mathematics

MDPI

Article

# Trends in Agroforestry Research from 1993 to 2022: A Topic Model Using Latent Dirichlet Allocation and HJ-Biplot

**Karime Montes-Escobar [1,2,*], Javier De la Hoz-M [2,3], Mónica Daniela Barreiro-Linzán [1], Carolina Fonseca-Restrepo [4], Miguel Ángel Lapo-Palacios [1], Douglas Andrés Verduga-Alcívar [1] and Carlos Alfredo Salas-Macias [5]**

1 Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad Técnica de Manabí, Portoviejo 130105, Ecuador; monica.barreiro@utm.edu.ec (M.D.B.-L.); miguel.lapo@utm.edu.ec (M.Á.L.-P.); douglas.verduga@utm.edu.ec (D.A.V.-A.)
2 Department of Statistics, University of Salamanca, 37008 Salamanca, Spain; jdelahoz@unimagdalena.edu.co
3 Facultad de Ingeniería, Universidad del Magdalena, Santa Marta 470004, Colombia
4 Departamento de Veterinaria, Facultad de Ciencias Veterinarias, Universidad Técnica de Manabí, Portoviejo 130105, Ecuador; carolina.fonseca@utm.edu.ec
5 Carrera de Agronomía, Facultad de Ingeniería Agronómica, Universidad Técnica de Manabí, Portoviejo 130105, Ecuador; carlos.salas@utm.edu.ec
* Correspondence: karime.montes@utm.edu.ec; Tel.: +593-99-150-13-48

**Abstract:** Background: There is an immense debate about the factors that could limit the adoption of agroforestry systems. However, one of the most important is the generation of scientific information that supports the viability and benefits of the proposed techniques. Statistical analysis: This work used the Latent Dirichlet Allocation (LDA) modeling method to identify and interpret scientific information on topics in relation to existing categories in a set of documents. It also used the HJ-Biplot method to determine the relationship between the analyzed topics, taking into consideration the years under study. Results: A review of the literature was conducted in this study and a total of 9794 abstracts of scientific articles published between 1993 and 2022 were obtained. The United States, India, Brazil, the United Kingdom, and Germany were the five countries that published the largest number of studies about agroforestry, particularly soil organic carbon, which was the most studied case. The five more frequently studied topics were: soil organic carbon, adoption of agroforestry practices, biodiversity, climatic change global policies, and carbon and climatic change. Conclusion: the LDA and HJ-Biplot statistical methods are useful tools for determining topicality in text analysis in agroforestry and related topics.

**Keywords:** text analysis; LDA; HJ-biplot; topic diversity; modeling method

**MSC:** 62-XX; 62-07; 62Pxx; 62P12

check for updates

**Citation:** Montes-Escobar, K.; De la Hoz-M, J.; Barreiro-Linzán, M.D.; Fonseca-Restrepo, C.; Lapo-Palacios, M.Á.; Verduga-Alcívar, D.A.; Salas-Macias, C.A. Trends in Agroforestry Research from 1993 to 2022: A Topic Model Using Latent Dirichlet Allocation and HJ-Biplot. *Mathematics* **2023**, *11*, 2250. https://doi.org/10.3390/math11102250

Academic Editors: Víctor Amor-Esteban and David Almorza-Gomar

Received: 7 April 2023
Revised: 6 May 2023
Accepted: 9 May 2023
Published: 11 May 2023



## 1. Introduction

At present, several investigations have shown that there is a high degree of pollution and environmental degradation [1–11] in ecosystems, mainly generated by human activities that are focused on meeting the needs of a constantly growing population [12]. Within this conglomerate of activities, it is possible to identify the relentless use of pesticides and chemical synthesis inputs in the agricultural sector, the excessive use of land, the deforestation of forests, the constant change of land use for agricultural exploitation, and the constant emissions generated by the industry that considerably strengthen the greenhouse effect, among others.

Therefore, society has seen the need to generate mechanisms that help mitigate these undesirable effects, leading the world to hold events such as the Rio de Janeiro summit (or the United Nations Conference on Environment and Development) (1992), the Program or

Agenda 21, the agreement on climate change and the agreement on biological diversity, the declaration on relative principles of the forest, the World Conference on Social Development (1995), the United Nations Framework Convention on Climate Change (1997), the Millennium Summit in Geneva (2000), the United Nations Conference on Sustainable Development, Rio +10 (2002), and the annual United Nations Conference on Climate Change (1995), among others, whose objectives are to try, through holistic approaches, to evaluate possible mitigation actions.

As a result of these events, based on scientific research, it has been established that the greatest threat in the medium term is the accumulation of greenhouse gases generated during human activities. In this aspect, the gas that causes the greatest concern, due to its abundance, is $CO_2$ [13,14], which can be stored by forests [15–22] or in the structure of trees in production systems [23–27].

In relation to this concept, and in order to mitigate unwanted effects, the inclusion of trees in productive systems or agroforestry techniques is proposed as an environmentally friendly alternative. While agroforestry research has provided the necessary scientific foundations for it to be considered a sustainable practice, there is still much to be done to efficiently adopt it by producers through extension and decision makers.

Within this order of ideas, there is an immense debate about the factors that could limit the adoption of agroforestry systems; many authors attribute this fact to the uniqueness and complexity of the technology [28]. On this subject, Ref. [29], after reviewing various studies regarding the adoption of agroforestry technology in 21 countries, found that preferences and resource endowments are the most frequent factors. However, it is undeniable that one of the elements that influences the adoption of production systems is the prior generation of scientific information [28,30,31] that supports the feasibility and benefits of the proposed techniques.

In any case, it is important to highlight that the development of research on agroforestry is being highly promoted by governmental and nongovernmental extension agencies that work together with farmers, which is an encouraging forecast for the future. However, the research agenda, until now, has given high priority to some issues and neglected others, which together could generate technologies for the benefit of the earth and its users [32].

In this work, we present a Latent Dirichlet Allocation (LDA) topic modeling method to identify and interpret scientific topics relating to existing topics or categories in a set of documents [33]. Thus, from a collection of documents called the *corpus*, LDA, based on Bayesian models, generates a probabilistic extension of the latent semantic analysis [33,34]. In addition, the LDA assumes an a priori sparse Dirichlet distribution on the topics in the document using Gibbs sampling [35]. Gibbs is able to determine the topic probabilities of the documents that mix various topics in different proportions. On the other hand, the HJ-Biplot method [36] is proposed in this work to obtain a more precise data evaluation, revealing the existing relationships between the analyzed data.

## 2. Materials and Methods

The procedure for the topic modeling analysis through Latent Dirichlet Allocation [37] was divided into four stages: (1) search and collection of articles; (2) preprocessing; (3) construction of the LDA model; (4) labeling of topics (Figure 1).

