ABORDAGENS TEÓRICAS E PRÁTICAS EM PESQUISA

COORDENADORES

Patricia Bieging

Raul Inácio Busarello

ISBN 978-85-7221-547-3

2025

Diego Durante Muhl

TECHNOLOGICAL TRENDS, GEOGRAPHICAL CONCENTRATION AND INNOVATION CYCLES IN BIOREFINERY PATENTS

DOI: 10.31560/pimentacultural/978-85-7221-547-3.3

ABSTRACT:

The global demand for renewable energy sources and the urgency to reduce reliance on fossil fuels have driven the development of more sustainable and regenerative technological pathways. Recent research has underscored the importance of biorefineries in the production of biofuels and industrially relevant compounds, emphasising the need to integrate these processes with effluent and waste treatment technologies. Despite the existing body of knowledge, a gap remains in the systematic and comprehensive analysis of the patent landscape related to various biorefinery routes. This study therefore aims to examine patents concerning biorefineries and to identify emerging technological trajectories. To this end, a methodological framework was developed, encompassing a systematic search of documents via Google Patents and the application of statistical and visualisation techniques within the R environment. The findings reveal a convergence between traditional chemistry and the growing demand for sustainability and carbon-negative solutions. Numerous patents focus on biomass and lignin, indicating increasing interest in biomass valorisation and cleaner process integration. Notably, CPC subclasses associated with physicochemical processes (B01, C07), biotechnology, and climate-change mitigation (C12, Y02) account for a significant proportion of recent filings. This suggests a synergy between advanced catalysis, the development of novel carbon-storing materials, and the adoption of multiproduct biorefinery approaches. The study underscores the potential of biorefineries to contribute to circular and regenerative value chains, offering a foundation for strategic decision-making in research and development, public policy, and sustainable technology investment.

Keywords: Biorefineries; Patents; Biomass; Sustainability; Bioeconomy.

1. INTRODUCTION

The growing global demand for renewable energy sources and the urgency to reduce reliance on fossil fuels have driven the development of more sustainable and regenerative technological pathways. Within this context, biorefineries have emerged as a promising strategy for the production of bioenergy, fuels, and high value-added chemical compounds from a diverse array of biomass feedstocks, including lignocellulosic residues, microalgae, agricultural and livestock waste, and industrial by-products. The capacity to integrate multiple processes within a single facility enhances both the economic viability and environmental performance of biorefineries, thereby contributing to the advancement of circular, carbon-negative, and resource-efficient value chains.

The growing global demand for renewable energy sources and the urgency to reduce reliance on fossil fuels have driven the development of more sustainable and regenerative technological pathways. Within this context, biorefineries have emerged as a promising strategy for the production of bioenergy, fuels, and high value-added chemical compounds from a diverse array of biomass feedstocks, including lignocellulosic residues, microalgae, agricultural and livestock waste, and industrial by-products. The capacity to integrate multiple processes within a single facility enhances both the economic viability and environmental performance of biorefineries, thereby contributing to the advancement of circular, carbon-negative, and resource-efficient value chains.

Recent studies have underscored the relevance of biorefineries in the production of biofuels and value-added industrial compounds, emphasising the importance of integrating these systems with wastewater and solid waste treatment technologies (Khoshnevisan *et al.*, 2021; Pessoa *et al.*, 2021). Innovation efforts also encompass high-rate vermicomposting methodologies for organic waste valorisation (Tauseef *et al.*, 2023), as well as catalytic and biochemical pathways for converting lignin into high-value hydrocarbons, such as aviation kerosene (Shen *et al.*, 2019). Several studies have demonstrated the feasibility of hybrid conversion routes capable of producing both fuels and bio-based chemical products—including organic acids, polyamides, and monomers (Dimian *et al.*, 2019)—thereby exemplifying the adoption of multiproduct configurations designed to maximise biomass utilisation and facilitate low-carbon, circular transitions.

Moreover, techno-economic assessments have demonstrated that the co-production of bio-kerosene and ethanol in lignocellulosic biorefineries can enhance the economic viability of existing facilities (Shen *et al.*, 2019). In parallel, other studies have examined the large-scale application of biotechnologies for the valorisation of wastes and by-products, including microalgae-based processes (Moscoviz *et al.*, 2018; Yadav *et al.*, 2022). Patent activity in the biorefinery sector indicates significant progress in converting biomass-derived sugars into high-value compounds, developing specialised reactors for vermicomposting, and implementing novel methods for recovering previously discarded by-products (Barragan-Ocana *et al.*, 2023; Tauseef *et al.*, 2023). Collectively, these contributions reflect a global shift towards interdisciplinary convergence, aimed at maximising the value extracted from each biomass fraction and bolstering the technological and economic competitiveness of biorefinery systems (Faba *et al.*, 2015; Tsvetanova *et al.*, 2024).

