illustrated in the figure, despite China's leading position, the field remains notably European-centric in diversity, suggesting that EU initiatives may be driving broader participation and collaboration in bio-based building materials research.

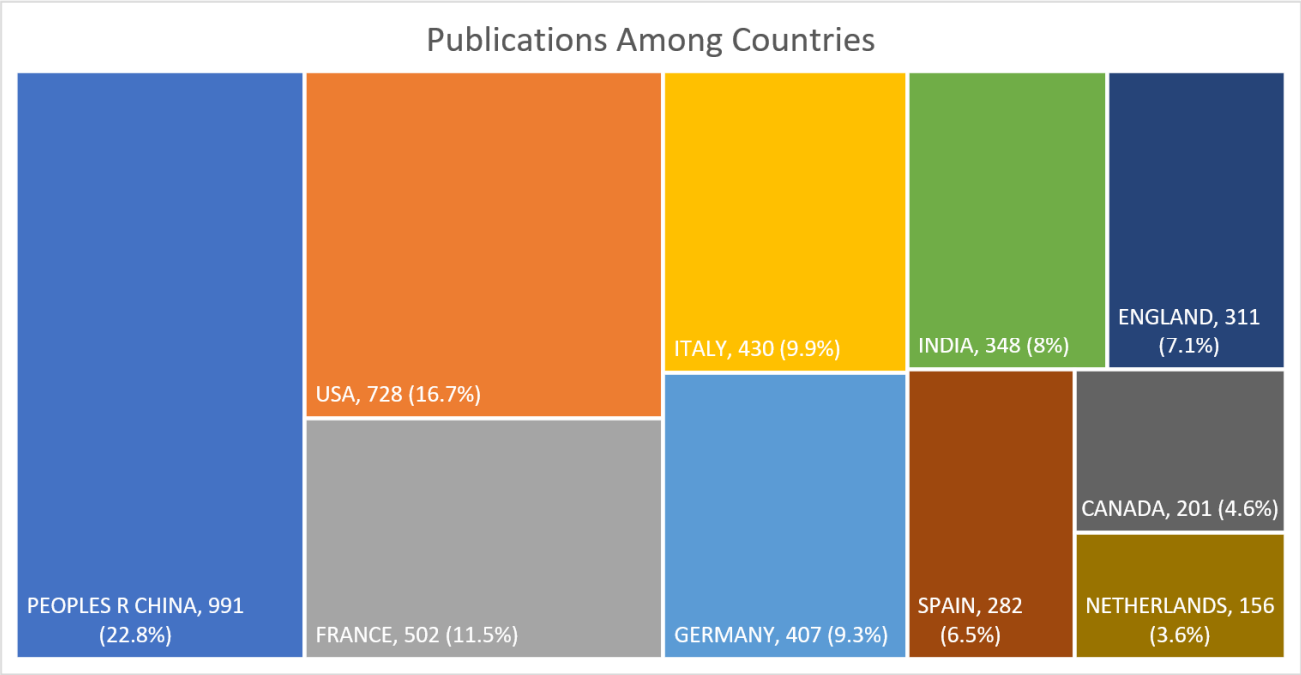

Figure 3. Number and percentage distribution of publications among the top 10 countries using WoS, based on the keywords set shown in Table 3.

3.2.3. Top Authors

A comparison between the results from the previous study [1] and Table 5 reveals notable differences in the ranking of the top contributing authors in the field of bio-based building materials. Although six of the nine authors identified in the previous study [1] remain within the top 10 authors in this study, most of them did not retain their original ranking positions. This shift highlights how differences in search strategies or keyword selection, even within the same database, can lead to varying bibliometric outcomes. It also suggests that databases may exhibit a certain degree of bias, potentially favoring well-established authors and influencing their visibility in bibliometric analyses.

It is important to note that attributing publications to individual authors in large bibliographic databases such as Web of Science can involve several identification challenges. Variations in how authors report their names across publications, as well as similarities between different authors' names, may lead to inconsistencies in attribution. In some cases, the database may merge records belonging to different authors with similar names, resulting in over-representation. In other cases, it may separate the publications of a single author into multiple name variants, causing under-representation. These well-known limitations indicate that top-author rankings should be interpreted cautiously, as they may not perfectly reflect individual research output.

Table 5. Number of publications for the top 10 contributing authors using WoS (based on Table 3 terminologies).

Author	Number of Publications	Institution	Country
Collet Florence	21	Rennes University	France
Blanchet Pierre	20	Laval University, Québec	Canada
Lanos Christophe	18	Rennes University	France
Bajare Diana	17	Riga Technical University	Latvia
Habert Guillaume	17	Swiss Fed Inst Technol, Zurich	Switzerland
Magniont Camille	17	LMDC Laboratory of Materials and Durability of Constructions, TOULOUSE	France
Walker Paul	16	University of Technology Sydney	Australia
Stevulova Nadezda	16	Tech Univ Kosice	Slovakia
Maalouf Chadi	15	University of Reims Champagne-Ardennes, GRESPI Laboratory	France
Faria Paulina	14	NOVA University Lisbon	Portugal
Sinka Maris	14	Riga Tech University	Latvia

3.3. VOSviewer Version 1.6.20 Analysis

The VOSviewer version 1.6.20 analysis offers a deeper understanding of research dynamics through keyword relationships and collaboration patterns. The following subsections present the co-occurrence keyword map and highlight the main co-authorship and citation networks.

3.3.1. Keyword Co-Occurrence Map

A comparison between the keyword co-occurrence network generated in this study (Figure 4) and the one presented in the previous study [1] reveals notable differences in keyword distribution and thematic clustering. In the current analysis, keywords are grouped into three main thematic clusters: (1) biological and chemical composition (blue), (2) material properties and behavior (green), and (3) modeling and simulation (red). These clusters indicate the major research directions in the field: the first focuses on understanding the intrinsic characteristics of bio-based materials, the second addresses their practical performance and behavior, and the third emphasizes computational approaches for predicting material performance. The emergence of these clusters can be attributed to the broader and more inclusive search strategy employed in this study, which allowed a wider range of publications to be collected and, consequently, revealed modeling and simulation as a central theme. Contrary to the previous findings [1], where the concept of "model" appeared only marginally in the keyword network. This comparison highlights how the choice of search terms can shape bibliometric outcomes and demonstrates that a more comprehensive keyword strategy can uncover emerging research areas that may otherwise remain underrepresented.

