Precision Agriculture Under a Bibliometric View

Wanderson de Vasconcelos Rodrigues da Silva
Post-graduation Program in Intellectual Property Science (PPGPI/UFS),
Federal Institute of Education, Science and Technology of Piauí (IFPI),
Piripiri, Brazil.
ORCID: https://orcid.org/0000-0002-7631-9830
Email: wanderson.vasconcelos@ifpi.edu.br

Renata Silva-Mann
Post-graduation Program in Intellectual Property Science (PPGPI/UFS),
Department of Agronomic Engineering (DEA/UFS),
São Cristóvão, Brazil.
ORCID: https://orcid.org/0000-0001-5993-3161
Email: renatamann@academico.ufs.br

Abstract

Precision Agriculture comprises techniques to monitor and control the differentiated application of agricultural inputs, considering the variability of cultivation areas over time to increase productivity and maintain environmental sustainability. Its current form considers the use of high-tech equipment to ensure food safety in the future and, therefore, constantly seeks research that produces innovations for the sector. However, there is a tremendous challenge in evaluating scientific development, given the large volume of information. This study aimed to carry out a scientific mapping of Precision Agriculture from a set of bibliometric techniques supported using the R bibliometrix tool. Based on this objective, the research questions were formulated and answered throughout qualitative quantitative and descriptive exploratory study. The data processing resulted 5,807 articles (13,705 authors) obtained from 1993 to 2020. Among the main results, there is constant growth in the number of publications, especially between 2016 and 2020; more significant concentration among countries, forming well-defined collaboration subnetworks through their institutions; presence of expressive central themes in the research with a high density of studies, such as the use of remote sensing combined with machine learning techniques, due to the growing trend in the amount of processed data.

Keywords: smart agriculture, scientific mapping, bibliometrics, systematic review, indicators.

1 Introduction

The objective of this research was to investigate the state and evolution of Precision Agriculture using bibliometrics and scientometrics analysis. This article presents an expanded review of studies associated with Precision Agriculture, allowing the development of quantitative and qualitative scientometrics indicators to answer the following questions: (1) Studies on Precision Agriculture have grown in terms of scientific publication? (2) Which researchers, journals, and countries published the most on the topic? (3) What are the main topics of interest to the studies, and their focus? (4) What are the main intellectual and social structures regarding Precision Agriculture?

To answer the four questions, a scientific analysis was performed using the open-source tool bibliometrix, a statistical package program in the R language capable of executing the recommended workflow for performing bibliometry analyzes and performing comprehensive scientific mapping analyzes (Aria & Cuccurullo, 2017).

This article is structured in section that offers a review of the literature on Precision Agriculture to construct a contextualization of the theme, describing its main concepts and applications. The methodological details are discussed, and the main results addressing the four research questions mentioned; finally, the properties and limitations of the study, as well as the suggestion for future research.

2 Literature review

Agriculture is characterized as one of the oldest activities. It has been essential to guarantee the demands of society, whether for food, consumer goods, energy and other resources. However, natural factors have always been crucial in determining its performance. This and other challenges prompted humanity to develop techniques and tools that enabled greater control over crop productivity, thus redefining agricultural practice. Thus, agriculture is gradually developed, and it has been this ability to innovate that will ensure food security for future generations (Coelho & Silva, 2009).

With the advancement of Information Technologies (IT), agriculture began to unite the physical and the digital world, creating an infrastructure capable of collecting different data to apply treatment to crops with higher precision and, consequently, ensuring efficiency at all stages of the production chain (Bassoi et al., 2020). Precision Agriculture (PA) is defined as a management strategy that collects, processes and analyzes data on the temporal and spatial variability of the productive unit, adding information to support decision-making. As a result, obtaining an increase in production efficiency improves economic yield and reduces the environmental impact of agricultural activity (International Society for Precision Agriculture, 2021).

Precision Agriculture is part of Intelligent Agriculture or Agriculture 4.0, as it refers to the use of high-tech equipment, ranging from its hardware (physical components) to the most varied specialized software (computer programs) and telecommunication networks (transmission of data) (Precision Agriculture in the 21st Century, 1997). Therefore, several tools help to identify and control the variability of crop areas, such as satellite navigation systems, sensors, and remote sensing techniques, measuring devices for soil properties, variable rate input technologies, methods of information and artificial intelligence, among others (Bassoi et al., 2020; Cisternas et al., 2020).