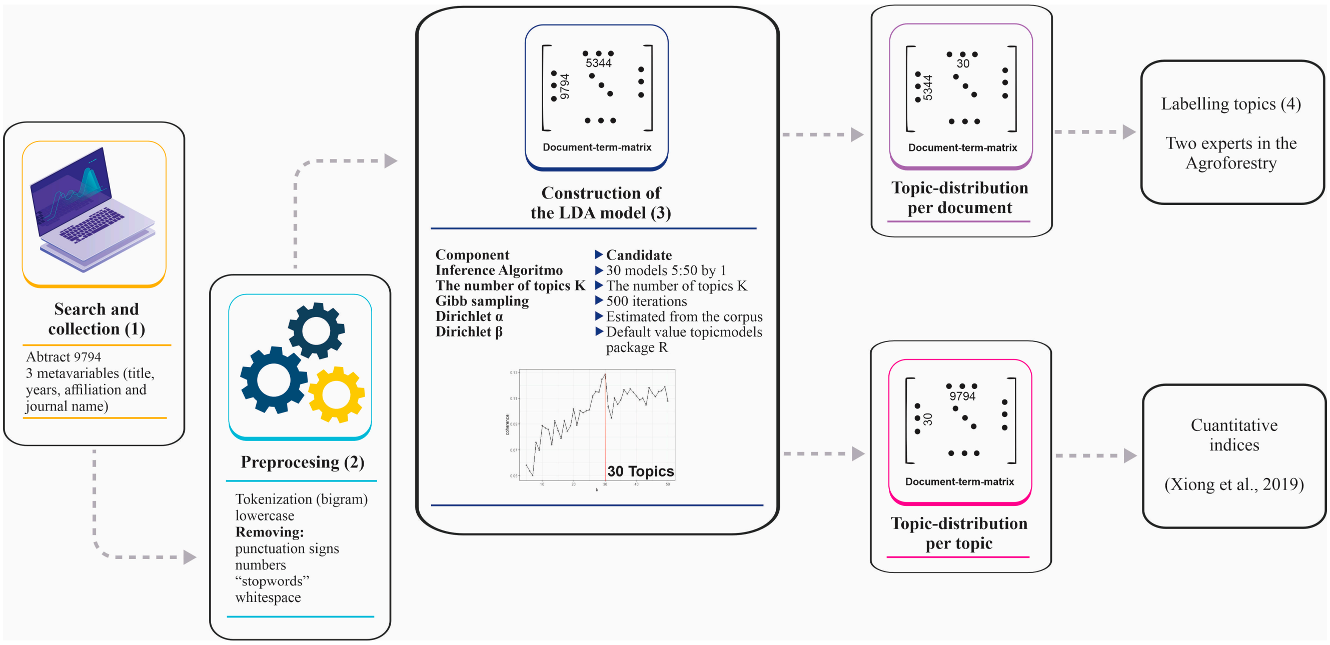


**Figure 1.** Scheme of the methodological process used in the identification of research topics in agroforestry through Latent Dirichlet Allocation [38].

### 2.1. Data Collection

The search for articles was carried out through Scopus, utilizing the following query: “TITLE-ABS-KEY (agroforestry) AND (EXCLUDE (PUBYEAR, 2024) OR EXCLUDE (PUBYEAR, 2023)) AND (LIMIT-TO (DOCTYPE, “ar”)) AND (LIMIT-TO (LANGUAGE, “English”))”. We decided to select this database since it is one of the most used by researchers [39]. The inclusion criteria focused on selecting research articles and reviews published in the English language between 1993 and December 2022. The search was conducted on 10 March 2023.

The preliminary database with the documents obtained after executing the search query contained 9966 documents; this initial sample was subjected to a filtering process, where documents that were repeated or misclassified or that contained no abstract were eliminated. The final sample obtained consisted of 9794 documents.

The procedure for the identification of topics through LDA was divided into three stages: (i) preprocessing, (ii) creation of the LDA model, and (iii) labeling topics.

### 2.2. Preprocessing Texts

Data processing in this part of the study was carried out using LDAShiny [40], an open-source package of the R programming language [41], which contains the development of a tool that provides a web-based graphical user interface to perform a review of the scientific literature under the Bayesian approach of Latent Dirichlet Allocation (LDA) and machine learning algorithms.

To increase the coherence of the topics, each abstract was tokenized using bigrams, which are the combination of consecutive unigrams. Although this process seems trivial, since the text is downloaded to the computer in readable format, it must be converted to lowercase and all punctuation marks, dashes, brackets, numbers, space blanks, and other characters removed. In addition, a standard list of words called “stopwords” was identified and eliminated, since their main function is to make a sentence grammatically correct (i.e., articles and prepositions).

### 2.3. Creation Model Latent Dirichlet Allocation

Topic models are latent variable models of documents that use correlations between words and latent semantic themes in a collection of documents [42]. This definition assumes that the expected number of topics, k, (i.e., latent variables) must be established a priori. Thus, selecting the right number of topics for a given collection of articles is not trivial. This challenge has been approached in a variety of ways, with the goal of striking a balance between the requirement for a large number of topics to cover all of the documents in the collection and the need for a small number of topics to ensure that the findings are intelligible. Simulations were run with k ranging from 5 to 50 and an inference process, Gibbs sampling [35], with 500 iterations was used. A topic coherence measure [43] was used to determine the quality of the LDA model, which is a measure of a topic model from the standpoint of human interpretability and is considered a more suitable measure than computational metrics such as perplexity [44].

### 2.4. Labeling Topics

First, a naive labeling algorithm provided by the package textmineR [45] based on bigrams was employed (these naive labels are based on P(bi-gram | topic) − P(bi-gram)). However, because these algorithms have a very limited capacity for understanding the latent meanings of human language, we decided to use manual labeling, which is considered a standard in topic modeling [46]. Two agroforestry professionals, with more than ten years of experience, manually identified the topics using two sources of information: the most common word lists (most likely) and a sample of the titles. Then, they summarized the 10 most-loaded articles.

### 2.5. Quantitative Indices Used to Analyze the Trend of Topics

We used some quantitative indices proposed by Xiong et al. [38], which were obtained by adding document–topic and topic–word distributions, in order to make the results and findings clear. The description of the indexes is as follows: the distribution of topics over time is obtained by:

$$\left(\theta_k^y\right) = \frac{\Sigma_{m\in y}\theta_{mk}}{n^y} \tag{1}$$

where $m\epsilon y$ represents articles published in a given year, $\theta_{mk}$ represents the proportion of the $k$-th topic in each item, and $n^y$ represents the total number of articles published in the year [38].

We used simple regression slopes for each topic, with the year as a dependent variable and the fraction of topics in the corresponding year as the response variable, to make it easier to characterize the topics in terms of their tendency. [33]. The slopes obtained for each topic were classified as positive or negative at a level of statistical significance of 0.01 and were classified as positive or negative trends, respectively.

Additionally, the VOSviewer software [47] was used to generate the co-authorship network, using countries as the unit of analysis, and employing the full counting method. Only countries with at least 20 documents were considered.

### 2.6. HJ-Biplot

Finally, the HJ-Biplot [36] was used. It is a multivariate statistical technique that allows a graphic representation of the information contained in the rows (individuals) and columns (variables) of a matrix of data [36]. This technique was chosen since it offers a more precise data evaluation, highlighting the relationships between the parts, years, and topics. For this analysis, the Multbiplot software [48] was used, which allowed us to have a fast and easy way to incorporate our tables from a *.xls format.

## 3. Results

The summary generated included basic statistics about the analyzed dataset; it is presented in Table 1. Documents stemmed from 1564 different journals and were published over the course of three decades. A total of 25,174 authors were involved in the scientific production on agroforestry. Among the papers, 771 were single-authored papers, whereas the overall collaboration index of the sample was 2.61.

**Table 1.** Statistics about the analyzed dataset.

| Description | Results |
|---|---|
| Main information about data | |
| Timespan | 1993:2022 |
| Sources (journals, books, etc.) | 1564 |
| Documents | 9794 |
| Annual growth rate % | 7.63 |
| Document average age | 9.8 |
| Average citations per doc | 22.48 |
| References | 1 |
| Document contents | |
| Keywords plus (ID) | 19,986 |
| Author's keywords (DE) | 20,415 |
| Authors | |
| Authors | 25,174 |
| Authors of single-authored docs | 654 |
| Authors collaboration | |
| Single-authored docs | 771 |
| Co-authors per doc | 4.41 |
| International co-authorships % | 37.99 |
| Document types | |
| Article | 9173 |
| Review | 621 |

Table 2 shows the most influential journals in terms of article count. These journals were distributed in different subject areas, such as environmental science, environmental stewardship, ecosystem services, economic investment for research, policies on sustainable land use, and the participation of entities, specialized policy, etc. The journal with the most published articles was *Agroforestry Systems*, which contributed about 16% of the documents analyzed.

