

BRAZILIAN BIOECONOMY: Overview of national scientific production



In this Issue:

Snowball Methodology

Overview of publications on
the bioeconomy

Characterization of the
network: thematic clusters

Cooperation with other
countries

Institutions that publish the
most in bioeconomy

Final Considerations

This is the second edition of the Bioeconomy Thematic Report. It aims to analyze the panorama of scientific production of the Brazilian bioeconomy. The first edition of the Report brought the panorama of publications worldwide and included a Brazilian analysis, albeit a superficial one. This edition aims to deepen the investigation of the Brazilian bioeconomy through a survey of the national scientific production. This Report considers the following definition of bioeconomy:

“The bioeconomy comprises all economic activity derived from bioprocesses and bioproducts that contribute to efficient solutions in the use of biological resources - facing challenges in food, chemicals, materials, energy production, health, environmental services and environmental protection - that promote the transition to a new model of sustainable development and societal well-being”

(ODBio, 2020)

In this edition, a methodology was developed to access a more robust and representative sample of national publications.

The development of a specific methodology was necessary due to the fact that the search in scientific databases for

the term **bioeconomy** does not capture the entire field of activity, as it is a **new** and **transversal** topic.

Snowball Methodology

The methodology used to create the network of scientific publications of the Brazilian bioeconomy was based on the snowball method. The snowball is used to generate search expressions that allow the identification of scientific productions related to a new and comprehensive expression. To generate the new search expressions, the keyword pairs that appeared in greater

co-occurrence in the initial network were used, and to filter out unrelated papers, the semantic similarity between the texts¹ was used. The process occurred as follows: initially, the network of Brazilian papers from the base of the Web of Science (WoS) that contained the term “bioeconomy” (seed network) was generated, then the pairs of keywords that appeared in a higher level of

co-occurrence were identified, and these pairs were used for new searches in the WoS base. These new papers were inserted in the initial network, and we checked which papers were connected by semantic similarity with the seed network. Non-connected papers were excluded. This process was repeated four times until it was verified that no new keyword pairs appeared².

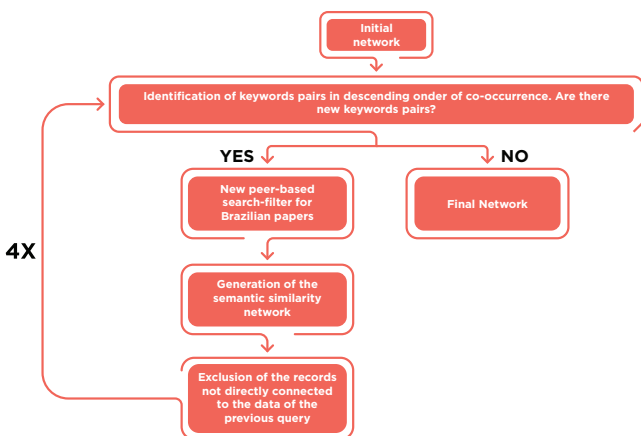


Figure 1: Snowball Methodology.

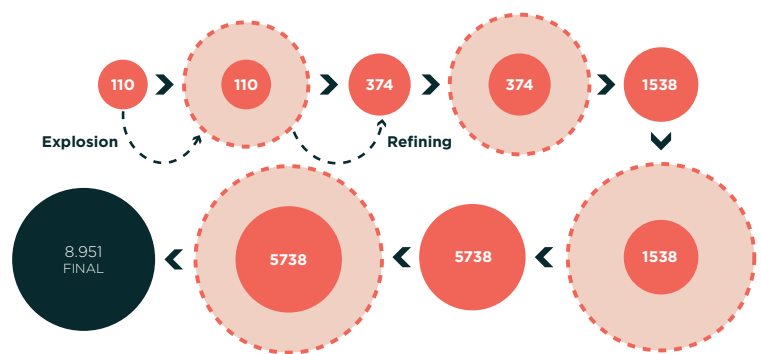


Figure 2: Evolution of the number of papers.

1 To measure the level of semantic similarity, the texts of the title, abstract, and keywords were used.

2 To define the pairs of keywords that would be selected, we used a criterion of which pairs had a co-occurrence value of at least half of the value of the pair that appeared in higher co-occurrence.

Figure 1 describes this process. In Figure 2, it is possible to observe in quantitative terms how the

snowball evolution occurred. The cycles oscillate between the “expansions” of papers and the

“refining” through the semantic similarity analysis with the previous network.

1. Overview

The network generated from the snowball methodology allowed us to survey 8,951 Brazilian papers from 1976 to 2021 and separate them into topic clusters (Figure 3). Figure 4 presents the evolution of the number of scientific publications on bioeconomy throughout the years as of 2006, where the growing interest in this topic was observed. The smaller number of publications in the year 2021 is due to the fact that the survey takes into account the publications registered in the WoS until August 2021.

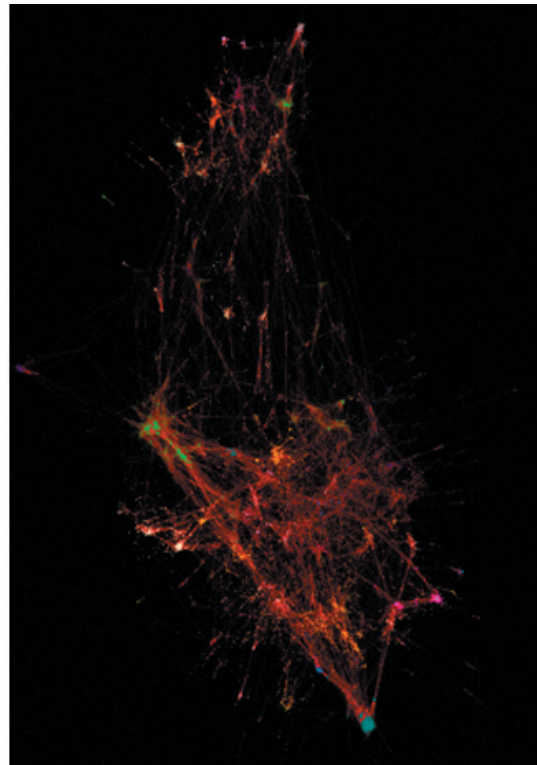


Figure 3: Bibliometric network of Brazilian papers on bioeconomy.

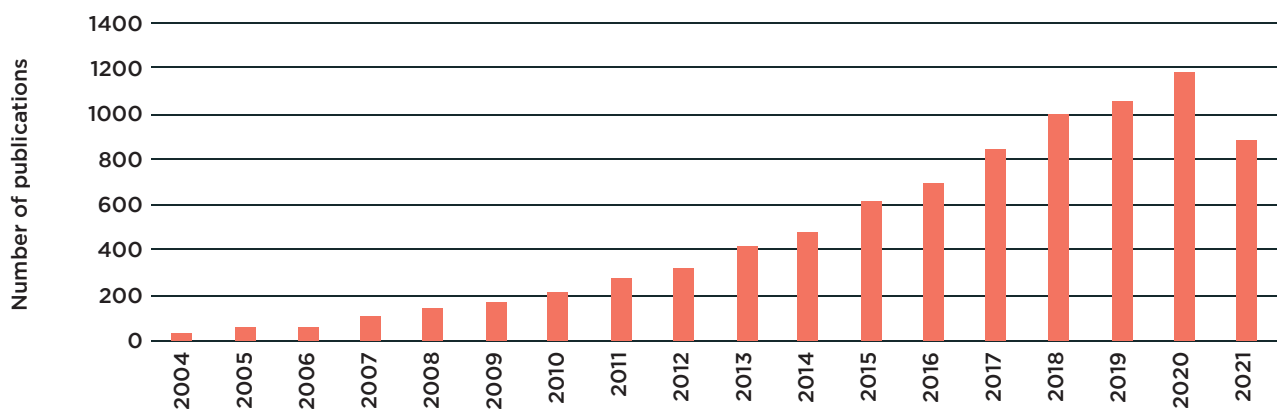


Figure 4: Number of scientific publications on the bioeconomy over the years.

A first approach to the Brazilian bioeconomy paper network can be to analyze the format of the semantic similarity network. The network is built in such a manner as to group similar topics. As a consequence, its layout may point out some possible topic areas. In Figure 6, we may distinguish two

macro-regions, an upper and a lower one, connected by a set of smaller clusters. An investigation into these macro-regions indicated that most of the network - the lower macro-region - deals with more technical issues related to biomass production and transformation processes into bioeconomy products.

Specific terms of Brazilian bioeconomy processes can be seen among the most cited keywords, such as **fermentation, pre-treatment, enzymatic hydrolysis, and pyrolysis**. Also, references to raw materials, such as **lignocellulosic biomass and microalgae**, and products, such as **ethanol and biodiesel**.