The adoption of new technologies in agriculture constitutes a process of constant innovation in the sector, and the use of Precision Agriculture techniques adds advantages to the production process compared to the conventional agricultural model, promoting its preference and sustainability. However, each edaphoclimatic region requires unique crop management. This results in that, not every technology can be within every stage of the process and immediately implement every available tool that is seriously costly (Coelho & Silva, 2009). It is recommended to implement one or two technologies in phenological phases of the production cycle and the continuous monitoring of results. The most important strategy is not to define the phases that the farmer must start, but instead recognize that Precision Agriculture provides the application of correctives, fertilizers, pesticides, seeds, etc. in the rational amount, focusing the proper place and at the ideal time, to obtain the efficiency. The costs and the technologies applied in the process enable your adoption (Kent Shannon et al., 2018; Prado, 2018).

Therefore, Precision Agriculture (PA) is not limited to a well-defined set of tools. Still, it is a strategy that reconfigures itself as technological innovations, scientific research, and agricultural sector transformation. Hence, the need to monitor and evaluate the PA's current stage of scientific development presented against the backdrop of changes in agriculture. This study assessed Precision Agriculture based on articles of indexed journals by bibliometrics and scientometrics indicators.

3 Data and Methods

The scientific map comprises several methodological steps and usually requires software, which makes the process complex and costly. In addition, researchers appreciate both quantitative and qualitative approaches to understanding and organizing knowledge about prior investigations. In this sense, bibliometrics stands out as a process of systematic methodology of scientific production, since it uses mathematical and statistical methods to classify scientific activity (Esfahani et al., 2019; Pritchard, 1969).

Bibliometrics is an objective method to access data and classify techniques, especially using a volume of information. These metrics can portray the long time and developing knowledge of science indicators and technology, inferring trends, listing the most influential research, authors/researchers, and institutions, and carrying out diagnoses at different levels (Diodato & Gellatly, 2013). Bibliographic data can be made available and visualized by different tools depending on the type of analysis intended. To map and analyze the information, we used the bibliometrix package. It is a tool of scientometrics and bibliometrics. The data are performed in R from scientific databases as Scopus (https://www.scopus.com), Web of Science (https://www.webofknowledge.com), PubMed (https://pubmed.ncbi.nlm.nih.gov), Dimensions (https://www.dimensions.ai) and Cochrane (https://www.cochranelibrary.com).

It is programmed in an open-source statistical language named bibliometrix, with algorithms that can be revised and improved by the developer community, making it one of the most powerful and flexible statistical software environments (Aria & Cuccurullo, 2017; Dervis, 2019).

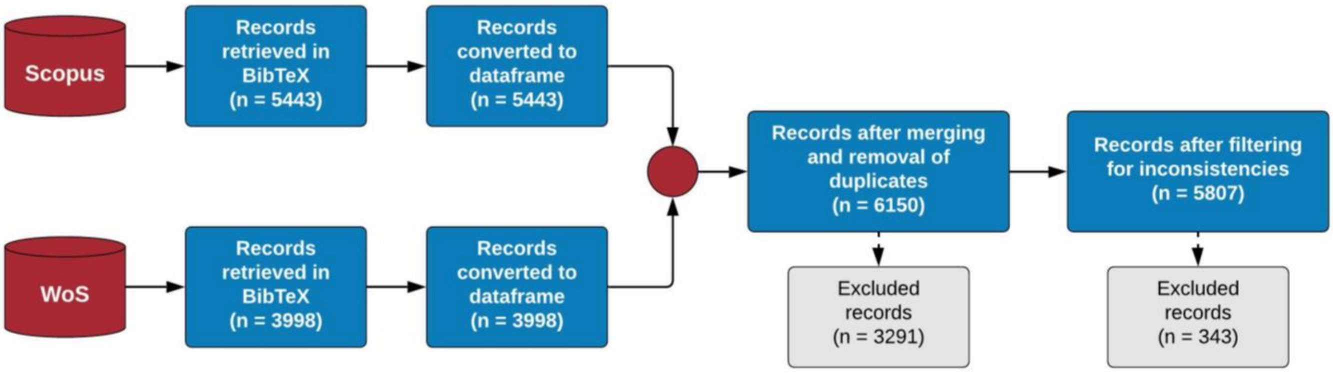
To answer the research questions, the scientific mapping workflow was divided into two major stages: (1) data collection and processing (section 3.1); and (2) analysis, visualization, and discussion (section 4).