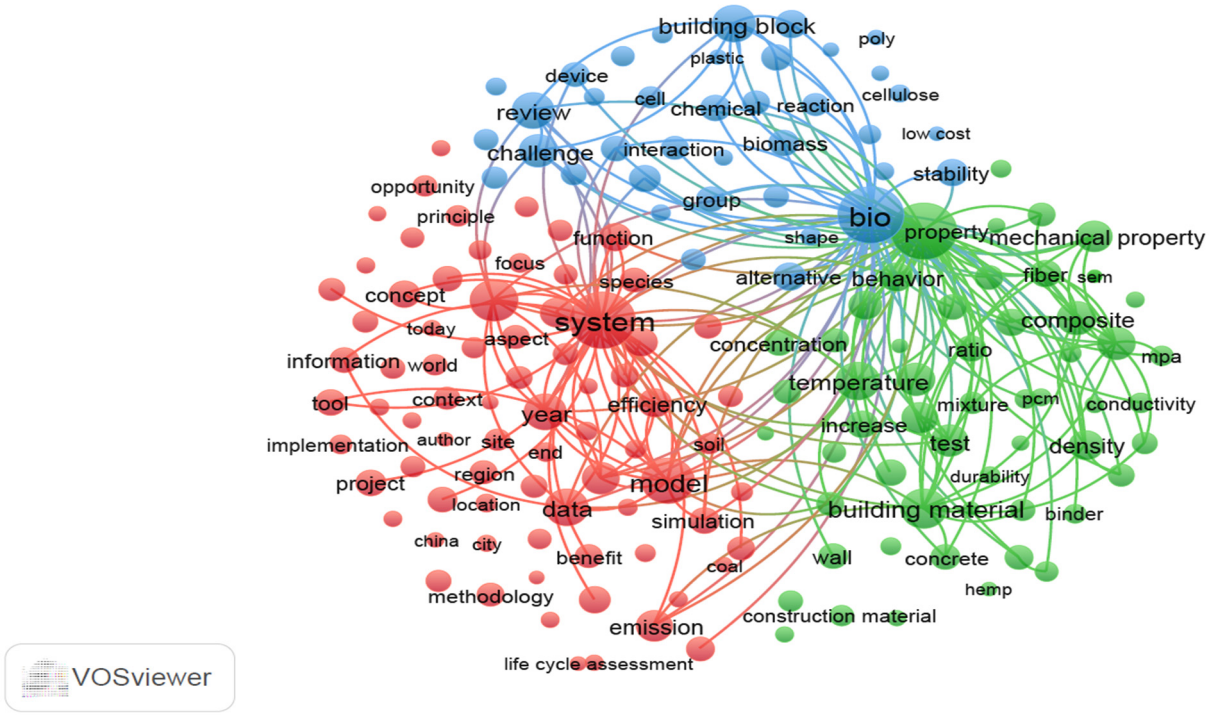

Figure 4. Keyword co-occurrence network using VOSviewer version 1.6.20 based on Table 3 terminologies.

3.3.2. Highlights on Co-Authorships and Citations

In terms of author collaboration, all researchers identified in the co-authorship network in the previous study [1] are also present in this study's network (Figure 5). However, the present analysis reveals a more expansive and interconnected network, including a larger number of authors who meet the threshold of five or more publications. This increased density reflects the broader and more inclusive search strategy employed in this study, allowing for a more comprehensive depiction of collaborative dynamics in the field.

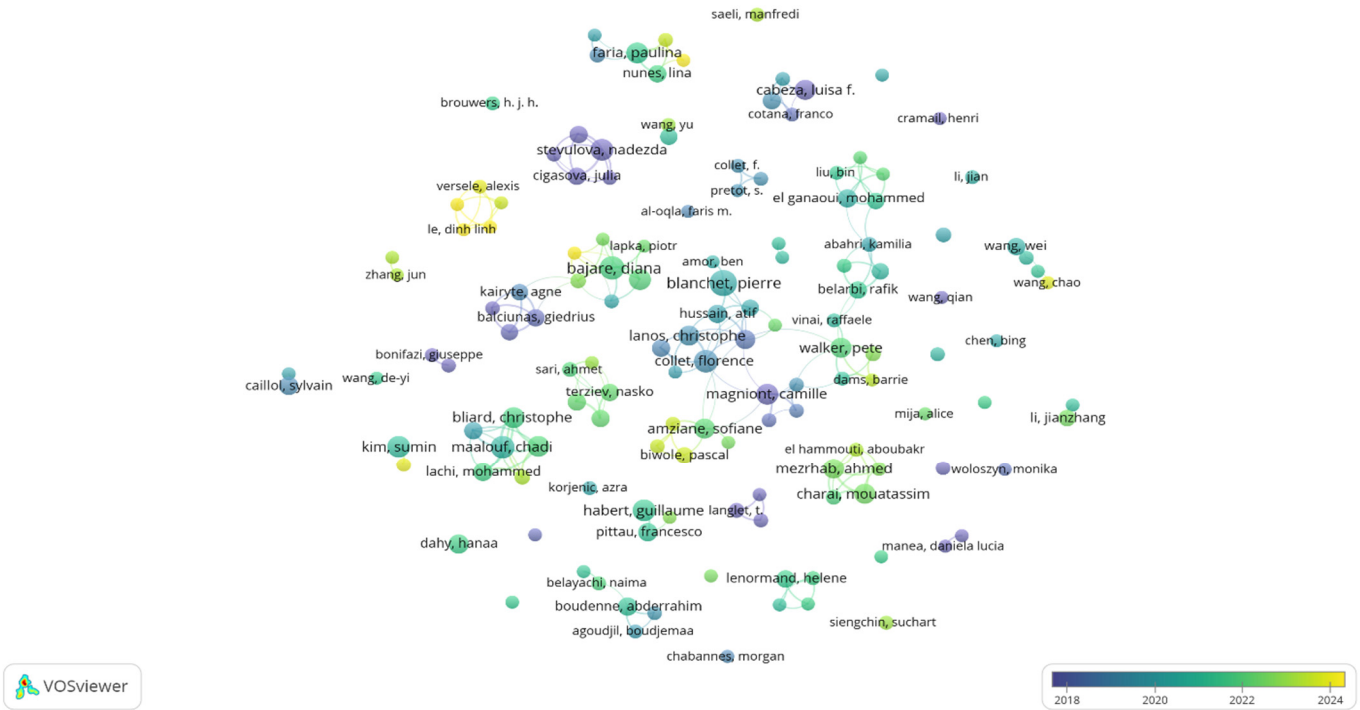

Figure 5. Co-authorship network using VOSviewer version 1.6.20 based on Table 3 terminologies.