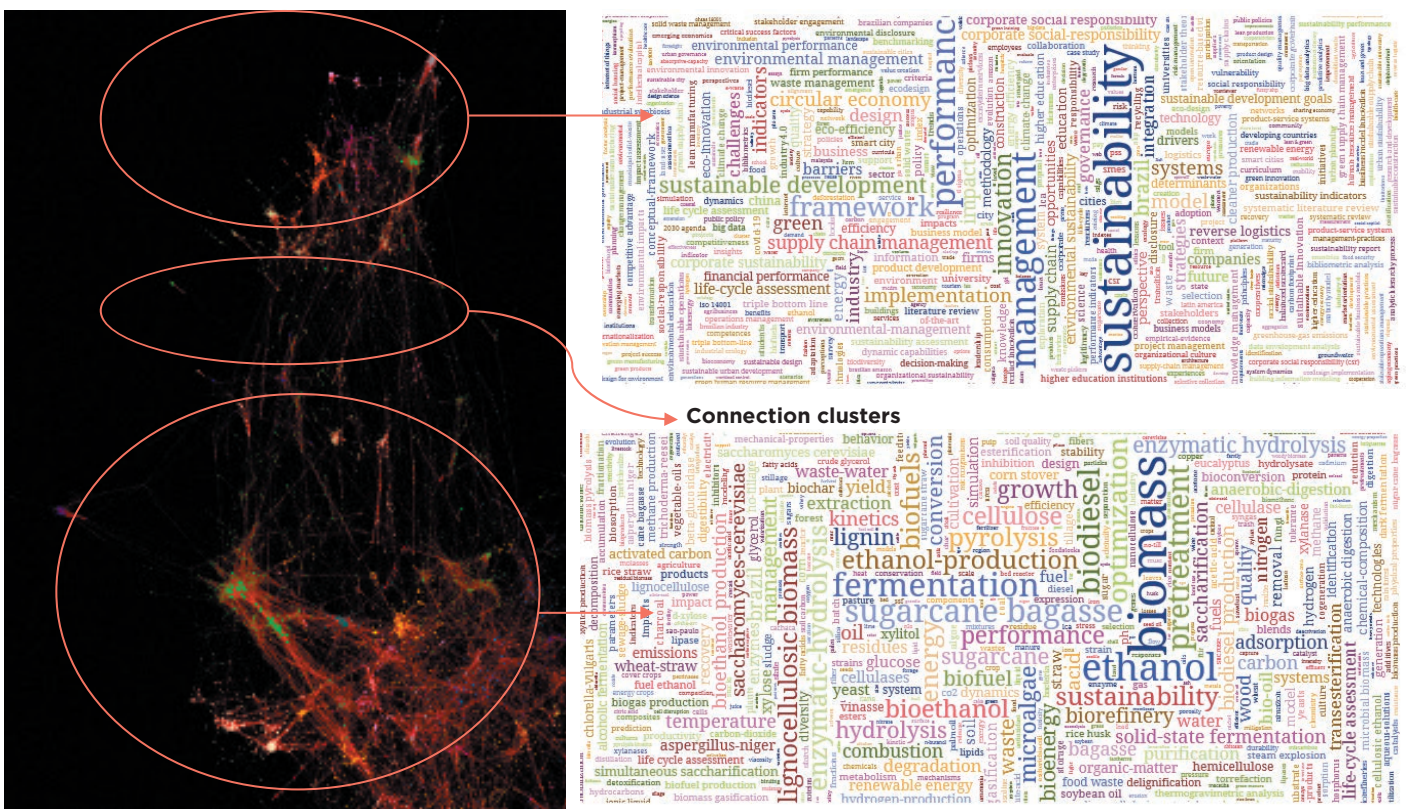


Figure 6: Macro-regions of the Brazilian scientific publications network in bioeconomy

The upper macro-region presents subjects focused on management and innovation. Amongst the most frequent keywords, we find:

1. Sustainability,

2. Management, 3. Performance, 4. Framework, 5. Innovation. Other words that also emphasize a different profile of the upper macro-region are **indicators, circular**

economy, governance, industry, and strategy. Aiming to explore the network topics further, the next section will present a discussion on 15 thematic clusters.

2. Characterization of the network: thematic clusters

To characterize the scientific publication network of the Brazilian bioeconomy, the 15 largest thematic clusters were

selected. For each cluster will be presented: description, word cloud, and graph of the number of publications. Figure

7 presents the analyzed clusters.

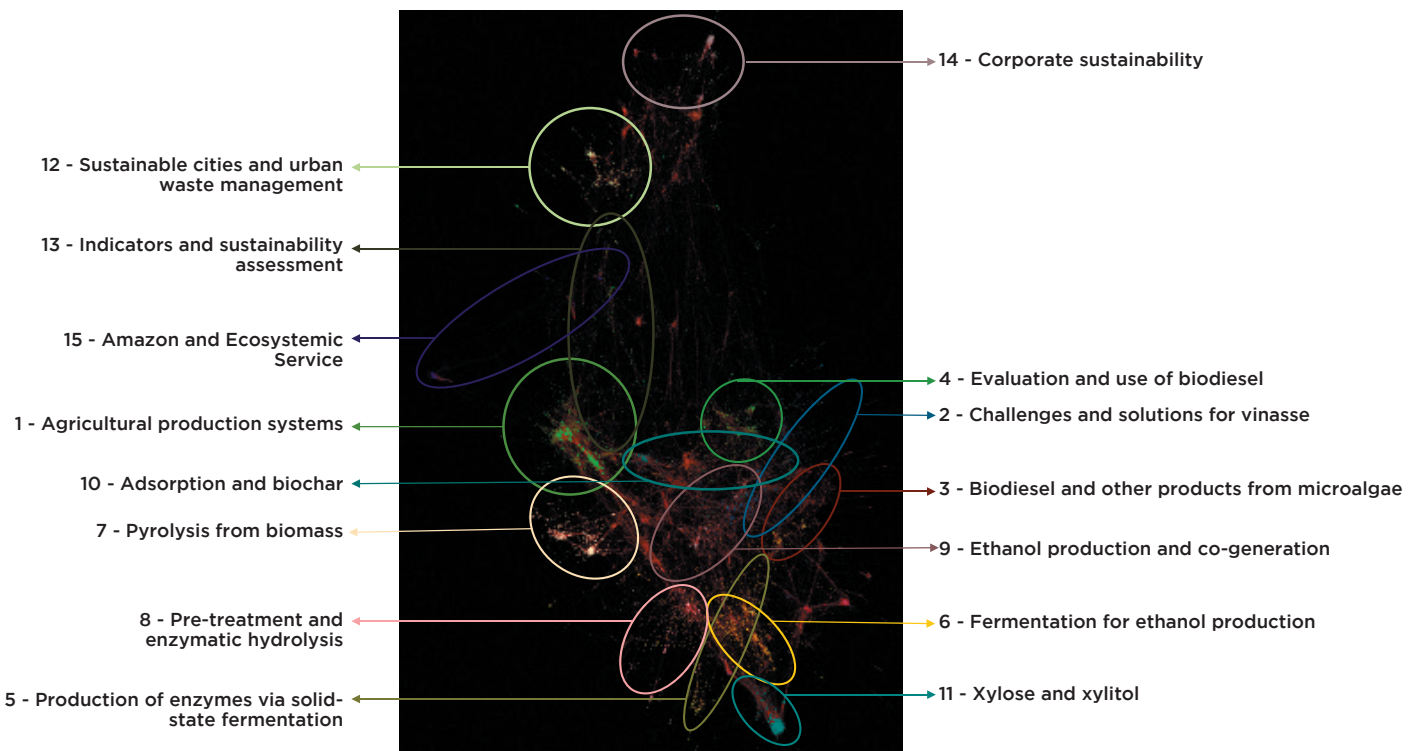


Figure 7: 15 thematic clusters analyzed.

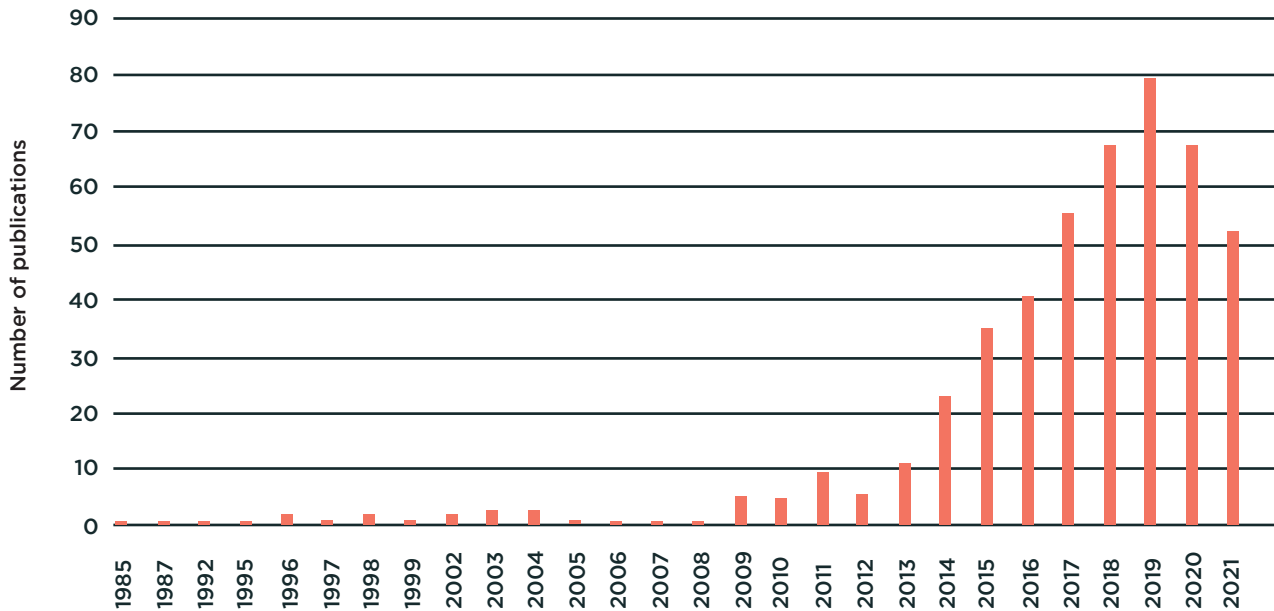


Figure 11: Number of publications in cluster 2 over time.