**Table 2.** Most influential sources of the documents analyzed.

| Sources | Articles |
|---|---|
| Agroforestry Systems | 1549 |
| Agriculture, Ecosystems and Environment | 307 |
| Sustainability (Switzerland) | 197 |
| Forest Ecology and Management | 188 |
| Forests | 149 |
| Plant and Soil | 118 |
| Forests Trees and Livelihoods | 112 |
| Science of the Total Environment | 82 |
| Agricultural Systems | 81 |
| Land Use Policy | 76 |
| Biodiversitas | 75 |
| Small-Scale Forestry | 75 |
| Journal of Sustainable Forestry | 74 |
| Land | 72 |
| Biodiversity and Conservation | 68 |
| Land Degradation and Development | 64 |
| Journal of Environmental Management | 62 |
| Plos One | 60 |
| Journal of Forestry Research | 56 |
| Range Management and Agroforestry | 53 |
| Agricultural and Forest Meteorology | 52 |
| Agronomy | 50 |
| Current Science | 50 |
| Environmental Management | 50 |
| Agronomy for Sustainable Development | 49 |

In terms of the annual number of publications in the evaluated period, we observed an annual growth rate of 7.63%. A low oscillation in the number of publications could be observed in the first decade, while in the last year the number of publications diverged from an exponential trend (Figure 2).

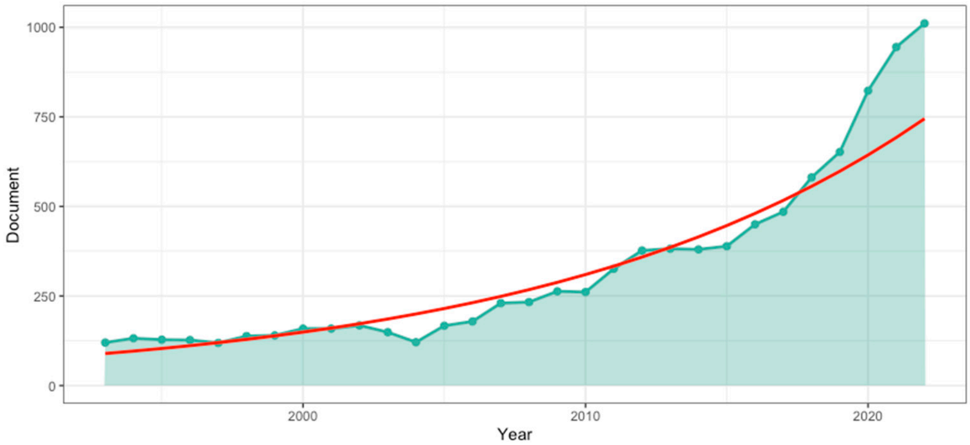


**Figure 2.** Annual trend of document about agroforestry. The red line is the exponential trend.

Figure 3 represents the countries that have published the most research on agroforestry in the world. The United States (1817), India (1114), Brazil (947), the United Kingdom (809), Germany (807), and France (638) were the five countries that appeared as the main sources of high-impact research available in the world scientific literature. They were followed by China (629), Indonesia (591), Kenya (587), Australia (464), Canada (456), and finally Spain (451), with a lower percentage of publications. Other countries did not show the same frequency of publications or the same impact.

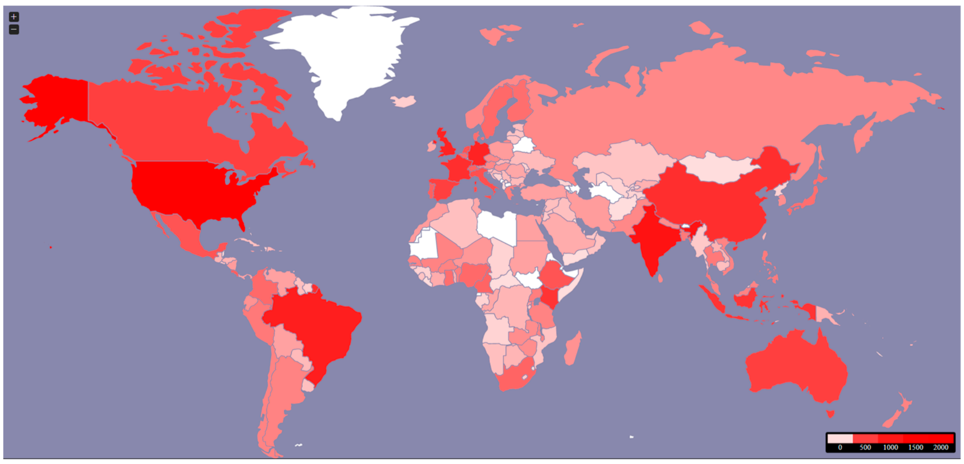

**Figure 3.** Worldwide distribution of 9794 articles on agroforestry research published for the years 1993–2022.

Figure 4 shows the network of international coauthor relationships among 71 countries, the largest set connected. There are seven colors in Figure 4, indicating that these 71 countries are grouped into five clusters. There is a stronger cooperative relationship between countries in the same cluster than between countries in different clusters. Of course, this does not mean that there is no cooperation between countries in different clusters, but that there may be some common research topics among the countries of the same cluster, which makes their cooperation closer. More frequent connections (represented by the thickness of the line) could be observed between the United States and Brazil.

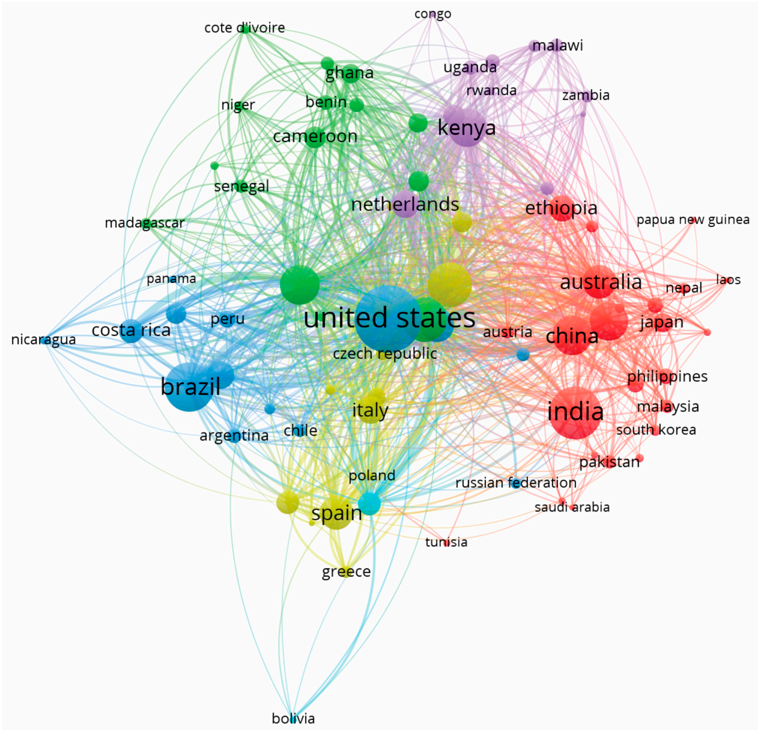


**Figure 4.** Network of international coauthor relationships.

*Topic Modeling Analysis*

Figure 5 shows the coherence score for all LDA models that were examined. According to the findings, the LDA model with the best coherence score had 30 themes ($k$ = 30).