Similarly, the co-citation network presented here in Figure 6 largely overlaps with the one published previously [1], yet it includes a wider array of publications and more interconnections. Both studies identify the most frequently cited publications as being concentrated between 2012 and 2018. This consistency suggests a strong foundational period for research in the field of bio-based building materials during those years.

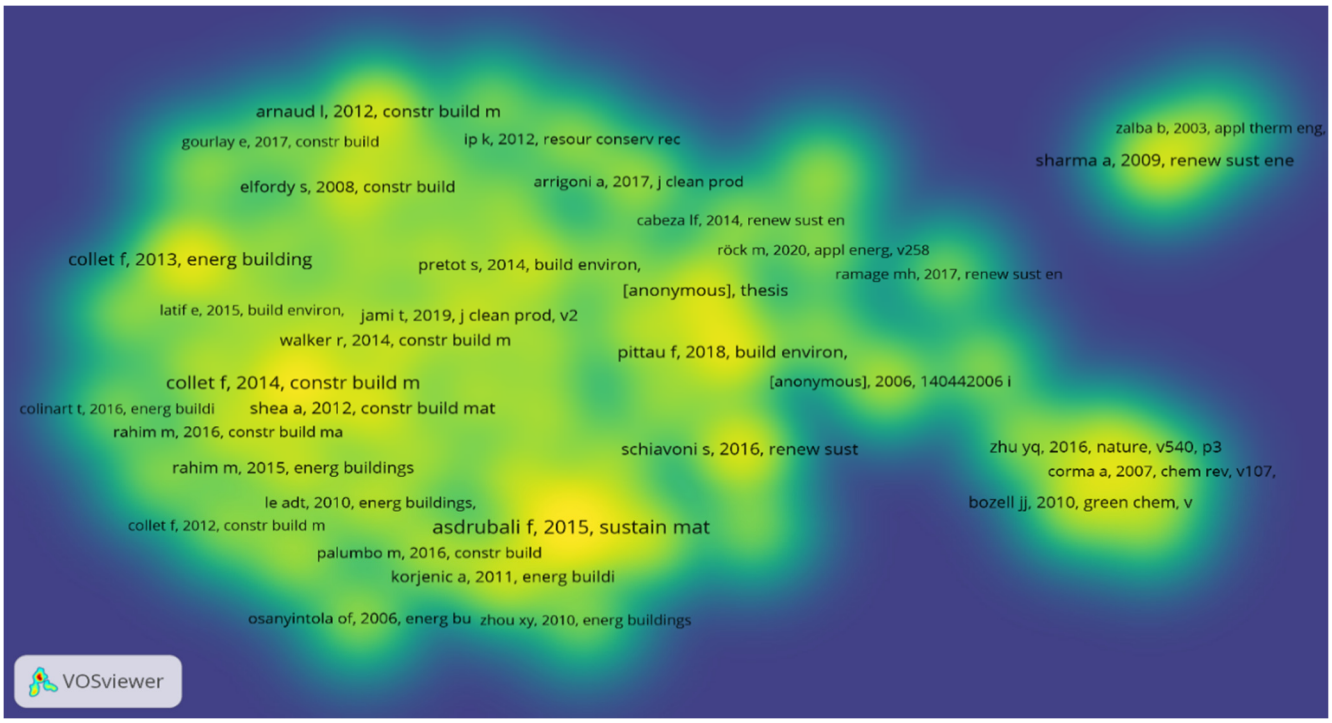

Figure 6. Co-citation network using VOSviewer version 1.6.20 based on Table 3 terminologies.

4. Bibliometric Results in Probabilistic Analysis

A bibliometric analysis is performed on the probabilistic approach to examine research activity in this field. The following subsections summarize the findings obtained from the Web of Science database and the insights generated through VOSviewer version 1.6.20 analysis.

4.1. Results from Web of Science

The following subsections present the publication trends, geographic distribution, and leading authors contributing to the field.

4.1.1. Publication Trends

Using the Web of Science (WoS) database, a total of 5485 publications were identified from 1900 to the present, based on a set of synonyms related to probabilistic analysis in building and construction materials (Table 6). The search terms used in this study are detailed in Table 6.

Table 6. Number of publications based on various terminology related to probabilistic analysis.

Topic	Year Published	Number of Publications
"Probabilistic Analysis in Building material" or "Stochastic Analysis in Building material" or "Statistical Analysis in Building material" or "Probabilistic Analysis in Construction material" or "Stochastic Analysis in Construction material" or "Statistical Analysis in Construction material"	1900–2025	5485

Based on the collected dataset, Figure 7 illustrates the temporal distribution of publications. The results show a clear upward trend, with publication numbers exceeding 600 in 2024, reflecting growing interest in probabilistic approaches as effective tools for addressing uncertainties and variabilities in material performance, more specifically

within the context of sustainable and bio-based construction materials. The fitted exponential trend line yields an R^2 value of 0.962, indicating a strong correlation between publication year and research output. This statistical fit confirms that the observed growth pattern is not random but represents a significant and consistent increase in scholarly attention to probabilistic methods over time.