Despite the dispersion of the cluster, it is possible to divide it into two main parts: the innermost section of the network and the peripheral section to the right. The innermost section of cluster 2 seems to have been formed around the term **vinasse**³, with the papers discussing proposed solutions for the vinasse from the production of ethanol, such as **microalgae**

cultivation, biogas production, lipid production, and concentration by evaporation. A few papers present simulations and optimization alternatives for the ethanol production process that includes the use or treatment of vinasse. Another group covers assessments of the environmental consequences of vinasse disposal. The peripheral part of the cluster seems to

have been more influenced by terms associated with microalgae. In this part of the cluster, there are several papers on the production of biodiesel from microalgae. Although cluster 2 is younger than cluster 1 - gaining relevance from 2013 - the number of publications grew rapidly, peaking in 2019.

³ It was the second most cited word - not keyword - in cluster 2. The first was the more generic term "production".

2.3. Cluster 3 - Biodiesel and other products from microalgae

Cluster 3 is found in the lower macro-region. It contains 466 papers and is considerably

spread out. Figures 12 and 13 present the word cloud and the number of publications

of the cluster over time, respectively.

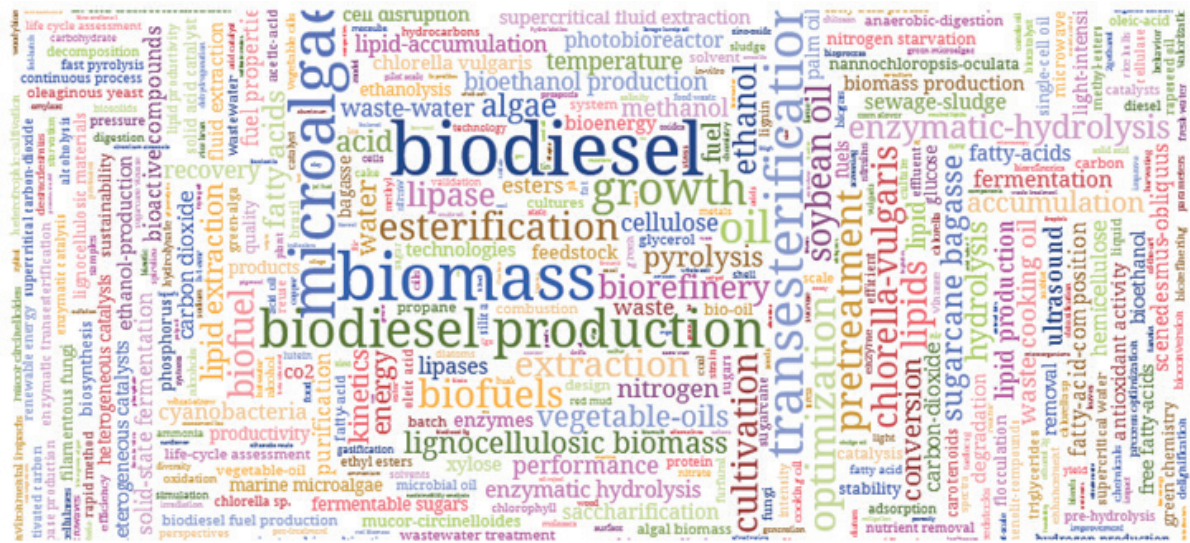


Figure 12: Word cloud for cluster 3.

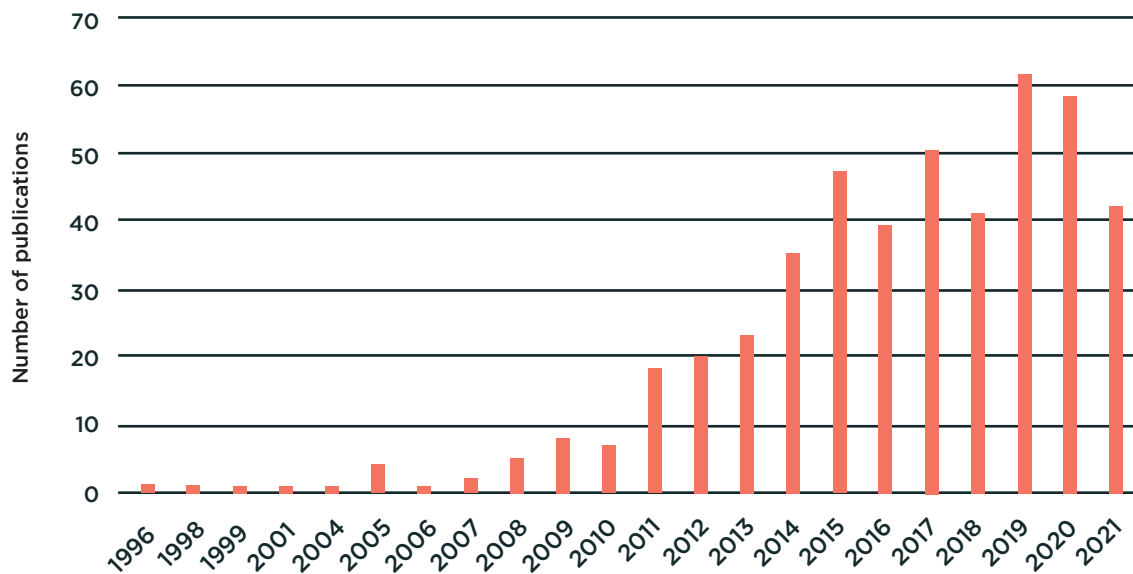


Figure 13: Number of cluster 3 publications over time.

Cluster 3 is quite close to cluster 2, mainly at the periphery of the network where the clusters overlap. This explains why cluster 3 also contains papers about the production of biodiesel from microalgae. However, unlike cluster 2, which seems

to have the vinasse as the focal point of the cluster, in this case, the focal point seems to be the production of biodiesel from microalgae. Among the most frequent topics in the cluster are studies on the potential of microalgae lipids; microalgae

production via industrial effluents; and biodiesel production. Cluster 3 has gained relevance since 2011, and despite the overall growth in the number of publications, has undergone oscillations since 2014.

2.4. Cluster 4 - Biodiesel assessment and use

Cluster 4 is found in the lower macro-region, very close to clusters 2 and 3. It contains 390 papers and is concentrated in a peripheral region of the network. Figures 14 and 15 show the word cloud and the number of publications of the cluster over time, respectively.

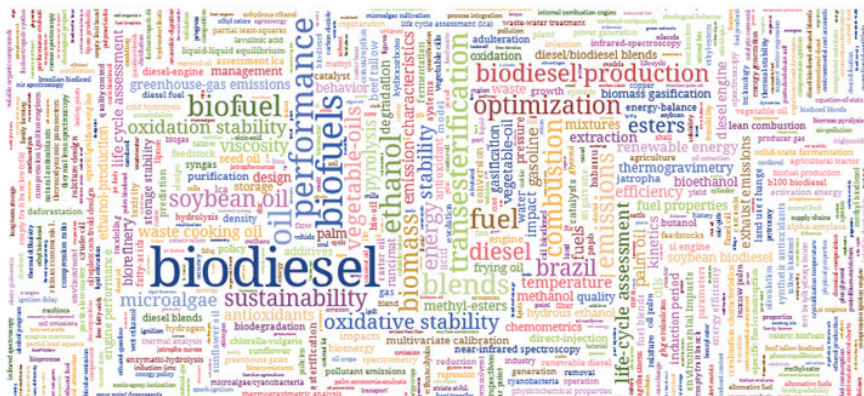


Figure 14: Word cloud of cluster 4.

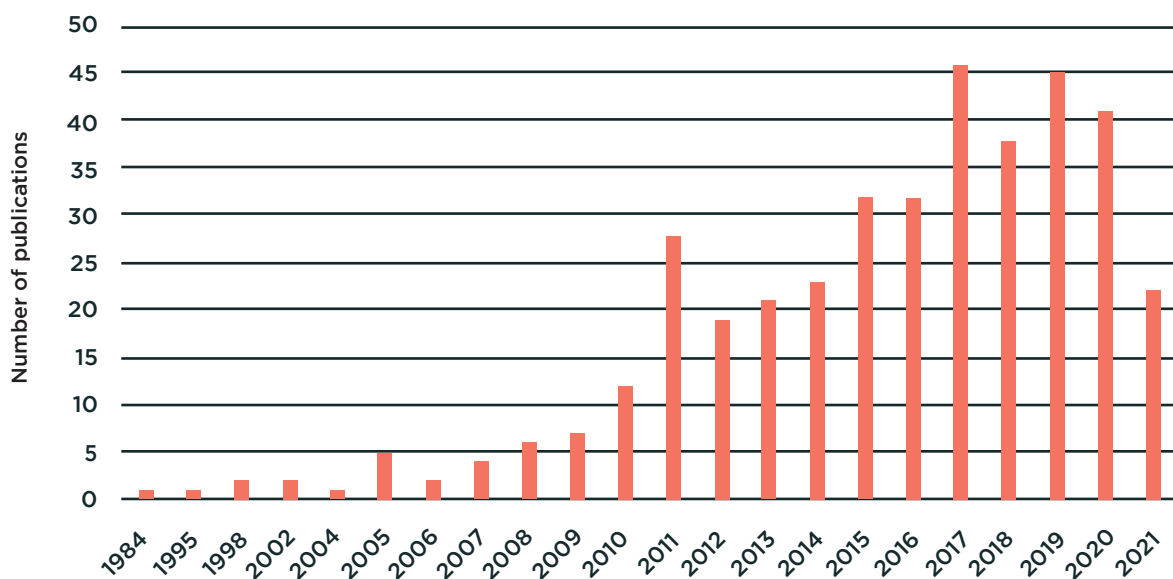


Figure 15: Number of publications in cluster 4 over time.