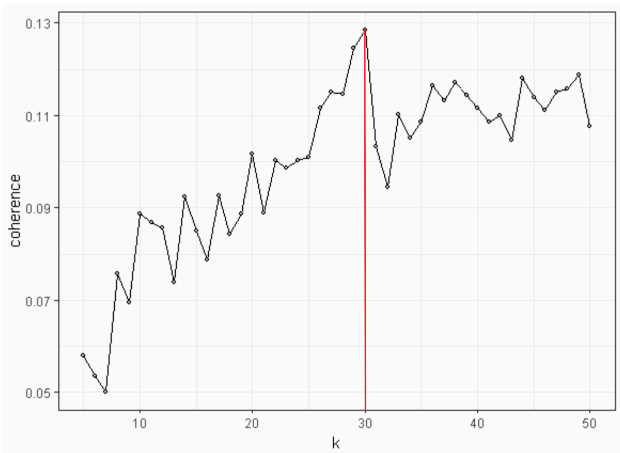


**Figure 5.** Coherence. The red line indicates the number of topics.

The topic distribution from the document $\theta_m$ was added to calculate the average probability $\theta_k^y$ of all the articles published in a given year and to identify the trends (Figure 6). We found that the probabilities of some topics increased gradually over time (color red); these topics were t_3 (raw materials production), t_11 (shade tree), t_12 (land cover), t_16 (adoption of agroforestry practices), t_17 (remote sensors), t_19 (fungal communities), t_21 (carbon and climatic change), t_22 (ecosystem services), t_24 (secondary forests), t_27 (biodiversity), t_28 (soil organic carbon), and t_29 (cork oak). Topics that had a decreasing behavior were (color blue): t_1 (tree species), t_2 (leaf litter), t_10 (alley cropping), t_14 (silvopastoral systems), t_15 (fine root), t_20 (soil water), t_23 (soil fertility), t_26 (production systems), and t_30 (aboveground biomass). The topics where there was no observed trend (color black) were: t_4 (intercrop systems), t_5 (rubber plantations), t_6 (food security), t_7 (fruit tree), t_8 (local knowledge), t_9 (agroforestry systems), t_13 (climatic change global policies), and t_25 (plant species).

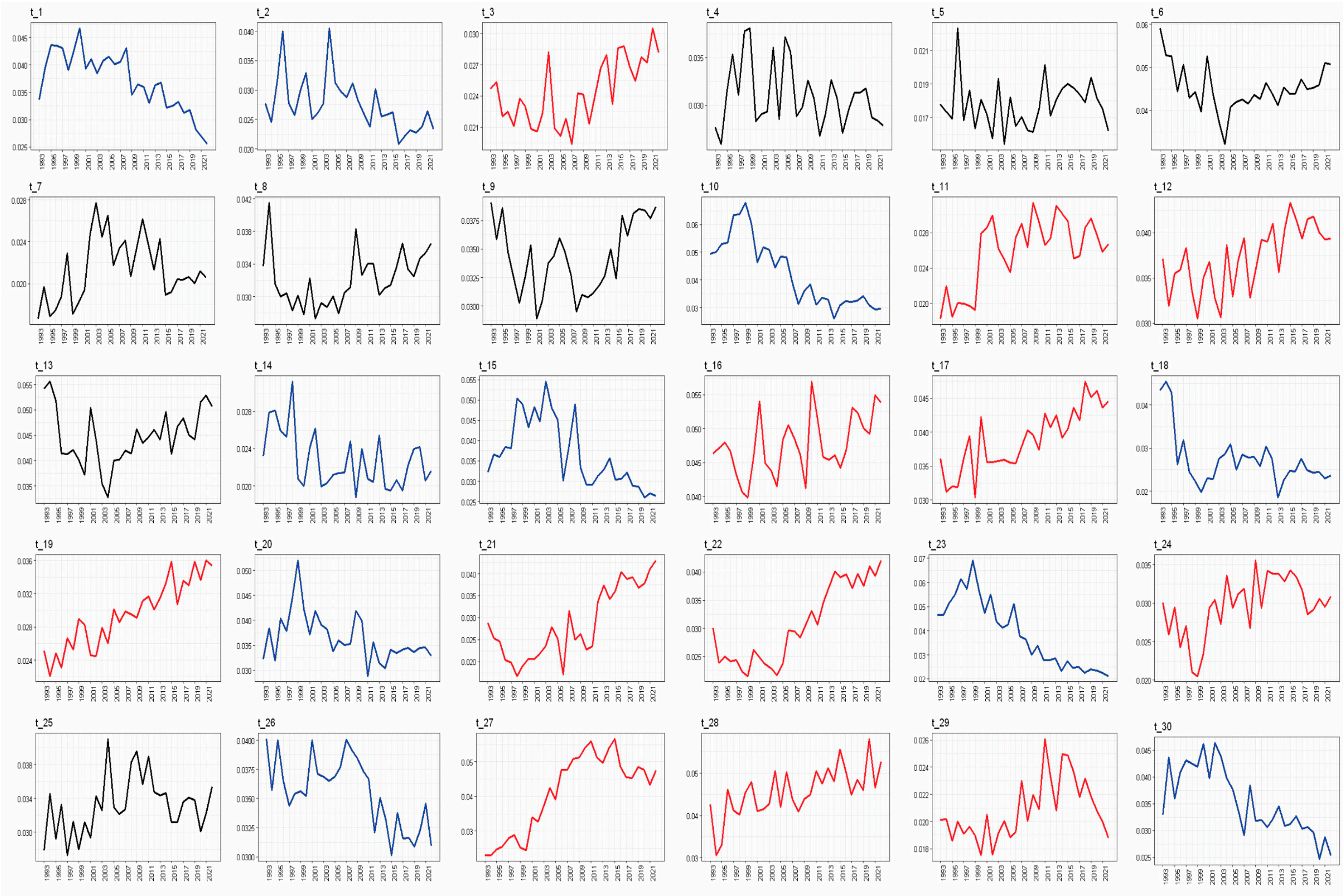


**Figure 6.** Trend of topics: topics that increased gradually over time (color red), topics that had a decreasing behavior (color blue), and topics where there was no observed trend (color black).

Our study found six groups through cluster analysis (Figure 7). The topics: t_16 (adoption of agroforestry practices), t_6 (food security), t_13 (climatic change global policies), t_28 (soil organic carbon), and t_27 (biodiversity) of group six obtained higher probability proportions for the years 1993 and 2022; the topics t_10 (alley cropping), t_23 (soil fertility), t_15 (fine root), t_1 (tree species), and t_30 (aboveground biomass) of group five obtained higher probability proportions for the years 1993 and 2008. The remaining topics had low probability statistics during the years of study.

In Table 3, it is possible to observe the topic names that were generated from the words with the highest number of repetitions. Words were ranked for relevance. After the searches within the articles according to each topic, the 20 words with the highest number of repetitions generated the prevalence rankings. The five main ones were: t_28 (soil organic carbon), t_16 (adoption of agroforestry practices), t_27 (biodiversity), t_13 (climatic change global policies), and t_4 (carbon and climatic change).

Our study also carried out the analysis with the LDA method, where the relationships between topics were observed. The existence of a possible relationship was obtained by statistical analysis of PC according to the matrix generated in the LDA model (Figure 8).

The grouping of topics and years was analyzed using the multivariate analysis method HJ-Biplot, using the theta matrix from the results, which was composed of probability values. First and foremost, a high quality of representation can be observed in the initial factorial plane, with 71.39% of the inertia being effectively absorbed and explained (Figure 9).