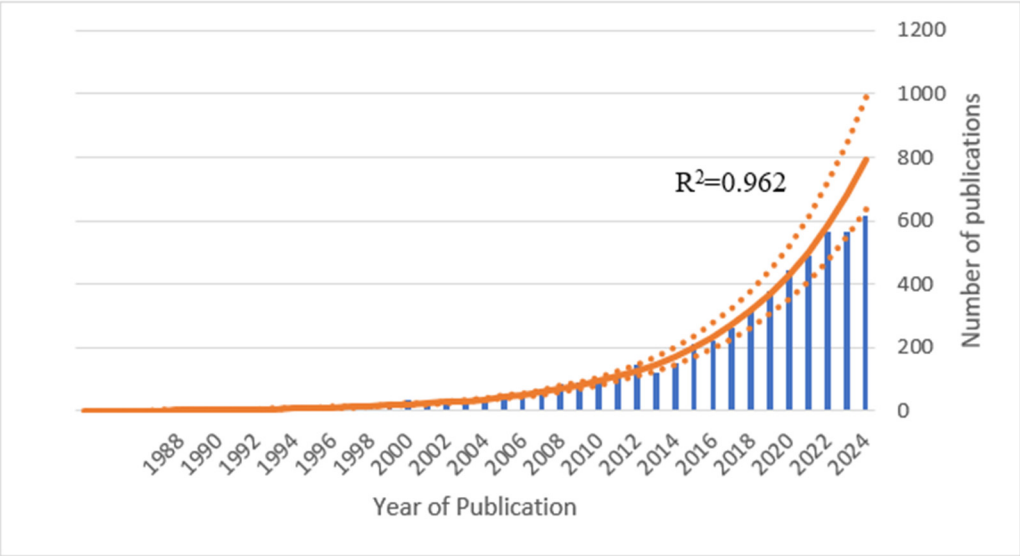

Figure 7. Number of publications over the years using WoS based on Table 6 terminologies, with fitted exponential trend and 95% confidence interval.

4.1.2. Geographic Distribution

The geographic analysis of publication outputs, shown in Figure 8, reveals that China and USA are the leading contributors by a significant margin. This trend reflects these countries' strong institutional investments and research capacity in this field. However, this dominance may also be partially influenced by language bias. As Web of Science predominantly indexes English language journals, contributions from non-English-speaking regions may be underrepresented, potentially skewing the global representation of research activity in this field.

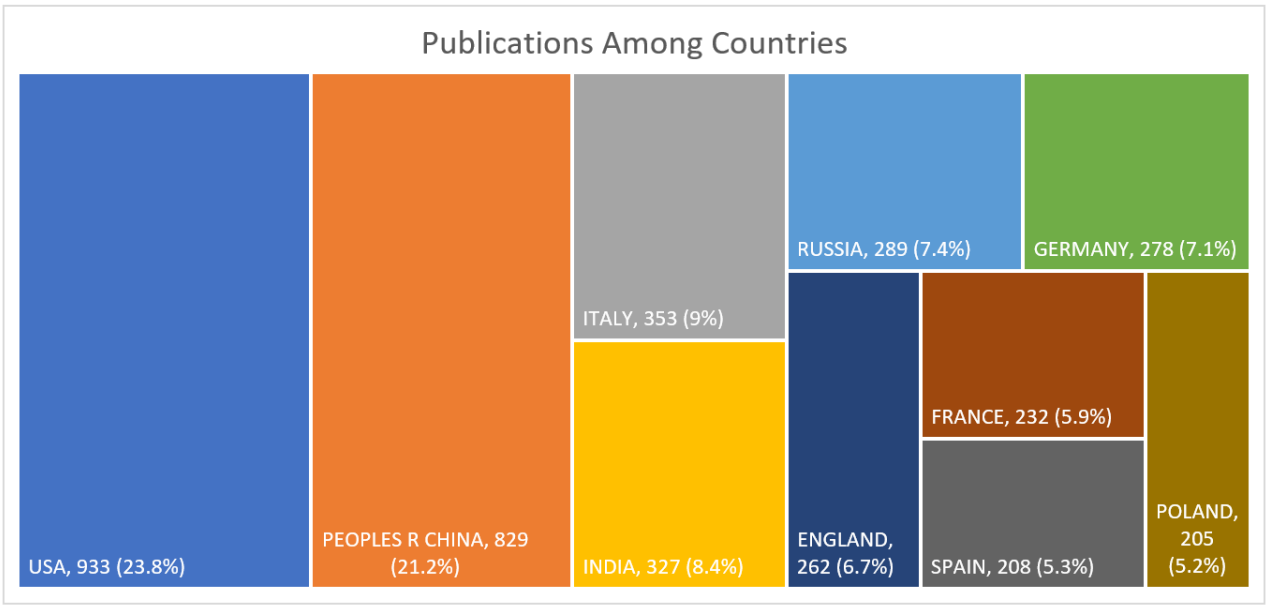

Figure 8. Number and percentage of publications using WoS based on Table 6 terminologies for the top 10 countries.

4.1.3. Top Authors

Table 7 presents the top 10 contributing authors based on the number of publications related to probabilistic approaches. The distribution suggests that the field is marked by a diverse and competitive author base, with many

researchers contributing at similar levels. Rather than being dominated by a few key figures, the field is driven by a broad and active group of experts, reflecting a vibrant and collaborative research environment.

Table 7. Number of publications for the top 10 contributing authors using WoS (based on Table 6 terminologies).

Author	Number of Publications	Institution	Country
Amin Muhammad Nasir	22	King Faisal University	Saudi Arabia
De Brito Jorge	18	University of Lisbon, CERIS	Portugal
Mohammed Ahmed Salih	17	University of Sulaimani	Iraq
Javed Muhammad Faisal	15	GIK Institute Engineering Science & Technology	Pakistan
Ahmad Waqas	14	Government College University Faisalabad	Pakistan
Althoey Fadi	12	Najran University	Saudi Arabia
Hanfi Mohammed	12	Ural Federal University	Russia
Stewart Mark G	12	University of Technology, Sydney	Australia
Chandrasekaran Akhila	11	SSN College of Engineering	India
Ahmad Ayaz	11	COMSATS University Islamabad	Pakistan

4.2. VOSviewer Version 1.6.20 Analysis

The following subsections present the co-occurrence keyword map and highlight the main co-authorship and citation networks.