Despite this growing and diversified body of research, important questions remain unanswered. Specifically, there is a clear gap in the systematisation and in-depth analysis of the patent landscape relating to biorefinery development. While the existing literature is substantial, it is often fragmented into specific case studies or isolated evaluations of individual technologies, lacking a comprehensive overview of key actors, CPC classifications, and the spatio-temporal patterns that shape innovation in this field. There is a pressing need for structured mapping capable of identifying cross-cutting technological trends and pinpointing areas of concentrated innovation activity aligned with circular and carbon-negative solutions.

In light of these limitations, this study is guided by the following research questions: **RQ1:** How is the productivity of authors, assignees, and countries distributed, and what patterns of concentration or dispersion can be observed within the biorefinery patent landscape? **RQ2:** How have the volume and timing of biorefinery patent filings evolved, and what major historical phases of technological development can be identified? **RQ3:** Which CPC subclasses have demonstrated the most significant growth, and what are the predominant and emerging technological themes within these categories? **RQ4:** What are the principal technological challenges addressed in biorefinery patents, and how have key concepts and innovation pathways shifted over time?

To address these questions, the primary objective of this study is to examine the technological development trajectories reflected in patents relating to biorefineries. To this end, four specific aims are pursued in alignment with the research questions: (i) to identify the key actors and analyse their roles within the global innovation landscape (RQ1, addressed in Section 3.1); (ii) to assess the distribution of patent productivity and patterns of concentration, with reference to Lotka's Law (RQ1, Section 3.2); (iii) to investigate the temporal evolution of patent filings and delineate the historical phases of technological development (RQ2, Section 3.3); and (iv) to examine changes in CPC classifications and identify the core technological concepts and emerging innovation pathways in biorefinery systems (RQ3 and RQ4, Sections 3.4 and 3.5).

The motivation for this study lies in the transformative potential of biorefineries to build sustainable and regenerative value chains, enabling the production of fuels, chemical inputs and high-value materials without depleting natural resources (Amidon *et al.*, 2011; Mühl *et al.*, 2024). An integrated understanding of who is driving innovation, which technologies are receiving the greatest investment, and how applications are evolving can inform funding strategies, public–private partnerships and the formulation of industrial policies. It is therefore hoped that this research will support the development of clearer pathways for scaling up and integrating biorefinery R&D strategies into the market, ultimately benefiting both the productive sector and society at large.

Beyond its technical and economic dimensions, this study is also aligned with the Sustainable Development Goals (SDGs) set out in the United Nations' 2030 Agenda. By promoting the efficient use of renewable resources, biorefineries contribute directly to SDG 7 (Affordable and Clean Energy) through the production of biofuels; to SDG 9 (Industry, Innovation and Infrastructure) by fostering clean and innovative industrial technologies; and to SDG 12 (Responsible Consumption and Production) by enabling waste valorisation and reducing environmental impact. By mapping technological trends in biorefinery-related patents, this study offers valuable insights to inform concrete actions in responsible and sustainable innovation—critical for addressing today's energy, environmental and climate challenges.

2. METHODOLOGY

To characterise patent activity in the field of biorefineries, a methodological framework was established, encompassing systematic document retrieval via Google Patents and the application of statistical and visual analyses within the R environment. The approach involved defining keyword-based search parameters, retrieving results segmented by CPC subclasses, cleaning and standardising the dataset, and implementing quality-assurance routines. The following steps outline how data collection and analysis were undertaken, detailing the construction of a unified database and the techniques employed to examine linguistic, temporal, geographical and technological classification aspects.

2.1. DATA COLLECTION

The study commenced with a search on the Google Patents platform using the keyword "biorefinery" to encompass a broad range of related technologies. The resulting search was refined into 19 subclasses based on the Cooperative Patent Classification (CPC) system. Each of these 19 queries combined the term "biorefinery" with a specific CPC subclass code, yielding an initial dataset of 13,639 patent documents across 11 variables, including identifier (ID), title, applicant, inventors and key dates (priority, filing, publication and grant), as well as hyperlinks to the original documents and representative figures.

2.2. DATA ANALYSIS

The first step involved transforming the dataset from a long to a wide format, enabling each patent entry to contain multiple columns for distinct CPC codes (Wickham, 2016). Language detection for patent titles was undertaken using probabilistic methods (Hornik et al., 2013), followed by the standardisation of linguistic variants and the generation of descriptive statistics on language distribution. Date fields (priority, filing, publication and grant) were converted to date format and used to extract year values, thereby allowing for the construction of historical series and the calculation of moving averages to highlight temporal trends (Chatfield, 2004).