BRAZILIAN BIOECONOMY:

Overview of national scientific production

BIOECONOMY OBSERVATORY • YEAR 2 • Nº 2 • FEBRUARY - 2022



The part of the cluster in the most central region of the network deals mainly with the production and characterization processes of enzymes produced through the solid-state fermentation process, with a focus on promoting hydrolysis of

biomass. The studies vary depending on the type of enzyme desired, the starting materials (such as different types of fungi), the types of analysis and characterization, and various forms of use. The peripheral section of the cluster shows similar

topics but focuses on the *Trichoderma reesei* fungus, as the central part focuses on *Aspergillus niger*. Cluster 5 saw a constant increase in publications between the years 2009 and 2016, starting to oscillate from 2017.

2.6. Cluster 6 - Fermentation for ethanol production

Cluster 6 is located in the lower macro-region near cluster 5 and

has 376 papers. Figures 18 and 19 present the word cloud and the

number of publications of the cluster over time, respectively.

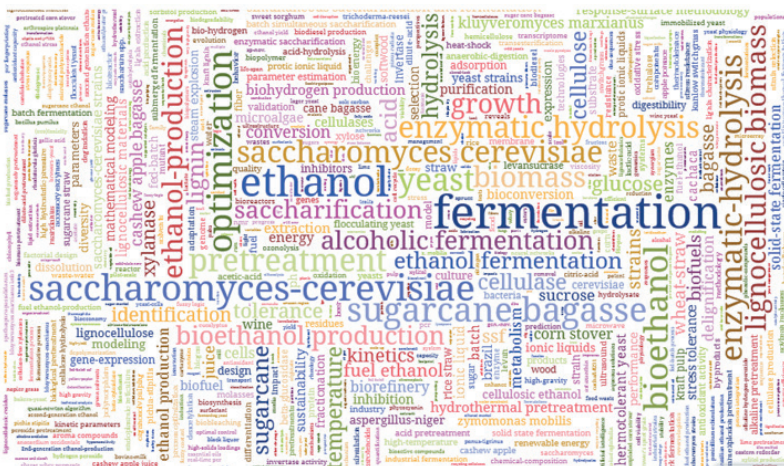


Figure 18: Word cloud for cluster 6.

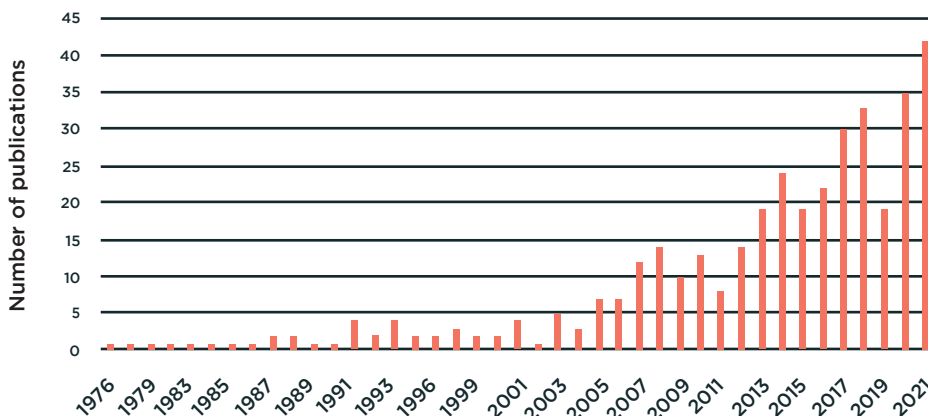


Figure 19: Number of publications in cluster 6 over time.

Cluster 6 contains topics related to the fermentation process for the production of ethanol. The papers discuss topics such as the use of different types of yeast; analysis of configurations of the fermentation process; yeast

characterization processes; and analysis of saccharification and fermentation of different residual biomass. Cluster 6 is one of the oldest clusters of the network, having records prior to the 1990s. Even with oscillations, the number of

publications has been growing since 2004, which saw the implementation of flex-fuel engines. It is remarkable that the number of publications in 2021 (until August) already exceeds that of previous years.

2.7. Cluster 7 - Pyrolysis from biomass

Cluster 7 contains 314 items and is located on the left periphery of the lower macro-region. The cluster is concentrated and has two main regions of higher concentration. Figures 20 and 21 present the Word cloud and the number of publications of the cluster over time, respectively.

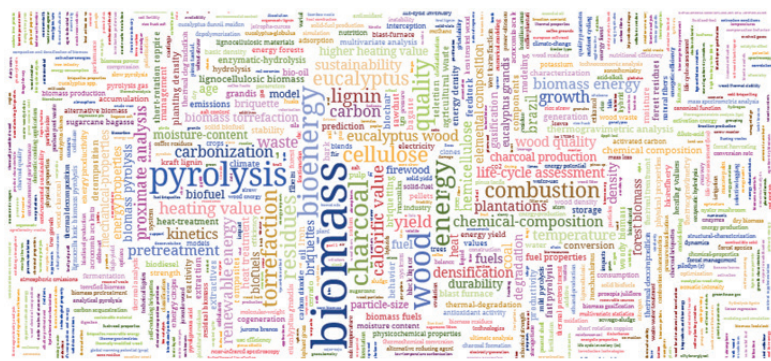


Figure 20: Word cloud for cluster 7.

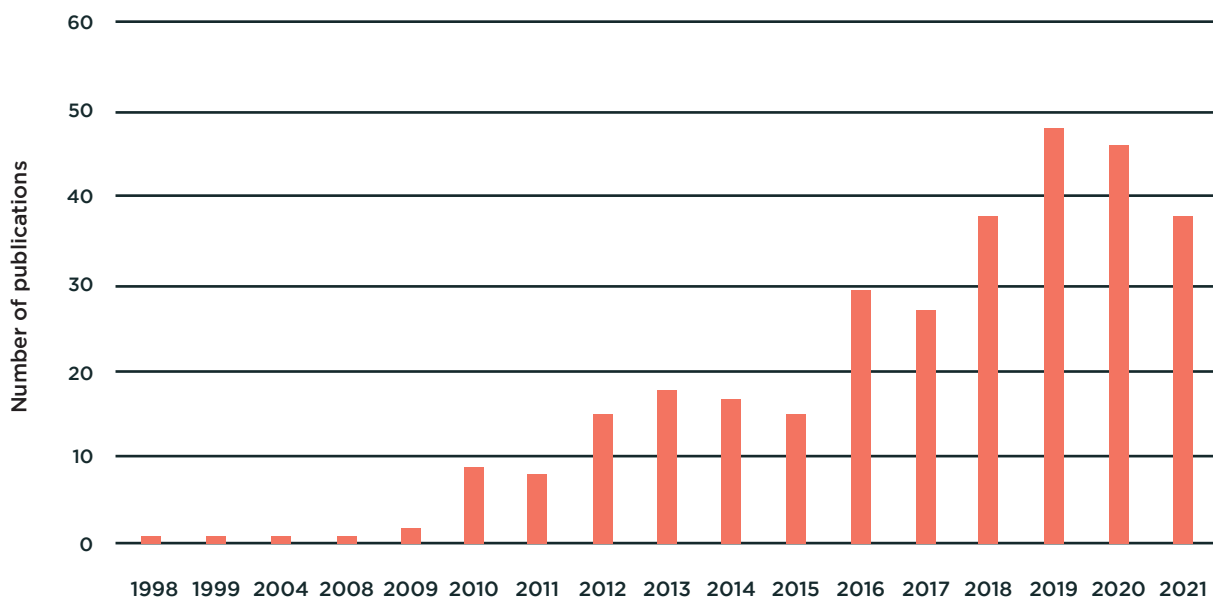


Figure 21: Number of publications in cluster 7 over time.

The topic of cluster 7 revolves around the pyrolysis process using mainly wood for energy purposes, such as charcoal. More specifically, the papers present topics such as the influence of the type of wood in charcoal production, the influence of pyrolysis

conditions on the quality of charcoal generated, the characterization of briquettes, pellets, and coals generated from biomass, and the influence of the mixture of biomass in the efficiency of the process and the products generated. The peripheral

and concentrated position of cluster 7 evidenced the characterization of a more specific topic, more loosely connected with the rest of the network. Cluster 7 is more recent and has evolved in terms of the number of publications since 2010.

2.8. Cluster 8 - Pre-treatment and enzymatic hydrolysis

Cluster 8 is located in the lower macro-region near cluster 6 of fermentation and

contains 275 papers. Figures 22 and 23 show the word cloud and the number of

publications of the cluster over time, respectively.



Figure 22: Word cloud of cluster 8.