3.1 Data collection and processing

The data collection and processing were divided into four substages (Figure 1). The first was the retrieval of information from two bibliographic databases: Scopus, from Elsevier; and Web of Science, maintained by Clarivate. Both are used to carry out scientific mappings in different areas of knowledge, either for individual studies or for comparison or construction of databases (Aria & Cuccurullo, 2017; Rodríguez-Soler et al., 2020).

Figure 1

Flow diagram for data collection and treatment



Source: the authors.

The data extraction definitions and guidelines should aid information recovery in the form of a protocol (Brereton et al., 2007); it was searched an advanced strategy developed, combining boolean logical operators as Precision Farming, in the title, abstract, and keywords of the documents (Table 1).

Table 1

Research criteria adopted in Scopus and WoS.

Scopus	Web of Science (WoS)
TITLE-ABS-KEY ("precision agriculture" OR "precision farming" OR "satellite farming" OR "satellite agriculture" OR "site specific crop management") AND DOCTYPE (ar) AND PUBYEAR < 2021	TI=("precision agriculture" OR "precision farming" OR "satellite farming" OR "satellite agriculture" OR "site specific crop management") OR AB=("precision agriculture" OR "precision farming" OR "satellite farming" OR "satellite agriculture" OR "site specific crop management") OR AK=("precision agriculture" OR "precision farming" OR "satellite farming" OR "satellite agriculture" OR "site specific crop management") AND DOCUMENT TYPE: (article) Defined time: 1945-2020

Source: the authors.

The survey was conducted on January 15, 2021, from articles published until 2020. WoS considers publications from 1945 onwards, which does not represent a limitation for the study, as the philosophy of Precision Agriculture, as currently known, dates to the 1980s (Molin et al., 2015).

In Scopus, there were 5,443 records identified, and 3,998 on the Web of Science. Data were exported from the databases in BibTeX (.bib), a file format that can be used to describe bibliographic data (Alkhateeb, 2010).

In the second substage, the BibTeX files extracted from the databases were individually loaded into the R Studio software, a programming environment in the R language used to work with the bibliometrix package. Then, both were converted into bibliographic Data Frames (data format suitable for working in the R bibliometrix tool).

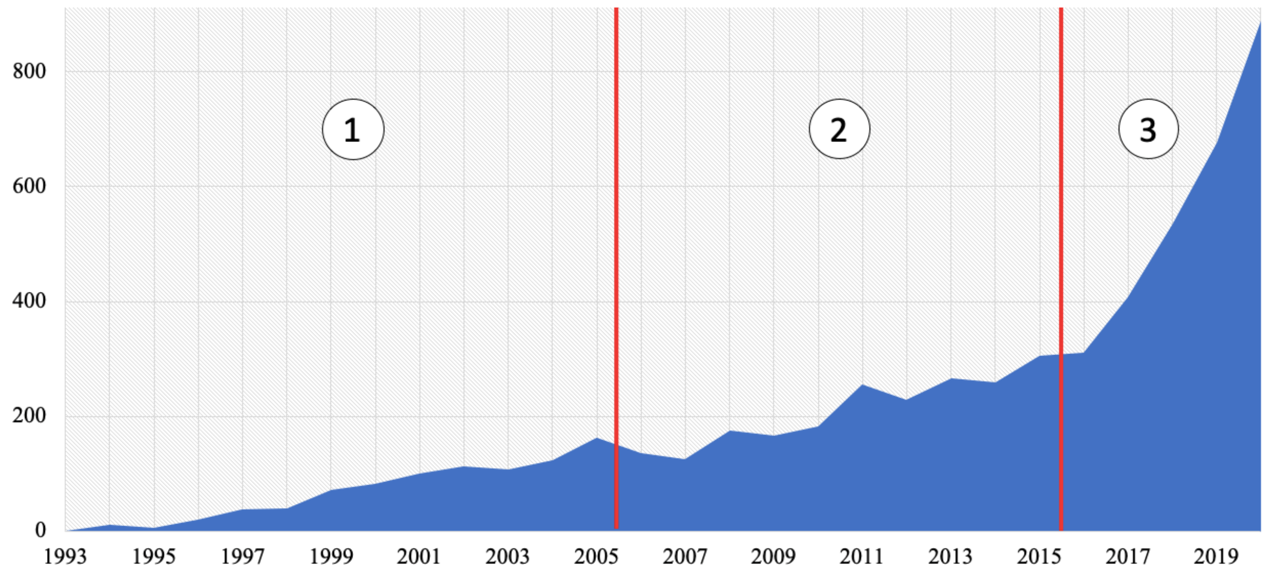
The two databases were overlapped into single frames in the third substage, and duplicate records were removed and reorganized. From a total of 9441 records, 3291 were excluded (Figure 1). Finally, in the last substage, data cleaning was performed. Aria and Cuccurullo (2017) highlight that bibliometrix does not have specific functions dedicated to data cleaning. However, most of the data made available by the databases are registered. Some methods can be reports to verify redundancies and inconsistencies. In this study, it was decided to perform a manual exclusion of 343 records.