Subsequently, a productivity analysis of authors, applicants and countries was undertaken, including character correction, standardisation of institutional abbreviations and name translation. Bar charts were produced to visualise patent-filing concentrations and to map key actors in the field. In parallel, Lotka's Law was applied to assess the distribution of productivity, using log–log scale transformations (Lotka, 1926; Potter, 1981). Co-occurrence heatmaps were also generated, linking authors, applicants and countries with CPC codes in order to identify potential technological synergies (Wilkinson & Friendly, 2009).

With regard to CPC classification, the records were restructured into a long format to quantify the annual frequency of each code and to support the construction of line graphs, stacked bar charts and area plots illustrating the evolution of subclasses over time (WIPO, 2025). Additional visualisations were created using the *ggplot2* package (Wickham, 2016) to highlight the relative importance of each CPC family.

To identify technological trends within each CPC subclass, patent titles were analysed. This process involved text normalisation (lowercasing, removal of punctuation and numbers, trimming of extra spaces and lemmatisation), followed by tokenisation and stop-word removal. These steps are fundamental to Natural Language Processing (NLP) techniques and are well documented in the literature (Manning *et al.*, 2008; Feldman & Sanger, 2006), ensuring consistent representation and minimising noise in subsequent analyses.

Publication cycles were defined on the basis of fluctuations observed in the temporal distribution of patent filings. The intervals were determined by identifying significant troughs in annual publication counts, with each decline followed by a reversal of trend interpreted as marking the end of one cycle and the beginning of the next. The following periods were identified: Cycle 1 (1900–1950): Initial period characterised by relative stability in publication volume. Cycle 2 (1951–1970): Period of limited growth. Cycle 3 (1971–1998): First major wave of patent publications. Cycle 4 (1999–2019): Second major wave of activity. Cycle 5 (2020–2025): Current period of focus.

Subsequently, the terms extracted from patent titles were aggregated and counted by publication cycle and CPC subclass, enabling the identification of the most frequent terms, as well as distinctions between common and subclass-specific terminology. These were visualised using bar charts and tile plots, facilitating the comparison of innovation patterns and technological trends over time. This integrated approach — combining text mining, statistical analysis and data visualisation — is grounded in classical patent analysis (Griliches, 1990) and contemporary innovation-management literature (Breschi & Malerba, 2023), offering a robust tool for identifying and monitoring technological trajectories.

Finally, frequency tables, language classifications and temporal statistics were structured and exported to support manuscript development and data transparency (Ooms, 2017). The results are presented from general to specific. First, the distribution of languages is examined to provide a global perspective on where and how patents are published. Next, the key actors — applicants and countries — are characterised, followed by a productivity analysis (Lotka's Law) and an assessment of the temporal behaviour of patent filings. Lastly, CPC codes are examined in detail, summarising how technological subclasses have evolved over time.

3. RESULTS AND DISCUSSION

Language detection was performed using the *textcat* library, which applies a probabilistic model based on character-sequence analysis of patent titles. English accounted for 85.98% of titles (7,637 records), followed by German (3.68%), Catalan (2.89%) and Danish (2.44%). Intermediate frequencies were observed for languages such as French, Scots and Latin, while Portuguese, Slovenian and Romanian appeared in smaller proportions.

The predominance of English is likely attributable to its widespread adoption in technical and scientific communication, as well as to its status as the standard language for international patent filings. The presence of German, Catalan, Danish and French suggests the existence of established research and innovation hubs in countries where these languages are spoken. By contrast, the relatively low occurrence of Portuguese and Romanian indicates that, although contributions emerge from diverse cultural contexts, the core of patent production remains concentrated in English-speaking countries or in regions that choose to publish in English to maximise global dissemination.

The growing use of English in both patent documents and scientific publications is frequently cited as a factor facilitating the international diffusion of research outputs (Pessoa et al., 2021). Scholars examining patent filings in the field of renewable energy have likewise observed that English is often employed to maximise both visibility and legal protection across multiple jurisdictions — an interpretation consistent with the distribution patterns identified in this study (Barragan-Ocana et al., 2023; Tauseef et al., 2023).

3.1. RQ1: KEY ACTORS

Information on authors, assignees and countries was consolidated through standardisation procedures, including abbreviation correction, name harmonisation and translation, in accordance with the long-format data-reorganisation approach described earlier. A co-occurrence matrix was subsequently generated, linking each entity (author, assignee or country) to the CPC subclasses in which it appeared, resulting in a unified heatmap. This visualisation illustrates the frequency of participation across key biorefinery subclasses, thereby facilitating the identification of major actors and potential areas of technological convergence.