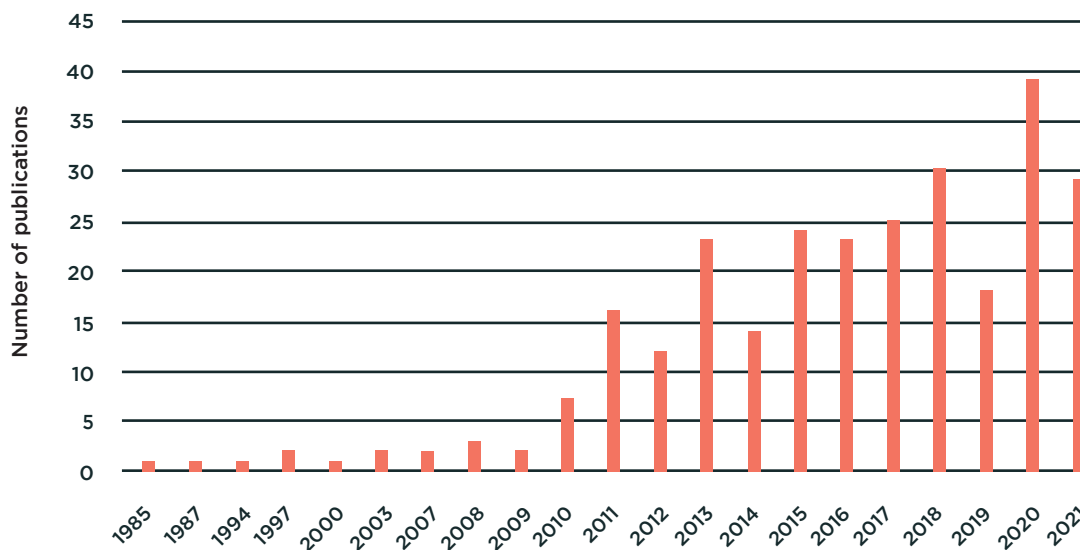


Figure 23: Number of publications in cluster 8 over time.

The topics of cluster 8 revolve around pre-treatment processes and enzymatic hydrolysis, mainly sugarcane residues. In general, the papers focus on the production

of second-generation ethanol. The topics range from the evaluation of different techniques to the characterization of pretreated and hydrolyzed biomass.

Several papers also present proposals for the optimization of these processes. Figure 23 shows an increasing trend in the number of papers on this topic since 2010.

2.9. Cluster 9 - Ethanol production and co-generation

Cluster 9 shows 256 papers and is located in the lower macro-region. The cluster is mainly concentrated in the

central part of the macro-region, spreading to the right-hand periphery. Figures 24 and 25 present the word

cloud and the number of publications of the cluster over time, respectively.

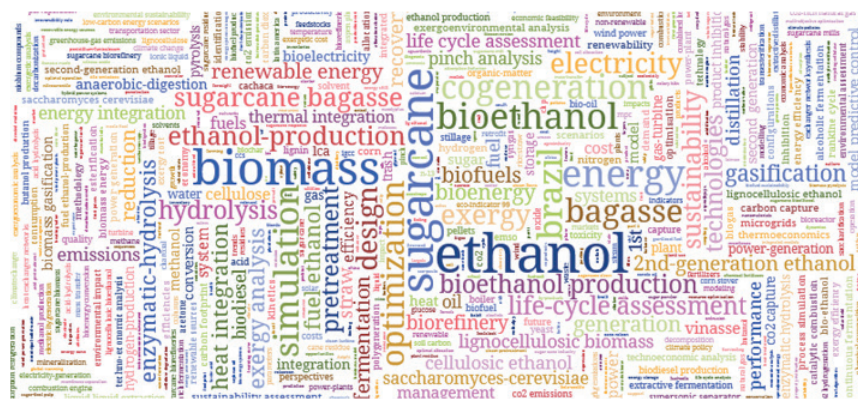


Figure 24: Word cloud for cluster 9.

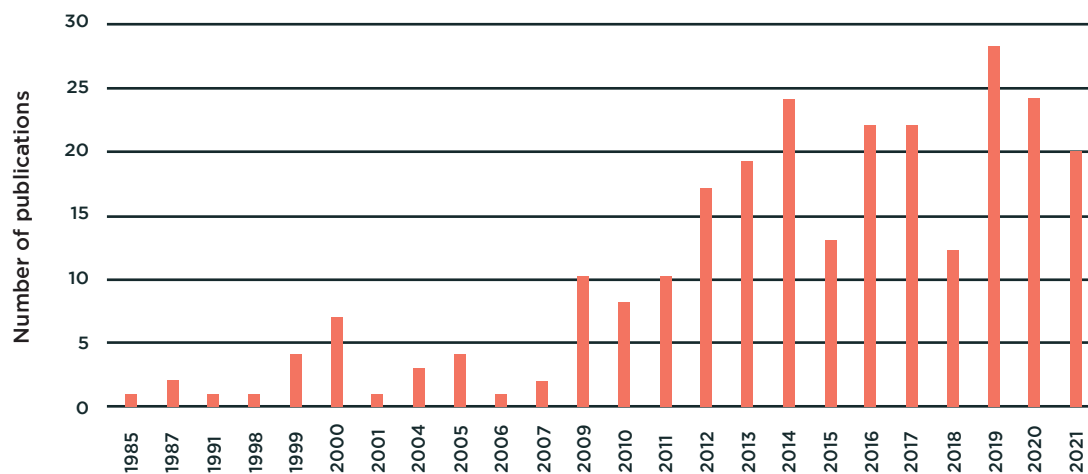


Figure 25: Number of publications in cluster 9 over time.



Cluster 9 is divided into two main areas. The largest area refers to ethanol production issues, mainly: integration alternatives between first and second-generation production, evaluation of different pre-treatment methods, and simulation

of different production configurations. As the papers approach the second area, the topic of cogeneration becomes more evident. More specifically, topics such as optimization of cogeneration processes, comparison between electricity generation

and ethanol production, and biomass alternatives for cogeneration are addressed. Cluster 9 had an initial peak in the number of papers published in 2000, but only began to grow again as of 2010, albeit with oscillations.

2.10. Cluster 10 - Adsorption and biochar

Cluster 10 is located at the upper border of the lower macro-region and is quite spread out. The cluster contains 245 papers. Figures 26 and 27 show the word cloud and the number of publications of the

cluster over time, respectively. For being a cluster quite spread out in the network, the topics vary slightly. In general, we can mention some of the topics addressed by the cluster: the use of biochar and activated

carbon in adsorptive processes; production, evaluation, and characterization of the biochar produced; analysis of greenhouse gas emissions for the production; and use of biochar in various processes.

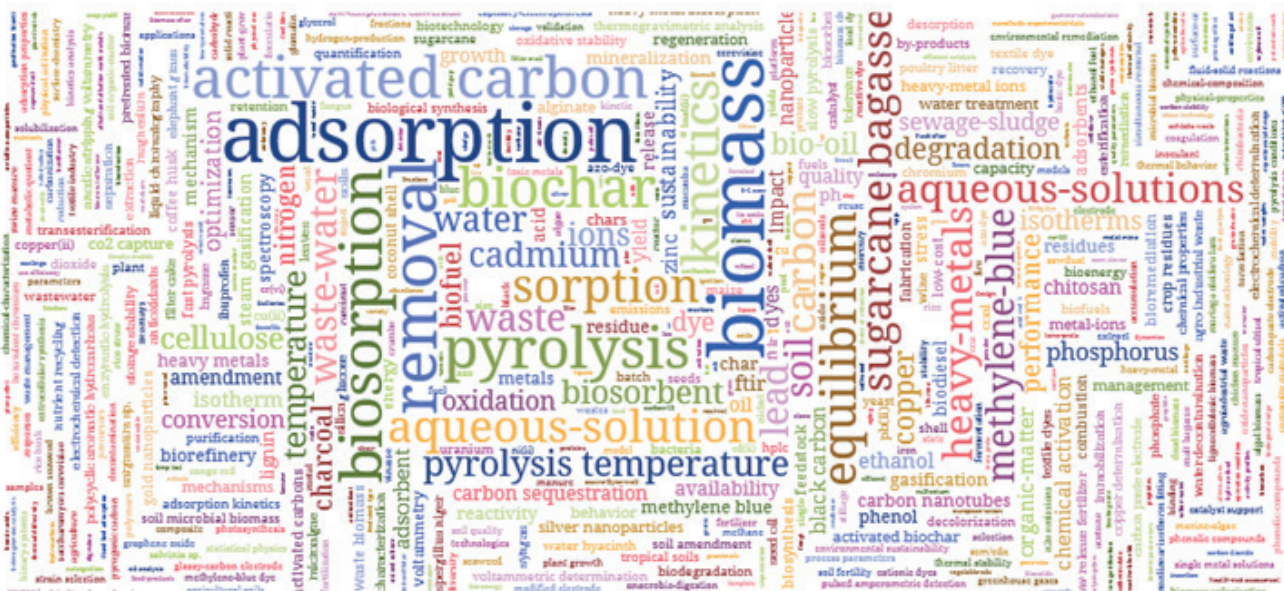


Figure 26: Word cloud for cluster 10.