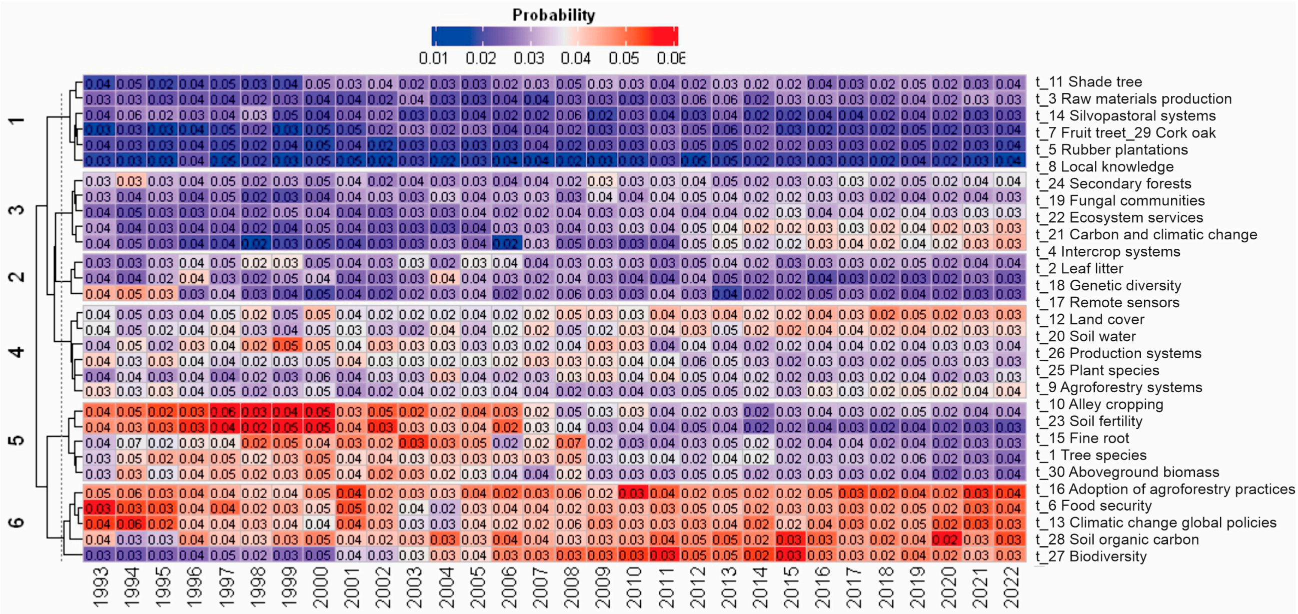


**Figure 7.** Representing the distribution of topics by year through the heat map.

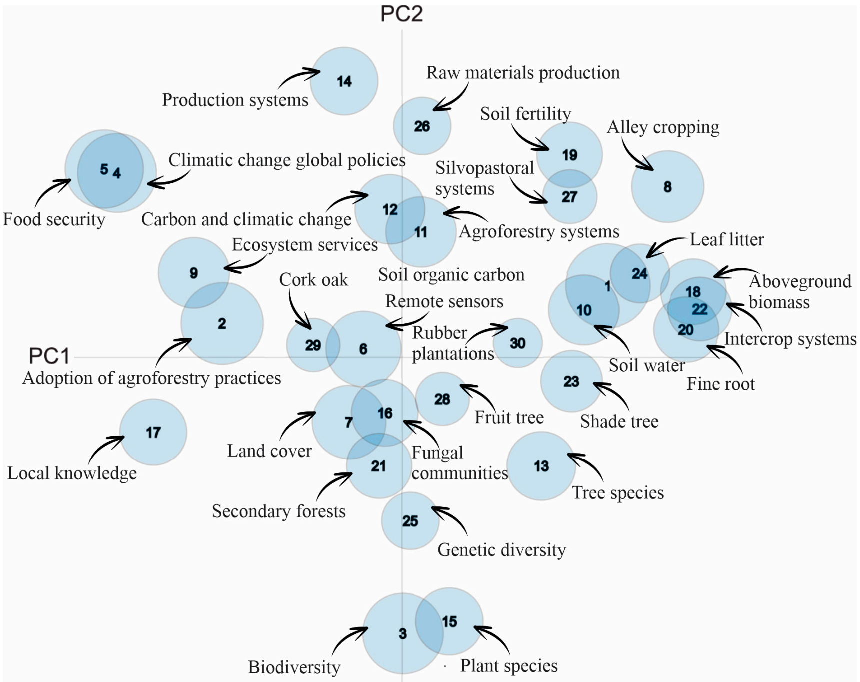


**Figure 8.** Topic similarity map that shows a two-dimensional representation (via multi-dimensional scaling).

**Table 3.** Frequent topics.

| TOPIC | LAB. ALGORIT | LABEL ASIGNED | P | R | ART. | TOPICS TERMS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| t_28 | soil_organ | soil organic carbon | 4.799 | 1 | 825 | soil | organ | depth | soc | properti |
| t_16 | agroforestri_practic | adoption of agroforestry practices | 4.931 | 2 | 785 | farmer | farm | agroforestri | practic | adopt |
| t_27 | speci_rich | biodiversity | 4.527 | 3 | 694 | divers | landscap | speci | habitat | conserv |
| t_13 | climat_chang | climatic change global policies | 4.656 | 4 | 511 | project | develop | polici | approach | base |
| t_21 | climat_chang | carbon and climatic change | 3.36 | 5 | 487 | carbon | climat | chang | stock | climat_chang |
| t_6 | food_secur | food security | 4.594 | 6 | 485 | develop | food | research | sustain | resourc |
| t_15 | fine_root | fine root | 3.265 | 7 | 483 | root | plant | growth | seedl | treatment |
| t_12 | land_cover | land cover | 3.866 | 8 | 392 | land | area | cover | agricultur | degrad |
| t_10 | allei_crop | alley cropping | 3.626 | 9 | 389 | crop | yield | maiz | field | allei |
| t_17 | remot_sens | remote sensors | 4.147 | 10 | 384 | model | data | method | estim | variabl |
| t_20 | soil_water | soil water | 3.508 | 11 | 382 | water | season | temperatur | rainfal | dry |
| t_11 | shade_tree | shade tree | 2.691 | 12 | 315 | coffe | shade | cacao | pest | shade_tree |
| t_3 | raw_materi | raw materials production | 2.597 | 13 | 303 | wood | energi | potenti | product | pine |
| t_26 | product_system | production systems | 3.408 | 14 | 298 | product | econom | cost | benefit | market |
| t_23 | soil_fertil | soil fertility | 3.081 | 15 | 298 | fertil | nutrient | fallow | increas | soil_fertil |
| t_4 | intercrop_system | intercrop systems | 3.017 | 16 | 282 | intercrop | plant | space | wheat | poplar |
| t_25 | plant_speci | plant species | 3.343 | 17 | 268 | speci | plant | nativ | divers | woodi |
| t_22 | ecosystem_servic | ecosystem services | 3.51 | 18 | 265 | manag | agricultur | ecosystem | servic | practic |
| t_30 | aboveground_biomass | aboveground biomass | 3.191 | 19 | 260 | biomass | year | growth | height | stand |
| t_18 | genet_divers | genetic diversity | 2.568 | 20 | 227 | popul | genet | select | variat | trait |
| t_2 | leaf_litter | leaf litter | 2.588 | 21 | 221 | litter | leaf | rate | content | concentr |
| t_8 | local_peopl | local knowledge | 3.326 | 22 | 216 | local | tradit | knowledg | region | import |
| t_14 | silvopastor_system | silvopastoral systems | 2.22 | 23 | 201 | pastur | grass | livestock | product | forag |
| t_24 | secondari_forest | secondary forests | 3.048 | 24 | 177 | forest | af | natur | tropic | restor |
| t_9 | agroforestri_system | agroforestry systems | 3.548 | 25 | 159 | system | agroforestri | agroforestri_system | base | monocultur |
| t_7 | fruit_tree | fruit tree | 2.127 | 26 | 155 | fruit | cocoa | africa | pollin | west |
| t_1 | tree_speci | tree species | 3.358 | 27 | 148 | tree | tree_speci | densiti | canopi | busi_media |
| t_29 | cork_oak | cork oak | 2.124 | 28 | 78 | term | forestri | long | agro | long_term |
| t_5 | rubber_plantat | rubber plantations | 1.782 | 29 | 58 | plantat | rubber | monocultur | palm | oil |
| t_19 | fungal_commun | fungal communities | 3.192 | 30 | 48 | effect | posit | factor | affect | influenc |