As shown in Figure 1, countries (in green) display the largest frequency blocks, particularly in the cases of the United States, China, the WIPO (representing international filings), as well as South Korea and Japan. Regarding assignees (in blue), organisations and universities such as The Regents of the University of California, Poet Research Inc., and various biotechnology companies stand out, reflecting active engagement across multiple CPC subclasses. Individual authors (in red), although more numerous (19,926 in total), tend to exhibit lower frequency values, indicating a more diffuse contribution from inventors across the technological spectrum.

This comparison reinforces the view that patent dynamics in the biorefinery sector encompass both major research centres and established corporations (e.g. the United States, China, and key industrial assignees), as well as a broad array of dispersed inventors. The participation of 56 countries underscores the global nature of this technology, while the identification of 5,341 assignees reflects considerable institutional diversity. The substantial number of individual authors further highlights the collaborative character of patent development, suggesting that both research teams and independent inventors may specialise in specific CPC subclasses. This global overview illustrates the mobilisation of a wide range of actors in biorefinery innovation, with overlaps that reflect not only chemical and biotechnological advancements but also sustainability-driven initiatives.

Previous studies have highlighted the importance of consolidated hubs—such as the United States, China, and Europe—in driving the development of clean technologies (Tsvetanova, Rennings & Broering, 2024; Yadav et al., 2022), thereby supporting the concentration patterns observed in this study. Corporate initiatives, such as those led by The Regents of the University of California and LanzaTech, are frequently cited in the literature for their roles in fostering knowledge transfer between companies and universities (Amidon et al., 2011; Dimian et al., 2019). The wide dispersion of authors aligns with research that emphasises the involvement of diverse research groups and underscores the need for interdisciplinary collaboration (Khemthong et al., 2021; Moscoviz et al., 2018).

3.2. RQ1: PRODUCTIVITY ANALYSIS

The productivity of authors, assignees and countries was assessed on the basis of individual patent counts, with each entity ranked in descending order according to the number of publications. These empirical frequencies were then compared with the theoretical distribution proposed by Lotka's Law, which holds that a small number of entities account for the majority of scientific output, while the vast majority contribute only marginally. The observed curves were plotted against a reference line (exponent = 2), a value typical in Lotka-based modelling (Lotka, 1926; Potter, 1981).

In Figure 2, authors display a productivity peak concentrated within a small group of individuals—such as Theodora Retsina, Lan Tang and Junxin Duan—followed by a steep decline down the ranking, consistent with the concentration hypothesis. Among assignees, Novozymes Inc. and Novozymes A/S lead in patent counts, followed by LanzaTech New Zealand and The Regents of the University of California, illustrating a pattern in which a few organisations dominate the landscape. Regarding countries, the United States leads with 2,903 patents, followed by China (1,641), Japan (895) and WIPO, reflecting a strong geographical concentration. Although there is a long tail of countries with lower patent counts, only 56 countries are represented, indicating that most nations do not hold a single biorefinery-related patent.

Authors and assignees display lower levels of concentration at the top, whereas countries exhibit a sharper concentration pattern—more pronounced than that predicted by Lotka's Law. The convergence of the empirical curves with Lotka's theoretical line reinforces the view that biorefinery research follows a power-law distribution, whereby a small number of entities account for the majority of outputs. This behaviour reflects the influence of centres of excellence, established corporations and national innovation policies. In practical terms, partnerships with these leading actors may play a pivotal role in technology diffusion, although there remains considerable scope for greater participation by emerging researchers, institutions and countries.

This disparity between strong country-level dominance and the more balanced distribution among individual authors and assignees supports the findings of Tsvetanova et al. (2024) and Shen et al. (2019), who highlight the presence of powerful technology hubs in leading nations alongside more dispersed activity across technological subdomains (Amidon et al., 2011; Yadav et al., 2022). The partial alignment with Lotka's Law echoes the debates raised by Moscoviz et al. (2018) regarding the geographic concentration of innovation and underscores the importance of policy measures aimed at engaging emerging regions.

A more precise modelling approach was applied using a power-law function, implemented through linear regression on the logarithmic scale of patent counts by rank, with each entity (author, assignee or country) ordered by productivity. The fitted equation was:

patent_count = K / rank^β (Eq. 1)

This formulation enabled comparison of the estimated β coefficients with Lotka's classical value ($\beta \approx 2$), which indicates a high concentration of outputs among a small number of agents.