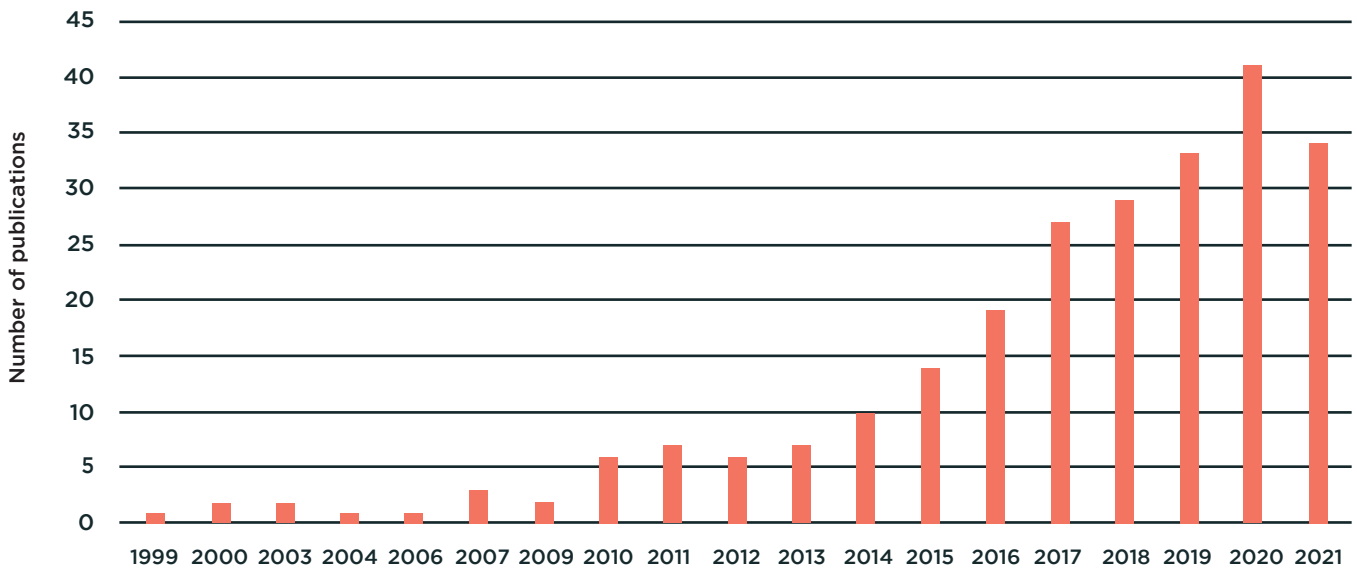


Figure 27: Number of publications in cluster 10 over time.

Waste materials from sugar and ethanol production, such as filter cake and sugarcane bagasse, have often been used as raw material for the production of biochar. Other residual feedstocks are also investigated, such as sewage

sludge. The pyrolysis process is also frequently mentioned for the production of biochar. In the most concentrated part of the cluster, it is possible to identify that the topic focuses on the evaluation and use of biochar as an alternative to

increasing soil fertility. This part of the cluster is close to cluster 1, referring to agricultural production systems. Since the early 2010s, Figure 27 shows a significant and sustained growth in the number of publications on this subject.

2.11. Cluster 11 - Xylose and Xylitol

Cluster 11 is located at the lower end of the network and is extremely concentrated. The cluster contains 244 papers. Figures 28 and 29 show the word cloud and the number of publications in the cluster over

time, respectively. Because it is highly concentrated, the topics in this cluster vary little. The papers mainly deal with xylose production and its conversion into xylitol. This production is often

contextualized in the ethanol industry and aims to use the pre-treatment and hydrolysis processes of biomasses such as sugarcane bagasse and straw for the production of xylose and subsequent



Figure 28: Word cloud for cluster 11.

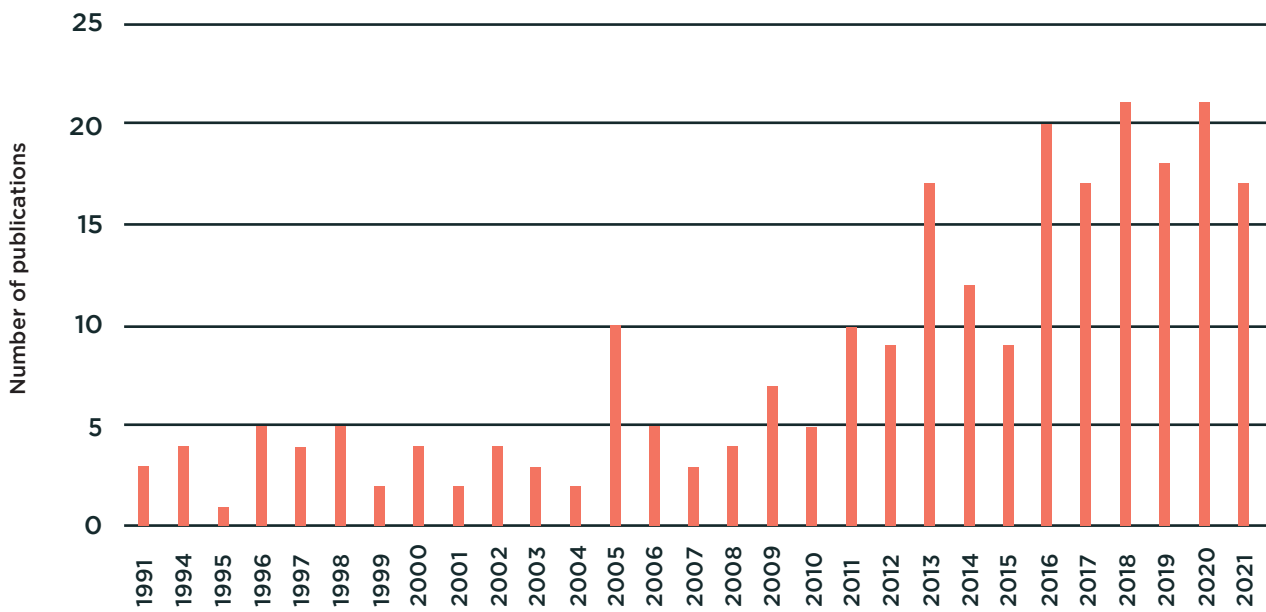


Figure 29: Number of cluster 11 publications over time.

conversion into xylitol as a product of a biorefinery. The studies mainly present techniques and ways to

optimize the process of obtaining xylose. Cluster 11 shows papers since 1991, but began to show growth in the

number of publications in 2007, reaching an apparent stabilization between 2016 and 2020.

2.12. Cluster 12 - Sustainable cities and urban waste management

Cluster 12 is located in the top macro-region of the network and contains 243 relatively concentrated papers. Figures 30 and 31 present the word cloud and the number of publications of the cluster over time, respectively.



Figure 30: Word cloud for cluster 12.

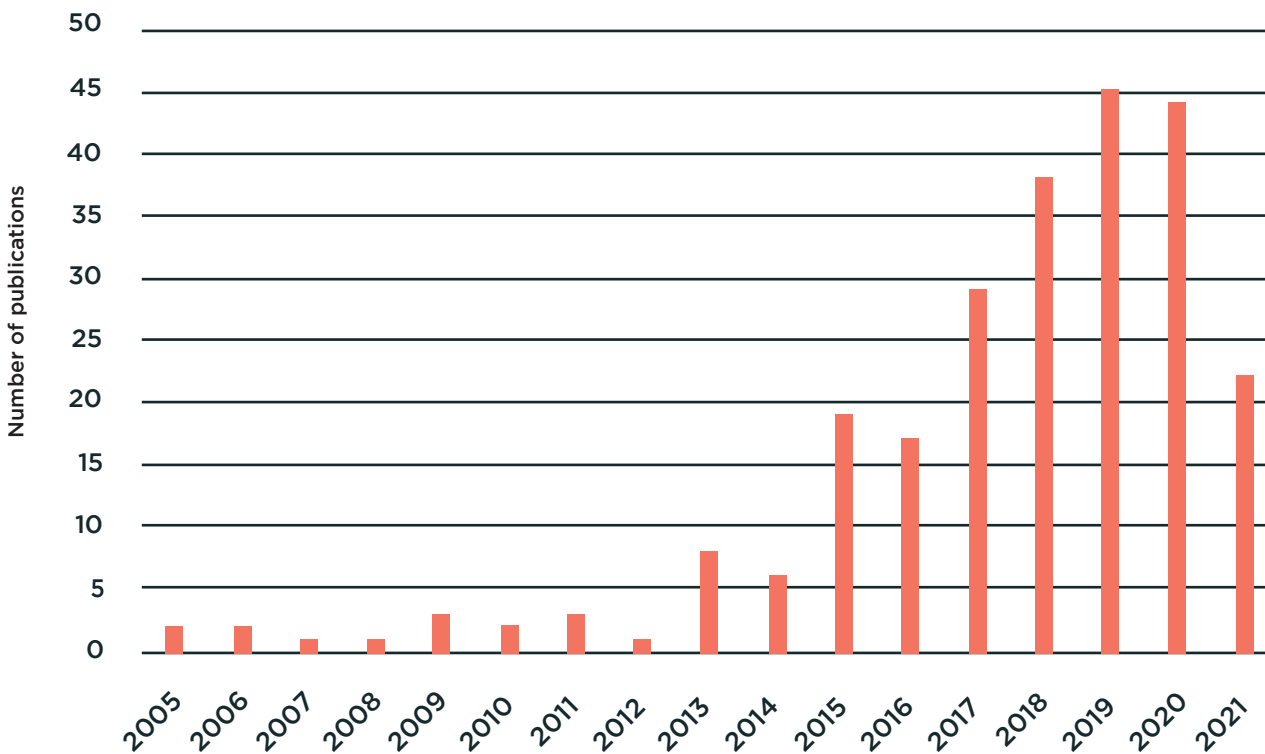


Figure 31: Number of publications in cluster 12 over time.