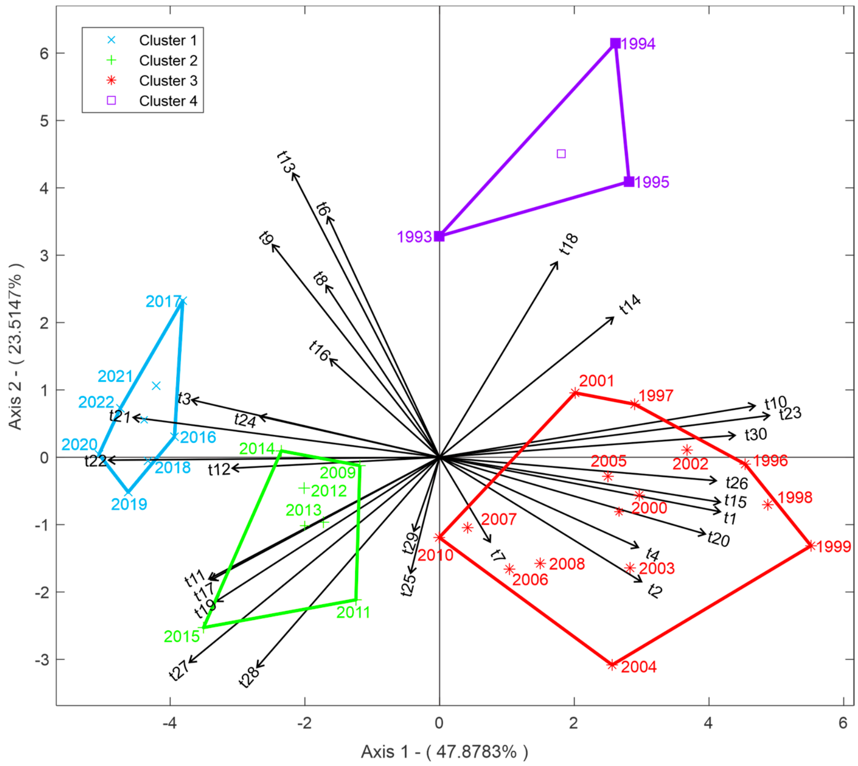


**Figure 9.** Association between topics by year using HJ-Biplot method.

The topics t_14 (silvopastoral systems) and t_18 (genetic diversity) had greater relevance in the years 1993, 1994, and 1995. The topics: t_10 (alley cropping), t_23 (soil fertility), t_30 (aboveground biomass), t_20 (soil water), t_26 (production systems), t_15 (fine root), t_1 (tree species), t_7 (fruit tree), t_4 (intercrop systems), t_2 (leaf litter), t_29 (cork oak), and t_25 (plant species) had greater relevance in the years 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, and 2010.

The topics: t_28 (soil organic carbon), t_27 (biodiversity), t_11 (shade tree), t_19 (fungal communities), t_17 (remote sensors), and t_12 (land cover) had greater relevance in the years 2009, 2010, 2011, 2012, 2013, 2014, and 2015. Finally, in the years 2016, 2017, 2018, 2019, 2020, 2021, and 2022, the investigation of the topics t_22 (ecosystem services), t_21 (carbon and climatic change), t_3 (raw materials production), t_24 (secondary forests), t_9 (agroforestry systems), t_16 (adoption of agroforestry practices), t_8 (local knowledge), t_6 (food security), and t_13 (climatic change global policies) was the most relevant in the last seven years.

## 4. Discussion

The results of this study were consistent with those established by [49], who mentioned the United States, India, Germany, Brazil, Kenya, France, Australia, and the United Kingdom as countries with the highest interest in agroforestry research. Furthermore, some of the most common topics in this study were consistent with this 2023 analysis.

It was possible to notice a gradual increase in interest in research on agroforestry issues; proof of this was the increase over time in the number of publications made, which went from 120 in 1993 to 1011 in 2022, with a total of 145 countries involved in the 9794 articles that made up the study sample. However, there could be other factors with significant impact, such as economic investment for research, policies on sustainable land use, and the participation of entities specialized in agroforestry. In this regard, the United States, for example, has key policies to promote agroforestry; this is the case of the USDA's National Agroforestry Center, which seeks to promote agroforestry science and its adoption. Similarly, Europe has included policies aimed at agroforestry issues in the common agricultural policy (CAP).

In this same sense, multistakeholder partnerships and agroforestry working groups and associations play extremely important roles in the generation and promotion of agroforestry science at different scales [50]. Considering the results of this study, countries such as India (which for the last 25 years has had a solid track record in agricultural and forestry research [51]), which has the Indian Council of Agricultural Research, the Indian Society of Agroforestry, and currently has the landmark national agroforestry policy that provides a framework designed to improve agricultural livelihoods; Germany that has the German Association for Agroforestry (DEFAF); France that has the French Agroforestry Association; Kenya that has the headquarters of World Agroforestry (ICRAF) are making efforts to develop agroforestry practices and economic growth, alleviate poverty, and make significant contributions in terms of environmental quality.

In this way, agroforestry is defined as a form of land use where it is possible to increase productivity, diversify production, and improve sustainability [52]. It is widely known as an agricultural production system where trees are grown together with annual crops and/or with animals, resulting in better complementary relationships between the components and an increase in multiple uses [53]. Additionally, it can provide high quality habitats, which is of great importance for the conservation of biodiversity. It can provide multiple alternatives for the commercialization of products, cushioning, in some cases, market fluctuations. The high levels of biodiversity also provide better ecological services, increasing local functions such as pest control [54] and pollination [55,56] and can improve soil fertility and reduce erosion [57,58].

Therefore, the relationship that existed between the results of the relevance analysis of the words used in the articles that made up the corpus under study was remarkable; the ten words that were the most relevant were: tree, system, soil, agroforestry, species, forest, crop, production, and land. In addition, it was possible to determine that there were topics that were gaining importance during the analysis period in the investigations; these being: soil organic carbon, adoption of agroforestry practices, biodiversity, climatic change global policies, carbon and climatic change, and food security.

## 5. Conclusions

This study provides relevant information about the evolution of science around the agroforestry topic over time and by country, providing a theoretical basis for the development of the discipline and could significantly contribute to decision making by researchers and research centers linked to the sector.

We found that the countries with the highest number of publications in agroforestry are the United States, India, Brazil, the United Kingdom, and Germany. Although there is generally a broad network of international cooperation, the United States and Brazil have more frequent connections in terms of international co-authorship relationships in the field of agroforestry.

Based on the results, we can conclude that there are certain topics related to raw material production, soil cover, adoption of agroforestry practices, biodiversity, and climate change that have gained importance over time, while other topics related to soil fertility, aerial biomass, and tree species have decreased in relevance. Bootstrap analysis also identified groups of topics that had higher probabilities in certain years, suggesting that certain topics may have been more relevant at certain times. Adoption of agroforestry practices, food security, global climate change policies, organic soil carbon, and biodiversity appear to be important topics both in 1993 and 2022, while topics such as alley cropping, soil fertility, fine roots, tree species, and aerial biomass appear to have been more relevant in 1993 and 2008. Overall, these results may be useful in identifying trends and areas of interest in agroforestry research, as well as guiding future research and policies related to agroforestry and sustainable land management.