4.2.1. Co-Occurrence Keyword Map

The keyword co-occurrence analysis (Figure 9) reveals three primary thematic clusters within the field of probabilistic analysis: (i) Red Cluster: Focused on data collection methods, encompassing terms related to surveys, experiments, and data sources. (ii) Green Cluster: Associated with material properties and parameterization, including references to structural and physical characteristics. (iii) Blue Cluster: Centered on simulation and uncertainty analysis, with emphasis on probabilistic modeling, uncertainty propagation, and simulation outputs. This clustering reflects the structured workflow typically involved in probabilistic studies. Beginning with data collection, followed by parameter characterization, and ending with uncertainty propagation interpretation.

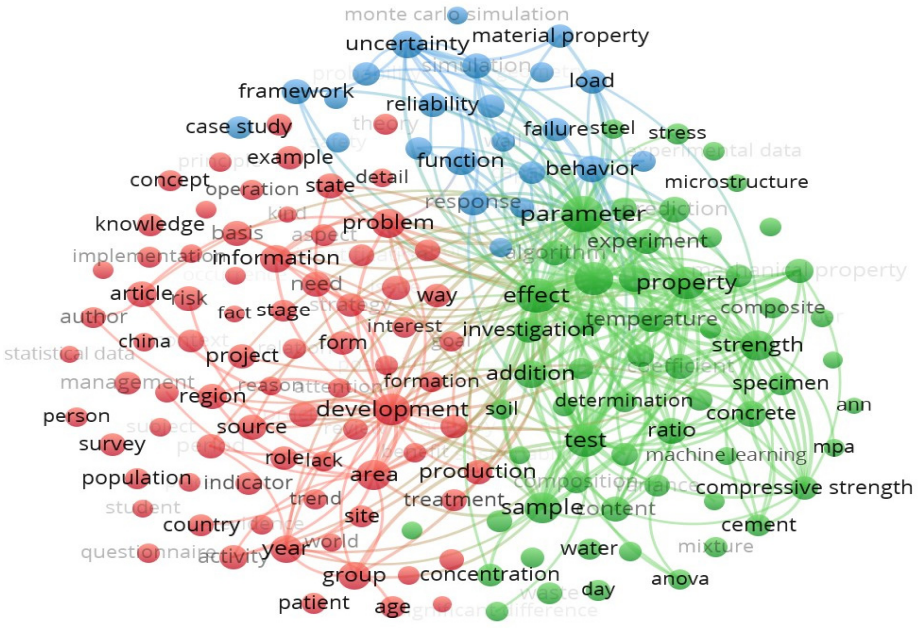

Figure 9. Co-occurrence network of keywords using VOSviewer version 1.6.20 based on Table 6 terminologies.

4.2.2. Highlights on Co-Authorships and Citations

The co-authorship network (Figure 10) highlights a well-connected group of researchers working on probabilistic methods in materials science. Despite the specificity of the search terms used, the results reveal a robust publication volume and a dense network of collaborative links. This highlights both the advancement of the field and the growing interest it continues to generate within the scientific community.

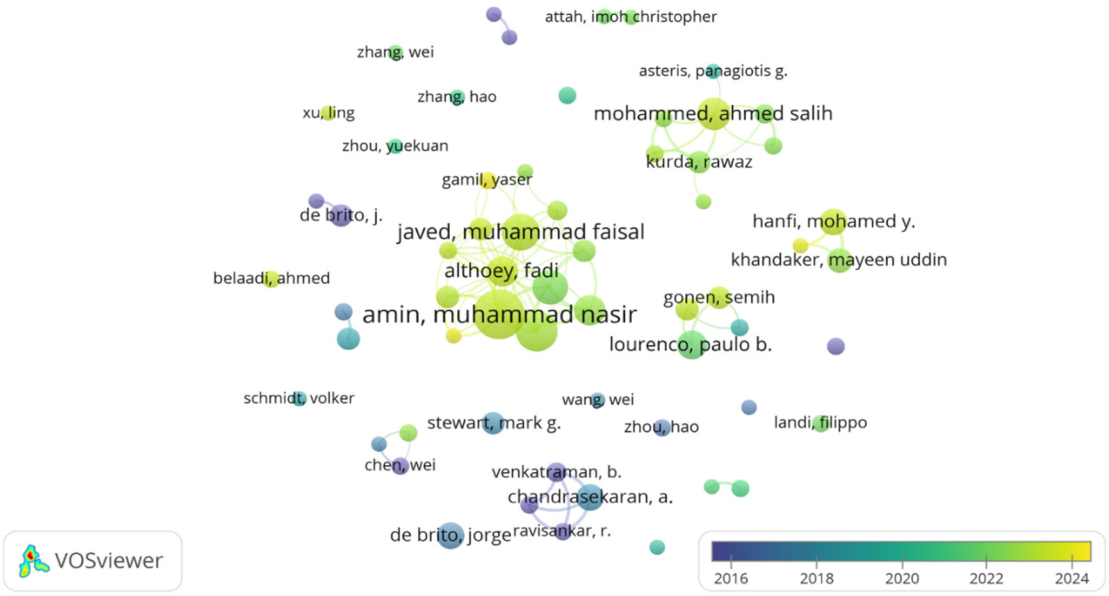

Figure 10. Co-authorship network using VOSviewer version 1.6.20 based on Table 6 terminologies.

The co-citation network presented in Figure 11 demonstrates that many foundational references in probabilistic research date back over two decades. These earlier works continue to be highly cited, confirming their influence. At the same time, the presence of recent publications indicates that the field is actively evolving. This juxtaposition of legacy and contemporary research highlights the dynamic and sustained growth of probabilistic approaches in materials science and construction engineering.