Cluster 12 revolves around topics related to better waste management, especially in urban environments. The cluster has a more concentrated area dedicated to papers on smart cities,

often citing the role of this type of technological development in ensuring sustainability. Other topics involve: reverse logistics, recycling, development of indicators and frameworks to

assess urban sustainability, and circular economy. Cluster 12 is more recent, with a growing number of publications since 2013, while apparently stabilizing recently.

2.13. Cluster 13 - Indicators and sustainability assessment

Cluster 13 contains 242 papers and is located at the boundary between the upper and lower macro-region, forming a connection cluster. Figures 32 and 33 show the word cloud and the number of publications of the cluster over time, respectively. Cluster 13 is spread out, and can be divided



Figure 32: Word cloud for cluster 13.

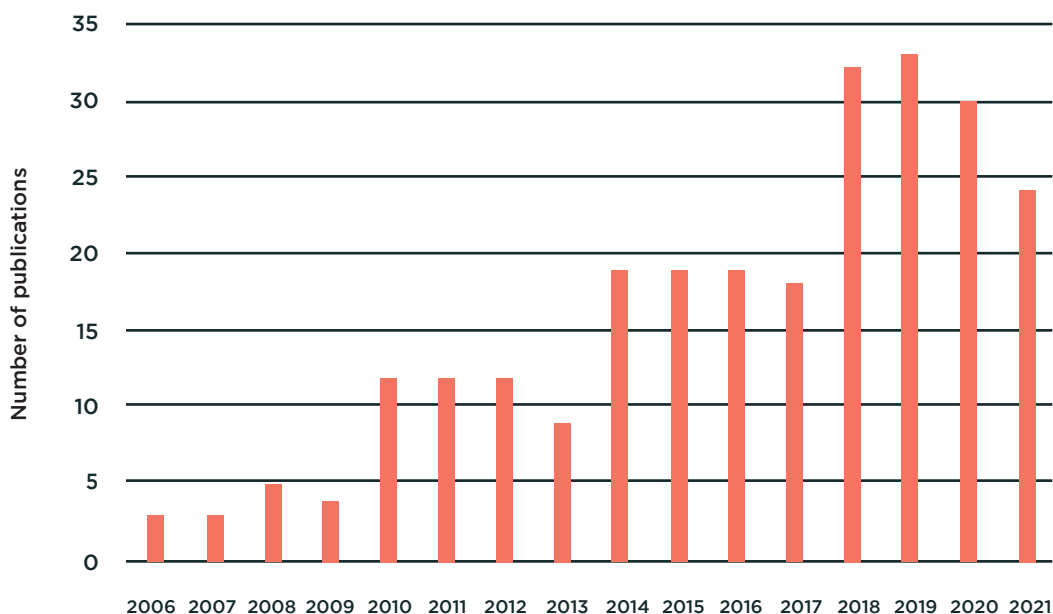


Figure 33: Number of publications in cluster 13 over time.

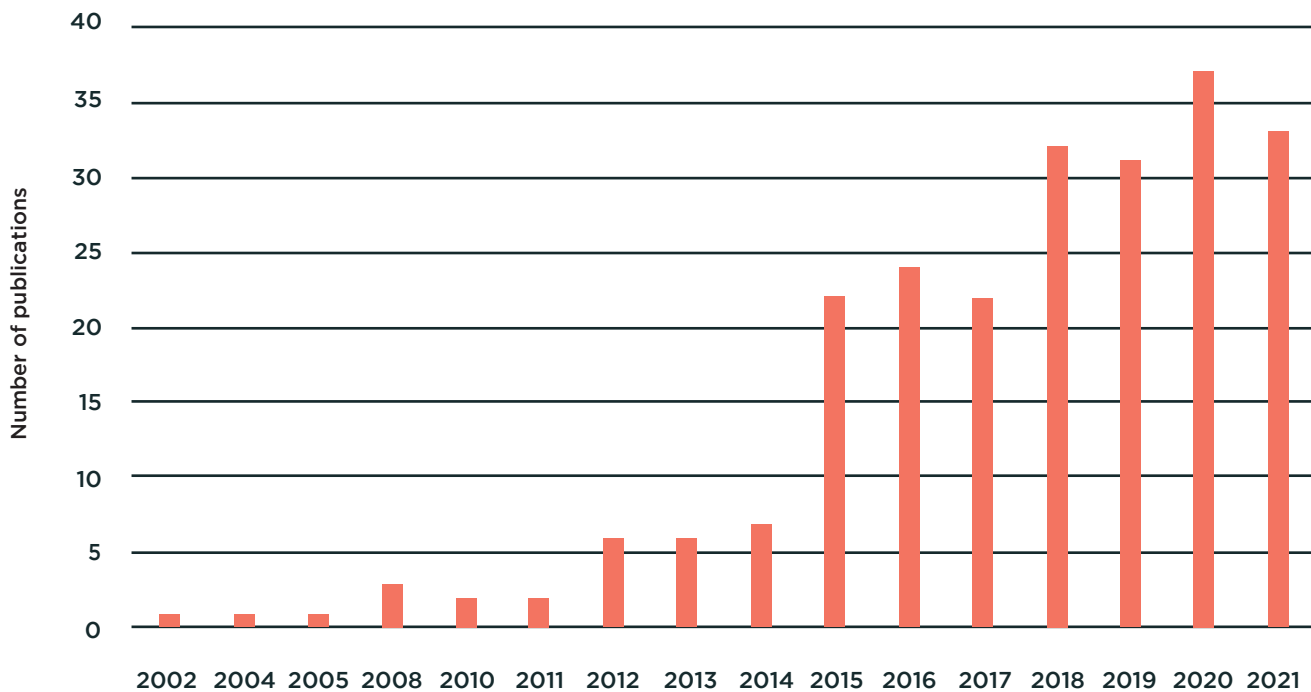
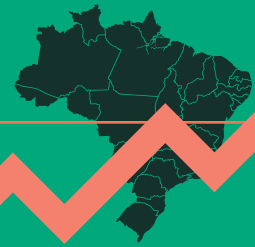


Figure 35: Number of publications in cluster 14 over time.

2.15. Cluster 15 - Amazonia and Ecosystem Services

Cluster 15 is more isolated than the others at the left end of the network. The cluster has a more concentrated region, but also spreads to the border between the upper and lower macro-region. Cluster 15 contains 191 items. Figures 36 and 37 present the word cloud and the number of publications in the cluster over time, respectively.

In the most concentrated and distant part of the network, topics revolve around the Amazon rainforest, especially seeking to assess how different actions affect the biome and ecosystem services. Several studies analyze the potential use of Amazonian products. In the more spread out part of the cluster, on the border between the two macro-regions, the

topic is ecosystem services, mainly impact assessment. Cluster 15 starts to emerge in the year 2000 with an increasing trend until the year 2008, when it stabilizes. The number of publications started growing again after 2013.

Considerations about cluster analysis

The analysis of the 15 largest clusters corresponded to 4946 publications, which is equivalent to 55.3% of the network. This means that several other topics were present in the network, but formed smaller clusters. Some examples are the clusters of innovation management (160 papers), circular economy (152 papers), and biohydrogen (150 papers). Figure 38 shows the evolution in the number of publications of the top 15 clusters over time.

Here, it is possible to highlight some points, such as the presence and resilience of

cluster 1 (agricultural production systems); the rapid growth of cluster 2 (challenges and solutions for vinasse), especially from 2013; the sustained and simultaneous growth of clusters 3 (biodiesel and other products from microalgae) and 4 (evaluation and use of biodiesel) from 2008; the peak of cluster 5 (enzyme production via solid-state fermentation) in 2018; the combined evolution of the number of publications of clusters 6 (fermentation for ethanol production) and 8 (pre-treatment and enzymatic hydrolysis) in the last decade; as well as the growth of the

relevance of sustainability issues, demonstrated by the growth in publications of clusters 13 (sustainability indicators and assessment) and 14 (corporate sustainability). Cluster analysis also showed that publications on the Brazilian bioeconomy seem to focus so far on energy products, mainly biofuels. The network macro-region is mostly focused on topics associated with the production of ethanol, biodiesel, and energy co-generation, followed by topics pertaining to agricultural systems, sustainability, and governance.

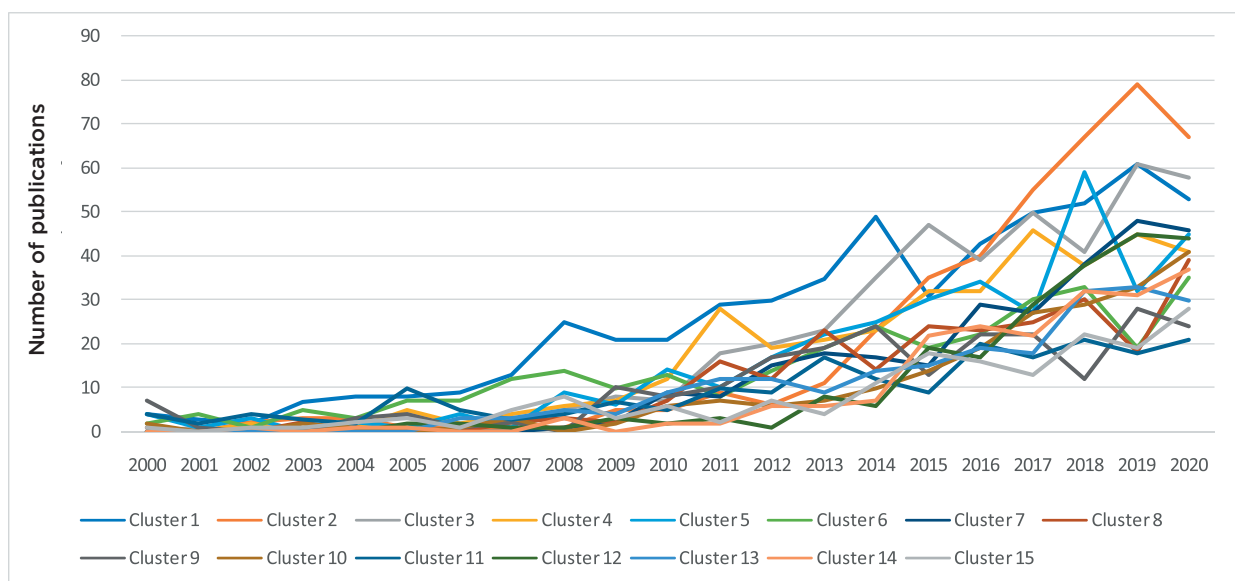


Figure 38: Evolution in the number of publications of the 15 largest clusters over time.