The results (Figure 3) show that, for authors ($\beta \approx 0.50$) and assignees ($\beta \approx 0.66$), the productivity decay was relatively gradual, suggesting lower levels of concentration at the top and a broader dispersion of inventors and organisations with moderate patent output. By contrast, for countries ($\beta \approx 2.37$), the curve steepens beyond

Lotka's classical threshold, indicating a sharp decline in productivity after the leading few and signalling the dominance of a highly concentrated group of nations.

Several studies confirm that this degree of national concentration exceeds what is typically observed among individual actors (Tsvetanova et al., 2024; Yadav et al., 2022), supporting the argument that leading countries benefit from superior innovation infrastructure and policy incentives conducive to high patent output (Amidon et al., 2011). Conversely, the dispersion among firms and inventors reflects the coexistence of multiple technological trajectories and a heterogeneous ecosystem of laboratories, start-ups and research groups (Faba, Diaz & Ordonez, 2015; Silva et al., 2020).

3.3. RQ2: TEMPORAL ANALYSIS

The temporal analysis was based on the extraction and standardisation of priority, filing, publication and grant dates, which were converted into date format, with the year of each event subsequently derived. Moving averages were calculated to smooth short-term fluctuations and to emphasise long-term trends. The resulting time series were plotted as separate lines, enabling comparison of the evolution of each temporal milestone associated with biorefinery-related patents.

Figure 4 shows the temporal distribution of priority, filing, publication and grant dates for biorefinery patents. Patent filings remained relatively stable until the mid-1970s, followed by moderate growth throughout the 1980s and a sharp increase from the early 2000s onwards. In more recent years, the publication curve (green) tends to slightly exceed the priority (blue) and filing (red) curves, while the grant line (purple) generally lags behind. Notable peaks occur around 2020, when publication volumes surpass 500 records, whereas priority and filing figures fluctuate between 400 and 600 documents.

The steep upward trend from the 2000s reflects intensified research and development efforts in the field of biorefineries, likely driven by renewable energy policies and advances in process engineering. The gap between the publication and grant lines highlights the natural time lag between patent application and official approval, as well as possible delays in administrative processing. Post-2020 fluctuations may be attributable to incomplete datasets, delays in the availability of documents or shifts in publication patterns—yet they nonetheless illustrate sustained and growing interest in the topic.

The surge observed in the 2000s has been attributed to policy incentives for renewable energy and technological breakthroughs (Barbosa et al., 2024; Faba, Diaz & Ordonez, 2015). Amidon et al. (2011) highlight the integration of multiple products as a key factor in enhancing the economic viability of biorefineries—an insight that aligns with the sharp increase in patent filings during this period. The recent plateau may indicate signs of technological maturity or a transition towards new research directions (Tsvetanova et al., 2024; Yadav et al., 2022).

3.4. RQ3: CPC CLASSIFICATION EVOLUTION

CPC classifications were analysed by restructuring the dataset into long format to identify the first and last years of publication for each subclass. This process involved mapping patents to their respective CPC codes and filtering the data from 1950 onwards to focus on periods with higher registration volumes. Historical time series and centred moving averages were then applied to compare, in chronological order, the publication intensity of each subclass. See Figure 5.

Subclasses such as C07C and D21C appear in earlier periods, whereas others, like Y02E and Y02P, have emerged more recently, displaying thicker bars that indicate higher annual publication averages. Core chemical classes (C07, C08, C12) and those associated with physico-chemical processes (B01) have maintained a long-standing presence but gained renewed momentum after the 2000s—coinciding with expanded research in biotechnology and sustainability. This expansion is consistent with the temporal patterns observed in priority, filing, publication and grant dates (Figure 4), suggesting that environmental mitigation technologies (Y02) and biomass utilisation have surged over the past decade, driving a marked increase in patenting activity.

Between 2006 and 2024, nearly all subclasses show a rapid increase in activity up to around 2015. Classes such as Y02E (clean energy), C07G (organic compound transformation) and C12 (biotechnological processes) lead in total publication volume. The stabilisation or slight decline observed after 2020 may indicate technological maturity or delays in indexing newly granted patents. In summary, this analysis reveals a historical growth trajectory shaped by economic, environmental and policy pressures that have stimulated innovation in biorefinery technologies.

These findings align with previous studies emphasising the importance of innovation in chemical processes (C07, C08, B01) and the growing centrality of environmental mitigation technologies (Y02) (Moscoviz et al., 2018; Tsvetanova et al., 2024). Furthermore, the late but intense emergence of Y02E and Y02P confirms a shift towards clean energy and efficient biomass utilisation (Shen et al., 2019; Yadav et al., 2022). The recent stabilisation may indicate maturation in specific technological trajectories and greater selectivity in patent filings (Barbosa et al., 2024). In essence, CPC classifications highlight the multifaceted role of biorefineries in bridging established chemical processes with emerging bioeconomy solutions (Mühl et al., 2024).