Finally, the LDA method employed in this review categorized each subject based on subjective observations of high probability words and therefore showed effectiveness in generating responses about the most common topics studied in agroforestry. In contrast, the HJ-Biplot method was able to group topics by year, identifying which agroforestry topics were most relevant and for which years.

**Author Contributions:** Conceptualization, K.M.-E. and C.A.S.-M.; methodology, K.M.-E. and J.D.l.H.-M.; software, K.M.-E. and J.D.l.H.-M.; validation, K.M.-E. and M.D.B.-L.; formal analysis, K.M.-E., J.D.l.H.-M. and C.A.S.-M.; investigation, K.M.-E.; resources, C.F.-R. and M.Á.L.-P.; data curation, K.M.-E. and J.D.l.H.-M.; writing—original draft preparation, C.F.-R.; writing—review and editing, M.Á.L.-P.; visualization, K.M.-E. and D.A.V.-A.; supervision, C.A.S.-M.; project administration, K.M.-E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This study was supported by the Departamento de matemáticas y Estadística. Facultad de Ciencias Básicas, Facultad de Ingeniería Agronómica, Facultad de Ciencias Veterinarias, Departamento de Veterinaria, Portoviejo, Manabí, Ecuador and Universidad del Magdalena, Santa Marta, Colombia.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ayub, M.A.; Usman, M.; Faiz, T.; Umair, M.; Haq, M.A.U.; Rizwan, M.; Ali, S.; Rehman, M.Z.U. Restoration of Degraded Soil for Sustainable Agriculture. In *Soil Health Restoration and Management*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 31–81. [CrossRef]
2. Carson, R. *Silent Spring 1962*; Mariner Books Classics: New York, NY, USA, 2009. Available online: https://www.amazon.com/Silent-Spring-Rachel-Carson/dp/0618249060 (accessed on 22 February 2023).
3. Das, S.; Aria, A.; Cheng, J.-O.; Souissi, S.; Hwang, J.-S.; Ko, F.-C. Occurrence and distribution of anthropogenic persistent organic pollutants in coastal sediments and mud shrimps from the wetland of central Taiwan. *PLoS ONE* **2020**, *15*, e0227367.
4. Dou, R.; Sun, J.; Deng, F.; Wang, P.; Zhou, H.; Wei, Z.; Chen, M.; He, Z.; Lai, M.; Ye, T.; et al. Contamination of pyrethroids and atrazine in greenhouse and open-field agricultural soils in China. *Sci. Total Environ.* **2020**, *701*, 134916. [CrossRef] [PubMed]
5. Litalien, A.; Zeeb, B. Curing the earth: A review of anthropogenic soil salinization and plant-based strategies for sustainable mitigation. *Sci. Total. Environ.* **2020**, *698*, 134235. [CrossRef]
6. Panwar, S. Vulnerability of Himalayan springs to climate change and anthropogenic impact: A review. *J. Mt. Sci.* **2020**, *17*, 117–132. [CrossRef]
7. Sanaullah, M.; Usman, M.; Wakeel, A.; Alam Cheema, S.; Ashraf, I.; Farooq, M. Terrestrial ecosystem functioning affected by agricultural management systems: A review. *Soil Tillage Res.* **2020**, *196*, 104464. [CrossRef]
8. Ramirez-Llodra, E. Deep-Sea Ecosystems: Biodiversity and Anthropogenic Impacts. In *The Law of the Seabed*; Brill Nijhoff: Leiden, Belgium, 2020; pp. 36–60.
9. Rivera-Rivera, D.M.; Escobedo-Urías, D.C.; Jonathan, M.P.; Sujitha, S.B.; Chidambaram, S. Evidence of natural and anthropogenic impacts on rainwater trace metal geochemistry in central Mexico: A statistical approach. *Water* **2020**, *12*, 192. [CrossRef]
10. Yang, L.; Zhou, Y.; Shi, B.; Meng, J.; He, B.; Yang, H.; Yoon, S.J.; Kim, T.; Kwon, B.-O.; Khim, J.S.; et al. Anthropogenic impacts on the contamination of pharmaceuticals and personal care products (PPCPs) in the coastal environments of the Yellow and Bohai seas. *Environ. Int.* **2020**, *135*, 105306. [CrossRef]
11. Wu, J.; Kurosaki, Y.; Du, C. Evaluation of climatic and anthropogenic impacts on dust erodibility: A case study in Xilingol Grassland, China. *Sustainability* **2020**, *12*, 629. [CrossRef]
12. Winch, D.; James, P. *An Essay on the Principle of Population: Or, a View of Its Past and Present Effects on Human Happiness: With an Inquiry into Our Prospects Respecting the Future Removal or Mitigation of the Evils Which It Occasions*; Cambridge University Press: Cambridge, UK, 1992.
13. Peters, G.P.; Andrew, R.M.; Canadell, J.G.; Friedlingstein, P.; Jackson, R.B.; Korsbakken, J.I.; Le Quéré, C.; Peregon, A. Carbon dioxide emissions continue to grow amidst slowly emerging climate policies. *Nat. Clim. Chang.* **2020**, *10*, 3–6. [CrossRef]
14. Sadatshojaie, A.; Rahimpour, M.R. $CO_2$ emission and air pollution (volatile organic compounds, etc.)-related problems causing climate change. *Curr. Trends Future Dev. (Bio-) Membr. Membr. Environ. Appl.* **2020**, 1–13. [CrossRef]
15. Canadell, J.G.; Jackson, R.B.; Ciais, P.; Besnard, S.; Yao, Y.; Yue, C.; Carvalhais, N.; Poulter, B.; Pugh, T.; Chave, J.; et al. The global carbon balance of forests based on flux towers and forest age data. In *AGU Fall Meeting Abstracts*; American Geophysical Union: Washington, DC, USA, 2019; p. B34D-01. Available online: https://ui.adsabs.harvard.edu/abs/2019AGUFM.B34D..01C/abstract (accessed on 22 February 2023).
16. Chambers, J.Q.; Higuchi, N.; Tribuzy, E.S.; Trumbore, S.E. 'Carbon sink for a century. *Nature* **2001**, *410*, 429. [CrossRef] [PubMed]