3. International cooperation

The construction of the network of scientific publications on the Brazilian bioeconomy took into account both exclusively national publications and publications made in partnership with other countries. Figures 39 and 40 present the countries that most cooperated with Brazil

in the generated network. The United States appears as the main partner (694 publications), followed by Portugal and England, with about half of the publications (just over 300). Figures 38 and 39 emphasize that Brazil publishes most in partnership with countries in North

America and Europe. In Latin America, Colombia, Mexico, Chile, and Argentina are the most relevant partnerships. Figure 40 shows that partnerships in publications are global, covering authors from countries on all continents.

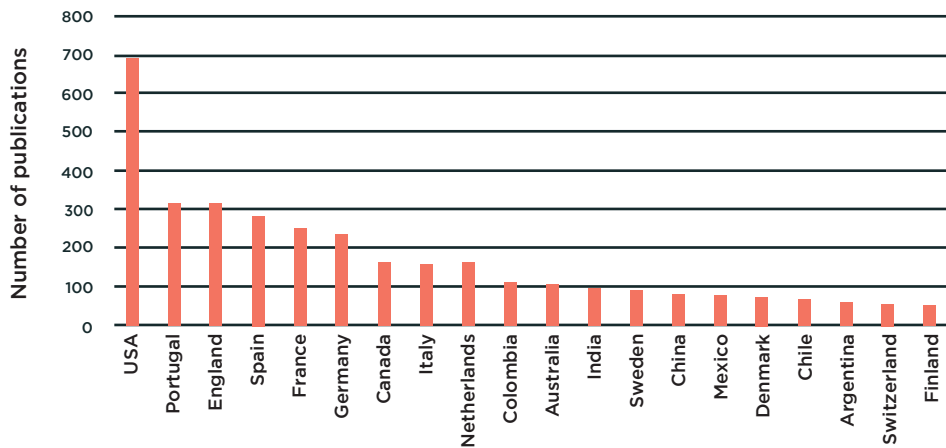


Figure 39: 20 countries that cooperated most with Brazil in the publications of the Brazilian bioeconomy network.

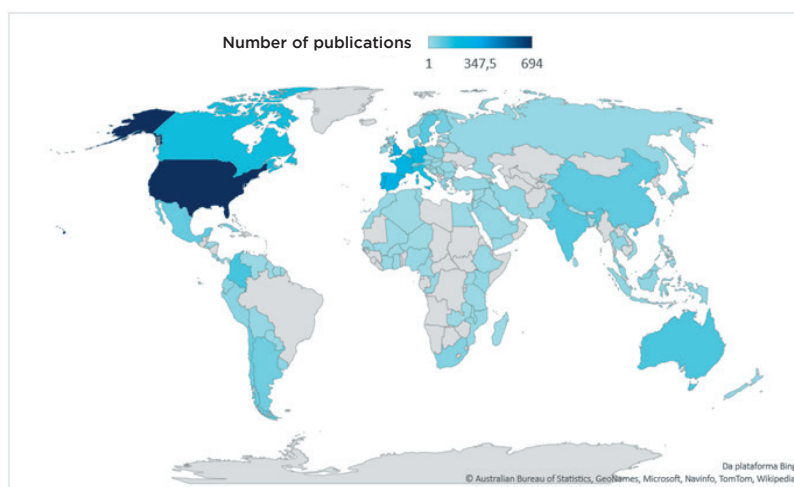


Figure 40: Map showing the countries that cooperated with the publications in the Brazilian bioeconomy network.

4. Brazilian institutions that most publish on bioeconomy

From the data collected, it was also possible to identify the Brazilian institutions that most published on bioeconomy through the network generated (Figure 41). It is clear that the

organizations are mainly public universities, except for the Brazilian Agricultural Research Enterprise (EMBRAPA) and the National Center for Research in Energy and Materials (CNPEM).

Also, please note that among the 15 organizations that most published on bioeconomy, there is a prevalence of institutions from the Southeast and South regions (Figure 42)⁴.

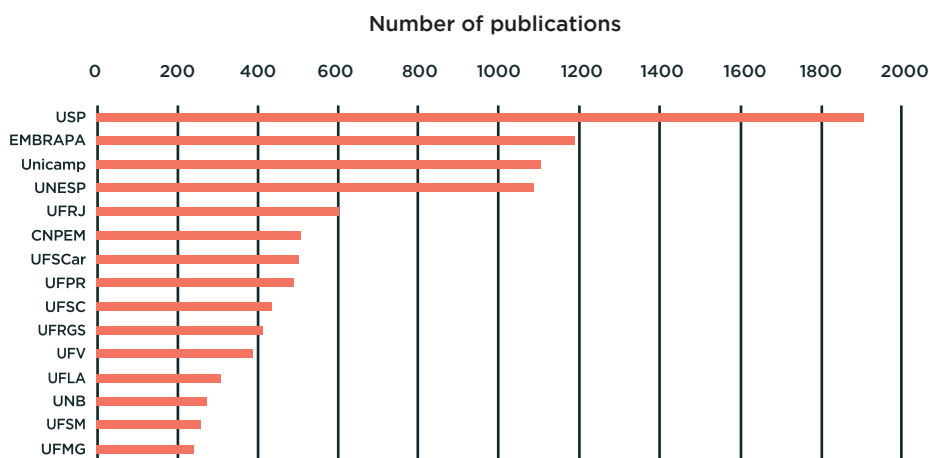


Figure 41: 15 Brazilian institutions that published most in the Brazilian bioeconomy network.

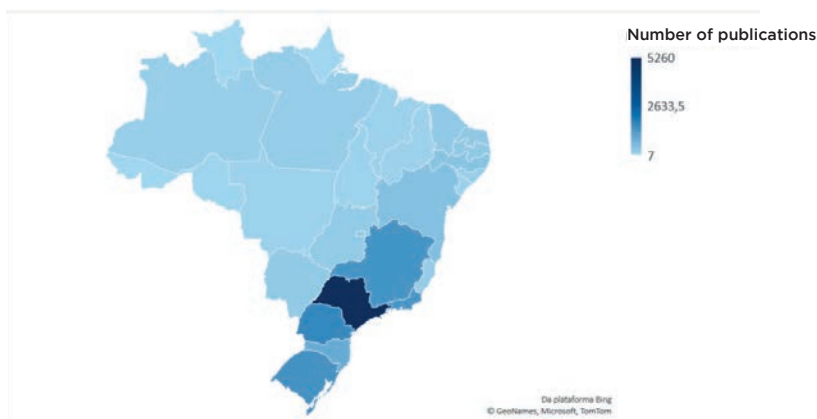


Figure 42: Map of the number of publications by state.

⁴ Institutions operating in more than one state were not accounted for in Figure 42.

5. Final considerations

This Report presented the panorama of Brazilian scientific production on bioeconomy, based on the semantic similarity network built from a snowball methodology for Brazilian publications on bioeconomy stored in the Web of Science (WoS) database. The results emphasized mainly: the predominance of studies focused on the energy sector, mainly related to the production of biofuels such as **ethanol** and **biodiesel**; the strong presence of studies focused on **agricultural**

techniques and systems, mainly related to the **sugarcane** sector; and the great interest in **waste reuse and repurposing**.

On the other hand, it is important to highlight the presence of **Cluster 15 -Amazon and ecosystem services**, which reveals that the network also contains studies on the bioeconomy of biodiversity. The snowball methodology retrieved papers discussing **management and innovation**, which are important areas of knowledge in the context of the bioeconomy. This is made evident by clusters

14 -Corporate sustainability and 13 -Indicators and sustainability assessment.

The data collected had some limitations, such as the search exclusively on the Web of Science database and partial characterization of the network. Future editions will aim to reduce the study's limitations and deepen the investigation of the Brazilian bioeconomy by seeking new data sources, such as other scientific and patent bases, in order to broaden the scope of research on the topic.

Thank you

CGEE/MCTI team of the ODBio project.

Staff

Chief Executive Officer

Marcio de Miranda Santos
(until 28/02/2022)

Fernando Cosme Rizzo Assunção
(from 03/01/2022)

Directors

Luiz Arnaldo Pereira da Cunha

Regina Silverio (Supervision)

Coordinator

Marcelo Poppe

CGEE Technical Team

Barbara Bressan

Daniella Fartes

Emilly Silva

Jackson Maia

João Pedro Arbache