3.5. RQ4: TECHNOLOGICAL TRENDS RELATED TO BIOREFINERIES

To identify technological trends within each CPC subclass, patent titles were analysed. The extracted terms were aggregated and counted by publication period and subclass, enabling the identification of the most frequent terms and the distinction between those commonly shared across subclasses and those specific to individual categories. Figure 6 summarises these findings.

In the first period (1900–1950), recurring keywords included hydrocarbon, acid, manufacture, aromatic and compound, reflecting a focus on fundamental chemical processes—particularly in the production of basic organic compounds and early-stage petroleum refining. During the same period, terms such as alkyl, ester and phenol (C07C), as well as oil, treatment and refine (C10G), also appeared, consistent with the expansion of the petrochemical industry at the time (Weissermel & Arpe, 2003).

In the next cycle (1951–1970), common terms shifted to acid, hydrocarbon, ester, prepare and catalytic, reflecting more advanced synthesis and refining routes—including the use of heterogeneous catalysts—consistent with the advent of zeolites and other high-efficiency solid materials (Thomas, 2003). Subclasses such as B01J and C10G support this evolution, featuring terms like aluminosilicate and crystalline (for catalysis), as well as heavy, crack and desulfurisation (for refining).

From 1971 to 1998, keywords continued to include acid and ester, now accompanied by prepare, patent and compound, indicating sustained interest in organic synthesis alongside growing attention to intellectual property. In subclass C07C, terms such as derivative, substitute and carboxylic emerged. This period also reveals an increasing focus on polymers and biomass-derived materials (C08B, C08H, C08L), with terms such as cyclodextrin, biodegradable and polysaccharide—echoing early developments in green chemistry (Clark & Deswarte, 2015).

The most notable transition occurred in the period from 1999 to 2019, during which common and transversal terms increasingly referred to biomass, lignin, acid, material and lignocellulosic. This clearly reflects heightened interest in lignocellulosic biomass valorisation, whether for the production of advanced fuels (C10G, C10L) or biodegradable materials (C08L). The use of enzymes and biotechnological pathways also gained prominence (C12N, C12P), consistent with the large-scale adoption of biorefinery systems and enzymatic processes (Sheldon, 2017).

In the most recent cycle (2020–2025), the most frequent terms—lignin, biomass, acid, material and composition—confirm the consolidation of biomass-based strategies. Emphasis was placed on technologies for lignin and cellulose conversion, as well as on the development of new, lower-impact composites and formulations (C10L, C08L). Meanwhile, subclasses such as B01D and B01J reflect innovations in purification, catalysis and desalination, while C12N and C12P highlight the growing interface with genetic engineering, including terms such as virus, pluripotent and genome.

In sum, the data depict a technological trajectory beginning with basic chemistry—focused on acids, hydrocarbons and aromatic compounds—progressing through increasingly sophisticated catalytic and synthetic processes, and culminating in the widespread adoption of lignocellulosic biomass and renewable materials. These findings support trends identified by Clark & Deswarte (2015), who emphasise the transition towards a circular economy grounded in biorefineries and high-value products derived from renewable sources. They are also consistent with resource- and waste-cycling strategies aligned with broader bioeconomy principles (Mühl et al., 2024).

4. CONCLUSIONS

This study consolidated and analysed a comprehensive dataset of patents related to biorefineries, highlighting the temporal evolution of publications and the concentration patterns across specific countries, inventors and assignees, as well as the CPC subclasses that have exhibited the most significant growth in recent years. A clear trend of convergence is evident between traditional chemistry and emerging sustainability imperatives, reflected in the dominant themes observed during the most recent cycles (1999–2019 and 2020–2025). Terms such as biomass, lignin, lignocellulosic, acid, material and composition have become increasingly prominent, indicating growing interest in biomass valorisation and the pursuit of cleaner technological pathways.

In particular, CPC subclasses related to physico-chemical processes (B01, C07) and biotechnology (C12, Y02) account for a substantial share of recent filings, suggesting a synergy between advanced catalysis, the development of novel materials and the expanding adoption of biorefinery systems. Temporal analyses reveal a marked increase in both the volume and complexity of patent activity from the early 2000s onwards, coinciding with the strengthening of policy incentives for renewable energy and intensified research into alternative resource strategies.