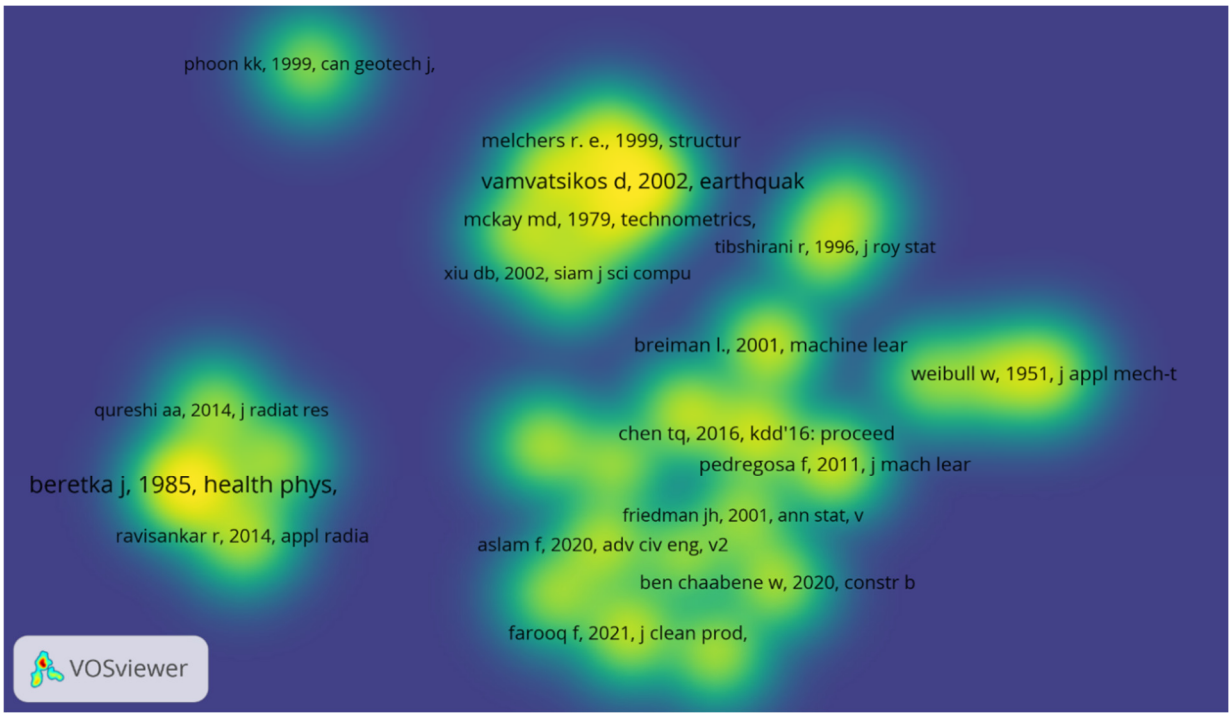

Figure 11. Co-citation network using VOSviewer version 1.6.20 based on Table 6 terminologies.

5. Intersectional Analysis: Bio-Based Building Materials & Probabilistic Approaches

The intersectional analysis between probabilistic analysis and bio-based building materials addresses the core problem of this research. The following subsections present the number and evolution of intersecting publications, followed by a focus on research gaps and future directions.

5.1. Number and Evolution of Intersecting Publications

A core objective of this study is to investigate the intersection between probabilistic analysis and bio-based building materials within the academic literature. To achieve this, an initial broad query was conducted using the Web of Science (WoS) database by combining the full set of keywords listed in Tables 3 and 6. However, this comprehensive search yielded no results, likely due to limitations in the database's capacity to handle complex Boolean combinations involving multiple keyword sets. To overcome this limitation, a refined approach was adopted wherein keywords from both domains were paired two at a time. The results of this targeted search are presented in Table 8, which details the number of publications identified for each specific keyword combination.

Table 8. Number of publications based on various combinations of terminologies.

Topic	Year Published	Number of Publications
Bio-based Building Material and Probabilistic Analysis	1900–2025	2
Bio Sourced Building Material and Probabilistic Analysis	1900–2025	0
Bio Renewable Building Material and Probabilistic Analysis	1900–2025	0
Plant Based Building Material and Probabilistic Analysis	1900–2025	6
Bio-based Construction Material and Probabilistic Analysis	1900–2025	2
Bio Sourced Construction Material and Probabilistic Analysis	1900–2025	0
Bio Renewable Construction Material and Probabilistic Analysis	1900–2025	0
Plant Based Construction Material and Probabilistic Analysis	1900–2025	4
Bio-based Building Material and Statistical Analysis	1900–2025	19
Bio Sourced Building Material and Statistical Analysis	1900–2025	3
Bio Renewable Building Material and Statistical Analysis	1900–2025	3
Plant Based Building Material and Statistical Analysis	1900–2025	36
Bio based Construction Material and Statistical Analysis	1900–2025	17
Bio Sourced Construction Material and Statistical Analysis	1900–2025	3
Bio Renewable Construction Material and Statistical Analysis	1900–2025	5
Plant Based Construction Material and Statistical Analysis	1900–2025	38
Bio-based Building Material and Stochastic Analysis	1900–2025	0
Bio Sourced Building Material and Stochastic Analysis	1900–2025	0
Bio Renewable Building Material and Stochastic Analysis	1900–2025	0
Plant Based Building Material and Stochastic Analysis	1900–2025	6
Bio-based Construction Material and Stochastic Analysis	1900–2025	0
Bio Sourced Construction Material and Stochastic Analysis	1900–2025	0
Bio Renewable Construction Material and Stochastic Analysis	1900–2025	0
Plant Based Construction Material and Stochastic Analysis	1900–2025	2

It is important to note that this method introduces potential for duplicate records across keyword combinations. Thus, while the total of search results is 147, the actual number of unique publications is likely lower. To ensure accuracy, all retrieved articles were reviewed manually to identify and eliminate duplicates.

Following this review, a total of 113 unique publications were found to explicitly address the intersection of probabilistic analysis and bio-based building materials. Refer to Figure 1 for the detailed screening process. A closer examination of these studies reveals that most rely on simulation-based probabilistic techniques, including Monte Carlo simulations, statistical uncertainty quantification, sensitivity analysis, and Bayesian inference, to evaluate variability and predict the performance of bio-based materials [26–28]. This filtered dataset serves as the foundation for subsequent discussion and highlights the current state of interdisciplinary research between these two critical fields. This iterative search methodology represents a novel contribution, as few bibliometric studies explicitly pair terminologies across two distinct research domains to reveal overlooked intersections.