17. Erb, K.-H.; Kastner, T.; Luyssaert, S.; Houghton, R.A.; Kuemmerle, T.; Olofsson, P.; Haberl, H. Bias in the attribution of forest carbon sinks. *Nat. Clim. Chang.* **2013**, *3*, 854–856. [CrossRef]
18. Huang, L.; Zhou, M.; Lv, J.; Chen, K. Trends in global research in forest carbon sequestration: A bibliometric analysis. *J. Clean. Prod.* **2020**, *252*, 119908. [CrossRef]
19. Pan, Y.; Birdsey, R.A.; Fang, J.; Houghton, R.; Kauppi, P.E.; Kurz, W.A.; Phillips, O.L.; Shvidenko, A.; Lewis, S.L.; Canadell, J.G.; et al. A Large and Persistent Carbon Sink in the World's Forests. *Science* **2011**, *333*, 988–993. [CrossRef] [PubMed]
20. Pugh, T.A.M.; Lindeskog, M.; Smith, B.; Poulter, B.; Arneth, A.; Haverd, V.; Calle, L. Role of forest regrowth in global carbon sink dynamics. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 4382–4387. [CrossRef]
21. Sedjo, R.A. Temperate forest ecosystems in the global carbon cycle. *Ambio* **1992**, *21*, 274–277. [CrossRef]
22. Zhu, K.; Song, Y.; Qin, C. Forest age improves understanding of the global carbon sink. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3962–3964.
23. Birhane, E.; Ahmed, S.; Hailemariam, M.; Negash, M.; Rannestad, M.M.; Norgrove, L. Carbon stock and woody species diversity in homegarden agroforestry along an elevation gradient in southern Ethiopia. *Agrofor. Syst.* **2020**, *94*, 1099–1110. [CrossRef]
24. Dhyani, S.K.; Ram, A.; Newaj, R.; Handa, A.K.; Dev, I. Agroforestry for carbon sequestration in tropical India. In *Carbon Management in Tropical and Sub-Tropical Terrestrial Systems*; Springer: Berlin/Heidelberg, Germany, 2019. [CrossRef]
25. Horrillo, A.; Gaspar, P.; Escribano, M. Organic Farming as a Strategy to Reduce Carbon Footprint in Dehesa Agroecosystems: A Case Study Comparing Different Livestock Products. *Animals* **2020**, *10*, 162. [CrossRef]
26. Ortiz-Ceballos, G.C.; Vargas-Mendoza, M.; Ortiz-Ceballos, A.I.; Briseño, M.M.; Ortiz-Hernández, G. Aboveground Carbon Storage in Coffee Agroecosystems: The Case of the Central Region of the State of Veracruz in Mexico. *Agronomy* **2020**, *10*, 382. [CrossRef]
27. Sari, R.R.; Saputra, D.D.; Hairiah, K.; Rozendaal, D.M.A.; Roshetko, J.M.; van Noordwijk, M. Gendered Species Preferences Link Tree Diversity and Carbon Stocks in Cacao Agroforest in Southeast Sulawesi, Indonesia. *Land* **2020**, *9*, 108. [CrossRef]
28. PSanchez, P.A. Science in agroforestry. *Agrofor. Syst.* **1995**, *30*, 5–55. [CrossRef]
29. Pattanayak, S.K.; Mercer, D.E.; Sills, E.; Yang, J.-C. Taking stock of agroforestry adoption studies. *Agrofor. Syst.* **2003**, *57*, 173–186. [CrossRef]
30. Coe, R.; Sinclair, F.; Barrios, E. Scaling up agroforestry requires research 'in' rather than 'for' development. *Curr. Opin. Environ. Sustain.* **2014**, *6*, 73–77. [CrossRef]
31. Kwesiga, F.; Akinnifesi, F.K.; Mafongoya, P.L.; McDermott, M.H.; Agumya, A. Agroforestry research and development in southern Africa during the 1990s: Review and challenges ahead. *Agrofor. Syst.* **2003**, *59*, 173–186. [CrossRef]
32. Nair, P.K.R. Directions in tropical agroforestry research: Past, present, and future. In *Agroforestry Systems*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 223–245. [CrossRef]
33. Griffiths, T.L.; Steyvers, M. Finding scientific topics. *Proc. Natl. Acad. Sci. USA* **2004**, *101* (Suppl. 1), 5228–5235. [CrossRef]
34. Hornik, K.; Grün, B. Topicmodels: An R package for fitting topic models. *J. Stat. Softw.* **2011**, *40*, 1–30.
35. Geman, S.; Geman, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *PAMI-6*, 721–741. [CrossRef]
36. Villardon, M.P.G. *Una Alternativa de Representación Simultánea: HJ-Biplot*. Qüestiió: Quaderns d'Estadística i Investigació Operativa. 1986, pp. 13–23. Available online: https://www.raco.cat/index.php/Questiio/article/download/26468/26302 (accessed on 22 February 2023).
37. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022. [CrossRef]
38. Xiong, H.; Cheng, Y.; Zhao, W.; Liu, J. Analyzing scientific research topics in manufacturing field using a topic model. *Comput. Ind. Eng.* **2019**, *135*, 333–347. [CrossRef]
39. Harzing, A.-W.; Alakangas, S. Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics* **2016**, *106*, 787–804. [CrossRef]
40. De la Hoz-M, J.; Fernández-Gómez, M.J.; Mendes, S. LDAShiny: An R Package for Exploratory Review of Scientific Literature Based on a Bayesian Probabilistic Model and Machine Learning Tools. *Mathematics* **2021**, *9*, 1671. [CrossRef]
41. Team, R. *R Studio—Google Académico*; RStudio: Integrated Development for R. RStudio Inc.: Boston, MA, USA, 2018.
42. Blei, D.M.; Lafferty, J.D. A correlated topic model of Science. *Ann. Appl. Stat.* **2007**, *1*, 17–35. [CrossRef]
43. Röder, M.; Both, A.; Hinneburg, A. Exploring the space of topic coherence measures. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 399–408.
44. Chang, J.; Boyd-Graber, J.; Gerrish, S.; Wang, C.; Blei, D.M.; Chang, J.; Boyd-Graber, J.; Gerrish, S.; Wang, C.; Blei, D.M. Reading tea leaves: How humans interpret topic models. In *Advances in Neural Information Processing Systems 22—Proceedings of the 2009 Conference, Vancouver, BC, Canada, 7–10 December; 2009*; Neural Information Processing Systems Foundation: San Diego, CA, USA, 2009. Available online: https://proceedings.neurips.cc/paper/2009/hash/f92586a25bb3145facd64ab20fd554ff-Abstract.html (accessed on 22 February 2023).
45. Jones, T.; Doane, W. *TextmineR: Functions for Text Mining and Topic Modeling*, R package version 3.0.4; 2019. Available online: https://rdrr.io/cran/textmineR/ (accessed on 22 February 2023).
46. Lau, J.H.; Grieser, K.; Newman, D.; Baldwin, T. Automatic labelling of topic models. In Proceedings of the ACL-HLT 2011—Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011.

47. Van Eck, N.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef] [PubMed]
48. Vicente-Villardón. *MULTBIPLOT: A Package for Multivariate Analysis Using Biplots*; Departamento de Estadística, Universidad de Salmanca: Salamanca, Spain, 2010.
49. Liu, W.; Yao, S.; Wang, J.; Liu, M. Trends and Features of Agroforestry Research Based on Bibliometric Analysis. *Sustainability* **2019**, *11*, 3473. [CrossRef]
50. EURAF. Agroforestry as Sustainable Land Use. In Proceedings of the 4th European Agroforestry Conference, Nijmegen, The Netherlands, 28–30 May 2018.
51. Puri, S.; Nair, P. Agroforestry research for development in India: 25 years of experiences of a national program. *Agrofor. Syst.* **2004**, *61–62*, 437–452. [CrossRef]
52. Somarriba, E. Diagnóstico y diseño agroforestal. *Agrofor. Am.* **1998**, *5*, 68–72.
53. Cannell, M.G.R. Soil Productivity Aspects of Agroforestry. By P. K. R. Nair. Nairobi: International Council for Research in Agroforestry (1984), pp. 83, price unstated. *Exp. Agric.* **1985**, *21*, 299. [CrossRef]
54. Perfecto, I.; Vandermeer, J.H.; Bautista, G.L.; Nunñez, G.I.; Greenberg, R.; Bichier, P.; Langridge, S. Greater predation in shaded coffee farms: The role of resident neotropical birds. *Ecology* **2004**, *85*, 2677–2681. [CrossRef]
55. Klein, A.-M.; Steffan-Dewenter, I.; Tscharntke, T. Bee pollination and fruit set of Coffea arabica and C. canephora (Rubiaceae). *Am. J. Bot.* **2003**, *90*, 153–157. [CrossRef]
56. Ricketts, T.H.; Daily, G.C.; Ehrlich, P.R.; Michener, C.D. Economic value of tropical forest to coffee production. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12579–12582. [CrossRef] [PubMed]
57. Buresh, R.J.; Tian, G. Soil improvement by trees in sub-Saharan Africa. *Agrofor. Syst.* **1997**, *38*, 51–76. [CrossRef]
58. Young, A. Soil productivity, soil conservation and land evaluation. *Agrofor. Syst.* **1987**, *5*, 277–291. [CrossRef]