Although a small number of countries continue to dominate overall patent-filing volumes, the broad dispersion of inventors and assignees reinforces the notion of a pluralistic and dynamic innovation ecosystem, in which diverse actors contribute value across various stages of the biorefinery chain. However, it is important to acknowledge that bibliometric and patent-mapping analyses are limited by the availability and completeness of data in the consulted databases, and are therefore subject to temporal gaps or delays in record inclusion. Furthermore, the multidisciplinary nature of many patents may lead to classification biases within the CPC system, as documents often span overlapping domains such as chemistry, biotechnology and process engineering.

Future research could include citation-network analyses to better capture knowledge diffusion and the formation of technological-collaboration structures. Expanding the scope to encompass filings submitted to smaller or regional patent offices would help assess whether meaningful innovations are under-represented in international databases. In parallel, qualitative case studies could provide deeper insights into the effectiveness of technology-transfer processes and the socio-environmental impacts of specific biorefinery pathways.

The integrated analysis of large-scale patent data and the identification of key technological and geographical trends provide critical input for strategic decision-making in research and development, public policy and responsible innovation. The updated findings reaffirm the potential of biorefineries to support the construction of sustainable and regenerative value chains, integrating chemical and biotechnological innovations aligned with circular-economy principles. Ultimately, strengthening biorefinery platforms and investing in biomass- and lignin-based technological routes may contribute not only to mitigating environmental impacts and closing material loops, but also to enhancing global energy security—benefiting both the scientific community and society as a whole.

5. REFERENCES

Amidon, T. E., Bujanovic, B., Liu, S., & Howard, J. R. (2011). Commercializing biorefinery technology: A case for the multi-product pathway to a viable biorefinery. **Forests**, 2(4), 929–947. https://doi.org/10.3390/f2040929

Barbosa, A. V. L., Paredes, M. L. L., Alijo, P. H. R., Sardou, A. C. O., Maia, J. G. S. S., & Bastos, J. B. V. (2024). Continuous design and technoeconomic assessment of commercial-scale biorefinery processes for the production of succinic acid. **Biofuels, Bioproducts and Biorefining**, 18(5), 1289–1305. https://doi.org/10.1002/bbb.2624

Barragan-Ocana, A., Merritt, H., Sanchez-Estrada, O. E., Mendez-Becerril, J. L., & del Pilar Longar-Blanco, M. (2023). Biorefinery and sustainability for the production of biofuels and value-added products: A trends analysis based on network and patent analysis. **PLOS ONE**, 18(1), e0279659. https://doi.org/10.1371/journal.pone.0279659

Breschi, S., & Malerba, F. (2023). **Clusters, networks, and innovation**. Oxford University Press.

Chatfield, C. (2004). **The analysis of time series**: An introduction (6th ed.). Chapman & Hall/CRC.

Clark, J., & Deswarte, F. (2015). The biorefinery concept. In **Introduction to chemicals from biomass** (pp. 1–29). John Wiley & Sons. https://doi.org/10.1002/9781118714478.ch1

Dimian, A. C., Iancu, P., Plesu, V., Bonet-Ruiz, A.-E., & Bonet-Ruiz, J. (2019). Castor oil biorefinery: Conceptual process design, simulation and economic analysis. **Chemical Engineering Research and Design**, 141, 198–219. https://doi.org/10.1016/j.cherd.2018.10.040

Faba, L., Diaz, E., & Ordonez, S. (2015). Recent developments on the catalytic technologies for the transformation of biomass into biofuels: A patent survey. **Renewable and Sustainable Energy Reviews**, 51, 273–287. https://doi.org/10.1016/j.rser.2015.06.020

Feldman, R., & Sanger, J. (2006). **The text mining handbook**: Advanced approaches in analysing unstructured data. Cambridge University Press. https://doi.org/10.1017/CBO9780511546914

Field, A. P., Miles, J., & Field, Z. (2012). **Discovering statistics using R. Sage**.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. **Journal of Economic Literature**, 28(4), 1661–1707.