5.2. Research Gaps and Future Directions

The findings underscore a significant research gap: despite the considerable volume of literature available on bio-based building materials and probabilistic analysis individually, studies that integrate both topics remain scarce. This lack of integration is surprising, given the inherent variability and uncertainty associated with bio-based building materials. This variability may only be addressed using probabilistic approaches.

The evident disconnect suggests a missed opportunity in current research practice. Bio-based building materials exhibit substantial heterogeneity due to factors such as biological origin, growth conditions, and processing methods. These characteristics naturally lend themselves to stochastic modeling and uncertainty quantification. However, the absence of probabilistic approaches in the existing literature represents a methodological risk, as ignoring variability can lead to misleading conclusions, overgeneralized findings, and potentially unreliable design recommendations.

By identifying this gap, the present study aims to raise awareness and encourage future research efforts to adopt probabilistic approaches when analyzing bio-based building materials, thereby contributing to the development of more reliable, efficient, and sustainable building solutions.

6. Conclusions

This bibliometric review provides a comprehensive examination of the academic literature related to bio-based building materials and probabilistic approaches, with a particular focus on identifying areas of intersection between the two. The study offers two principal contributions to the field. First, it provides practical insights into the design of effective search strategies for bibliometric studies. By systematically exploring keyword selection and term combinations, this study demonstrates how to overcome common challenges in bibliometric data collection.

Second, and more significantly, this review reveals a clear research gap in the literature concerning the intersection of bio-based building materials and probabilistic approaches. While both topics are individually well-established and the subject of extensive academic inquiry, their convergence remains remarkably underexplored. This is especially notable given the inherent variability in bio-based building materials, a characteristic that aligns naturally with the strengths of probabilistic modeling. The findings underscore the importance of integrating probabilistic methods into sustainable materials research. Bio-based building materials, shaped by natural and environmental factors, exhibit a level of uncertainty that deterministic approaches struggle to capture accurately. Probabilistic techniques offer a more realistic framework for understanding, predicting, and optimizing the performance of these materials in construction applications.

Thus, it is essential that future research bridges this gap by embracing uncertainty quantification, reliability assessment, and sensitivity analysis as standard tools in the study of sustainable construction materials. This would enable researchers to move beyond deterministic assumptions and toward more realistic decisions.

Encouraging the adoption of probabilistic methods will not only improve the reliability and predictability of sustainable materials but will also accelerate their acceptance and application in real-world construction projects. As the global demand for environmentally friendly materials continues to rise, integrating probabilistic analysis into the core of materials research becomes an urgent and strategic necessity.

The integration of probabilistic approaches carries practical implications for both industry and policy. For construction practitioners, it can lead to more informed decision-making, risk-aware design choices, and optimized resource allocation. For policymakers, it provides a stronger evidence base to support regulations and incentives that promote the adoption of bio-based building materials. As the global demand for environmentally friendly materials continues to rise, embedding probabilistic analysis into the core of materials research becomes not only a scientific necessity but also a strategic pathway toward more sustainable and resilient construction practices.

Looking ahead, the forthcoming phase of this research will focus on quantitative validation of bio-based material properties using Bayesian inference, Monte Carlo simulation, and Polynomial Chaos Expansion (PCE) for sensitivity analysis. These approaches will rigorously quantify uncertainties, identify the most influential parameters, and provide actionable insights to optimize material performance and support risk-informed design in sustainable construction.

Statement of the Use of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the author(s) used ChatGPT in order to improve language clarity and readability. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgments

The authors acknowledge Séverine Rosa Latapie for her input related to the comparative Analysis. No external administrative, technical, or financial support was received.

Author Contributions

Conceptualization, S.D., A.A.C., T.A.B. and J.A.; Methodology, S.D.; Software, S.D.; Validation, S.D., A.A.C., T.A.B. and J.A.; Formal Analysis, S.D.; Investigation, S.D.; Resources, A.A.C., T.A.B. and J.A.; Data Curation, S.D.; Writing—Original Draft Preparation, S.D.; Writing—Review & Editing, S.D., A.A.C., T.A.B. and J.A.; Visualization, S.D.; Supervision, A.A.C., T.A.B. and J.A.

Ethics Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

All data generated or analyzed during this study are included in this published article. Additional information can be provided upon reasonable request from the corresponding author.

Funding

This research received no external funding.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Latapie SR, Abou-Chakra A, Sabathier V. Bibliometric Analysis of Bio-and Earth-Based Building Materials: Current and Future Trends. *Constr. Mater.* **2023**, *3*, 474–508. doi:10.3390/constrmater3040031.
2. Lagouin M, Magniont C, Sénéchal P, Moonen P, Aubert JE, Laborel-préneron A. Influence of Types of Binder and Plant Aggregates on Hygrothermal and Mechanical Properties of Vegetal Concretes. *Constr. Build. Mater.* **2019**, *222*, 852–871. doi:10.1016/j.conbuildmat.2019.06.004.
3. Mouton L, Allacker K, Röck M. Bio-Based Building Material Solutions for Environmental Benefits over Conventional Construction Products-Life Cycle Assessment of Regenerative Design Strategies (1/2). *Energy Build.* **2023**, *282*, 112767. doi:10.1016/j.enbuild.2022.112767.
4. Ürge-Vorsatz D, Khosla R, Bernhardt R, Chan YC, Vérez D, Hu S, et al. Advances Toward a Net-Zero Global Building Sector. *Annu. Rev. Environ. Resour.* **2020**, *45*, 227–269. doi:10.1146/annurev-environ-012420-045843.
5. Ali KA, Ahmad MI, Yusup Y. Issues, Impacts, and Mitigations of Carbon Dioxide Emissions in the Building Sector. *Sustainability* **2020**, *12*, 7427. doi:10.3390/su12187427.
6. Huang G. Prediction and Optimization of the Mechanical and Thermal Properties of Bio-Based Concrete: Analytical and Numerical Multi-Scale Approaches. Doctoral Dissertation, INSA de Toulouse, Toulouse, France, 2023.
7. Bourbia S, Kazeoui H, Belarbi R. A Review on Recent Research on Bio-Based Building Materials and Their Applications. *Mater. Renew. Sustain. Energy* **2023**, *12*, 117–139. doi:10.1007/s40243-023-00234-7.
8. Magniont C, Escadeillas G, Coutand M, Oms-Multon C. Use of Plant Aggregates in Building Ecomaterials. *Proc. Eur. J. Environ. Civ. Eng.* **2012**, *16*, s17–s33. doi:10.1080/19648189.2012.682452.
9. Séverine R. Modélisation et Optimisation Des Performances Thermiques Des Bio et Géosourcés Par Approche Multi-Échelle : Apport à La Valorisation d'une Large Gamme de Co-Produits. Ph.D. Thesis, Université Toulouse III-Paul Sabatier, Toulouse, France, 2024.
10. Al-Bittar T. Probabilistic Analysis of Shallow Foundations Resting on Spatially Varying Soils. Doctoral Dissertation, UNIVERSITÉ DE NANTES, Nantes, France, 2012.

11. Mao N, Al-Bittar T, Soubra AH. Probabilistic Analysis and Design of Strip Foundations Resting on Rocks Obeying Hoek-Brown Failure Criterion. *Int. J. Rock Mech. Min. Sci.* **2012**, *49*, 45–58. doi:10.1016/j.ijrmms.2011.11.005.
12. Ghavami S, Naseri H, Safi Jahanshahi F. Enhanced Prediction and Uncertainty Modeling of Pavement Roughness Using Machine Learning and Conformal Prediction. *Infrastructures* **2025**, *10*, 166. doi:10.3390/infrastructures10070166.
13. Youssef E, Cornou C, Youssef D, Massih A, Al-Bittar T. Nonstationary Shear-Wave Velocity Randomization Approach to Propagate Small-Scale Spatial Shear-Wave Velocity Het-Erogeneities into Seismic Response. *J. Geotech. Geoenviron. Eng.* **2024**, *150*, 10. doi:10.1061/JGGEFK.GTENG-11884.
14. Marsh E, Hattam L, Allen S. Stochastic Error Propagation with Independent Probability Distributions in LCA Does Not Preserve Mass Balances and Leads to Unusable Product Compositions-a First Quantification. *Int. J. Life Cycle Assess.* **2025**, *30*, 221–234. doi:10.1007/s11367-024-02380-0.
15. Boshoff N, Akanmu MA. Scopus or Web of Science for a Bibliometric Profile of Pharmacy Research at a Nigerian University? *S. Afr. J. Libr. Inf. Sci.* **2018**, *83*, 12–22. doi:10.7553/83-2-1682.
16. Van Eck NJ, Waltman L, Dekker R, Van Den Berg J. A Comparison of Two Techniques for Bibliometric Mapping: Multidimensional Scaling and VOS. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 2405–2416. doi:10.1002/asi.21421.
17. Liu Y, He H. Scientometrics of Scientometrics Based on Web of Science Core Collection Data between 1992 and 2020. *Information* **2023**, *14*, 637. doi:10.3390/info14120637.
18. Mongeon P, Paul-Hus A. The Journal Coverage of Web of Science and Scopus: A Comparative Analysis. *Scientometrics* **2016**, *106*, 213–228. doi:10.1007/s11192-015-1765-5.
19. Abrizah A, Zainab AN, Kiran K, Raj RG. LIS Journals Scientific Impact and Subject Categorization: A Comparison between Web of Science and Scopus. *Scientometrics* **2013**, *94*, 721–740. doi:10.1007/s11192-012-0813-7.
20. Tarazi A. Comparative Analysis of the Bibliographic Data Sources Using PubMed, Scopus, Web of Science, and Lens. *High Yield Med. Rev.* **2024**, *2*, 1–10. doi:10.59707/hymrUNHW4628.
21. Singh VK, Singh P, Karmakar M, Leta J, Mayr P. The Journal Coverage of Web of Science, Scopus and Dimensions: A Comparative Analysis. *Scientometrics* **2021**, *126*, 5113–5142. doi:10.1007/s11192-021-03948-5.
22. Anker MS, Hadzibegovic S, Lena A, Haverkamp W. The Difference in Referencing in Web of Science, Scopus, and Google Scholar. *ESC Heart Fail* **2019**, *6*, 1291–1312. doi:10.1002/ehf2.12583.
23. Wouters P, Costas R. *Users, Narcissism and Control-Tracking the Impact of Scholarly Publications in the 21st Century*; SURFfoundation: Utrecht, The Netherlands, 2012.
24. Clermont M, Dyckhoff H. Coverage of Business Administration Literature in Google Scholar Analysis and Comparison with EconBiz, Scopus and Web of Science 1. *Bibliometr.-Prax. Und Forsch.* **2012**, *1*, 1–54. doi:10.2139/ssrn.2016850.
25. Kirby A. Exploratory Bibliometrics: Using VOSviewer as a Preliminary Research Tool. *Publications* **2023**, *11*, 10. doi:10.3390/publications11010010.
26. Bui R, Goffart J, McGregor F, Woloszyn M, Fabbri A, Grillet AC. Uncertainty and Sensitivity Analysis Applied to a Rammed Earth Wall: Evaluation of the Discrepancies between Experimental and Numerical Data. *E3S Web Conf.* **2020**, *172*, 17004. doi:10.1051/e3sconf/202017217004.
27. Shi D, Xu Y, Demartino C, Xiao Y, Spencer BF. Cyclic Behavior of Laminated Bio-Based Connections with Slotted-in Steel Plates: Genetic Algorithm, Deterministic Neural Network-Based Model Parameter Identification, and Uncertainty Quantification. *Eng. Struct.* **2024**, *310*, 118114. doi:10.1016/j.engstruct.2024.118114.
28. Migoni Alejandre E, Koskamp G, van de Leur M, Wandl A, van Timmeren A. Quantifying the Life Cycle Emissions of Hybrid Structures with Advanced Bio- and Conventional Materialization for Low-Embodied Carbon Urban Densification of the Amsterdam Metropolitan Area. *J. Clean. Prod.* **2024**, *483*, 144273. doi:10.1016/j.jclepro.2024.144273.