Hornik, K., Mair, P., Rauch, J., Geiger, W., Buchta, C., & Feinerer, I. (2013). The textcat package for n-gram based text categorisation in R. **Journal of Statistical Software**, 52, 1–17. https://doi.org/10.18637/jss.v052.i06

Khemthong, P., Yimsukanan, C., Narkkun, T., Srifa, A., Witoon, T., Pongchaiphol, S., Kiatphuengporn, S., & Faungnawakij, K. (2021). Advances in catalytic production of value-added biochemicals and biofuels via furfural platform derived lignocellulosic biomass. **Biomass and Bioenergy**, 148, 106033. https://doi.org/10.1016/j.biombioe.2021.106033

Khoshnevisan, B., Duan, N., Tsapekos, P., Awasthi, M. K., Liu, Z., Mohammadi, A., Angelidaki, I., Tsang, D. C. W., Zhang, Z., Pan, J., Ma, L., Aghbashlo, M., Tabatabaei, M., & Liu, H. (2021). A critical review on livestock manure biorefinery technologies: Sustainability, challenges, and future perspectives. **Renewable and Sustainable Energy Reviews**, 135, 110033. https://doi.org/10.1016/j.rser.2020.110033

Lotka, A. J. (1926). The frequency distribution of scientific productivity. **Journal of the Washington Academy of Sciences**, 16(12), 317–323.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). **Introduction to information retrieval**. Cambridge University Press.

Moscoviz, R., Trably, E., Bernet, N., & Carrere, H. (2018). The environmental biorefinery: State-of-the-art on the production of hydrogen and value-added biomolecules in mixed-culture fermentation. **Green Chemistry**, 20(14), 3159–3179. https://doi.org/10.1039/c8gc00572a

Mühl, D. D., da Silva, M. V. B., & de Oliveira, L. (2024). How to build a bioeconomic food system: A thematic review. **Circular Economy and Sustainability**. https://doi.org/10.1007/s43615-024-00387-1

Ooms, J. (2017). **writexl**: Export data frames to Excel "xlsx" format (Version 1.5.1) [R package]. https://doi.org/10.32614/CRAN.package.writexl

Pessoa, L. C., Deamici, K. M., Pontes, L. A. M., Druzian, J. I., & Assis, D. de J. (2021). Technological prospection of microalgae-based biorefinery approach for effluent treatment. **Algal Research**, 60, 102504. https://doi.org/10.1016/j.algal.2021.10250

Potter, W. (1981). **Lotka's law revisited**. Library Trends. https://www.semanticscholar.org/paper/Lotka's-Law-Revisited-Potter/e7703e0d6b7b0a9bd14d49f11daf949e584590b5

Sheldon, R. A. (2017). The E factor 25 years on: The rise of green chemistry and sustainability. **Green Chemistry**, 19(1), 18–43. https://doi.org/10.1039/C6GC02157C

Shen, R., Tao, L., & Yang, B. (2019). Techno-economic analysis of jet-fuel production from biorefinery waste lignin. **Biofuels, Bioproducts and Biorefining**, 13(3), 486–501. https://doi.org/10.1002/bbb.1952

Silva, S. C., Ferreira, I. C. F. R., Dias, M. M., & Barreiro, M. F. (2020). Microalgae-derived pigments: A 10-year bibliometric review and industry and market trend analysis. **Molecules**, 25(15), 3406. https://doi.org/10.3390/molecules25153406

Tauseef, S. M., Tabassum-Abbasi, Patnaik, P., Abbasi, T., & Abbasi, S. A. (2023). A novel high-rate vermicomposting machine for downstream processing of biorefinery waste and other forms of organic solid waste. **Biomass Conversion and Biorefinery**, 13(5), 3683–3692. https://doi.org/10.1007/s13399-021-01522-w.

Thomas, V. M. (2003). Product self-management: Evolution in recycling and reuse. **Environmental Science & Technology**, 37(23), 5297–5302. https://doi.org/10.1021/es0345120

Tsvetanova, L., Rennings, M., & Broering, S. (2024). Actors and sectors in technological innovation systems: Patterns of knowledge development in the field of second generation biorefineries. **Technology Analysis & Strategic Management**. https://doi.org/10.1080/09537325.2024.2320239

Weissermel, K., & Arpe, H.-J. (2003). **Industrial organic chemistry** (4th completely rev. ed.). Wiley-VCH. https://doi.org/10.1002/9783527619191

Wickham, H. (with Sievert, C.). (2016). **ggplot2**: Elegant graphics for data analysis (2nd ed.). Springer International Publishing.

Wilkinson, L., & Friendly, M. (2009). The history of the cluster heat map. **The American Statistician**, 63(2), 179–184. https://doi.org/10.1198/tas.2009.0033

World Intellectual Property Organization (WIPO). (2025, January 31). **IPC/CPC definitions**. https://ipcpub.wipo.int/?notion=scheme&version=20250101&symbol=none&menulang=en&lang=en

Yadav, K., Vasistha, S., Nawkarkar, P., Kumar, S., & Rai, M. P. (2022). Algal biorefinery culminating multiple value-added products: Recent advances, emerging trends, opportunities, and challenges. 3 **Biotech**, 12(10), 288. https://doi.org/10.1007/s13205-022-03288-y