4 Analysis, visualization, and discussion

In this study, 5,807 documents were published from 1993 to 2020 (28 years). Although the annual scientific production in Precision Agriculture (PA) has shown over the years increasing tendency, this evolution has not been constant. Approximately half of all articles were published in the last five years, an average of 564 articles per year. In previous years, the maximum did not exceed 305 documents.

Figure 2

Annual scientific production on PA divide into time intervals



Source: the authors.

Although a scientometry is performed to portray a static image of a research field, it can be divided into intervals to capture scenarios over time (Rodríguez-Soler et al., 2020). Thus, the trend lines of publications at different intervals were divided into three periods (Figure 2): Period 1 (1993-2005); Period 2 (2006-2015); and Period 3 (2016-2020).

A descriptive scenario of the study's database was articulated, considering the authors (in charge of Precision Agriculture), journals, and documents' theme. We evaluated periods 1, 2, and 3 independently and after the Total Period. The use of the bibliometrix R package allowed different techniques to extract conceptual, social and intellectual networks (section 4.2). To study the conceptual structure of Precision Agriculture and its evolution over time, a bibliometric study for periods 1, 2, 3, and the total was considered. When the period is not mentioned, it is regarded as the holy period.

4.1 Descriptive analysis of bibliographic data

Table 2 summarizes the main results of the descriptive data, based on the bibliometrix R package and tabulated later.

Table 2

Main information about the study data

Description	Period 1 1993-2005	Period 2 2006-2015	Period 3 2016-2020	Total 1993-2020
Years	13	10	5	28
Articles	885	2,102	2,820	5,807
Average of articles per year	68.08	210.2	564	207.39
Journal	220	559	800	1,227
Average citations per article	42.46	24.9	7.705	19.23
Keyword Plus	2,558	9,304	10,631	17,306
Author's Keywords	1,740	4,958	7,337	11,920
Authors	1,756	5,164	8,812	13,705
Author appearances	2,919	8,638	13,493	25,050
Authors of single-authored articles	106	106	109	309
Authors of multiple-authored articles	1,650	5,058	8,703	13,396
Articles per author	0.504	0.407	0.32	0.424
Authors per article	1.98	2.46	3.12	2.36
Co-authors per article	3.3	4.11	4.78	4.31
Collaboration index	2.16	2.54	3.22	2.45

Source: the authors.

The annual scientific production on Precision Agriculture (Fig. 2) and the main information (Table 2) increased over the years. The journal of The American Society of Agricultural and Biological published in 1993 the first article about PA. The main theme was soil property and its possible applications to estimate unsampled points in the field (Han et al., 1993). The publications in this journal grew at an annual percentage rate of 27.44%, reaching 889 documents in 2020.

The total number of authors summed 13,705; 309 presented a single author. The number of authors per article and co-authors grew for each period, totaling the average value of 2.36 and 4.31, respectively. On the other hand, the number of articles per author decreased. Consequently, the increased Collaboration Index (CI) remained stable in the three previous periods (Table 2).

Table 3

Most relevant authors in Precision Agriculture

	Productivity		Fractionated production	
Classification	Authors	Articles	Authors	Articles
1	Li M	50	Yang C	14.13
2	Li Y	48	Sudduth K	12.30
3	Sudduth K	48	Molin J	12.27
4	Wang Y	47	Li M	11.54
5	Wang X	46	He Y	9.79
6	Wang J	45	Wang Y	9.41
7	He Y	43	Wang X	9.25
8	Li X	43	Li Y	9.25
9	Zhang X	41	Wang J	9.11
10	Molin J	39	Shearer S	8.94

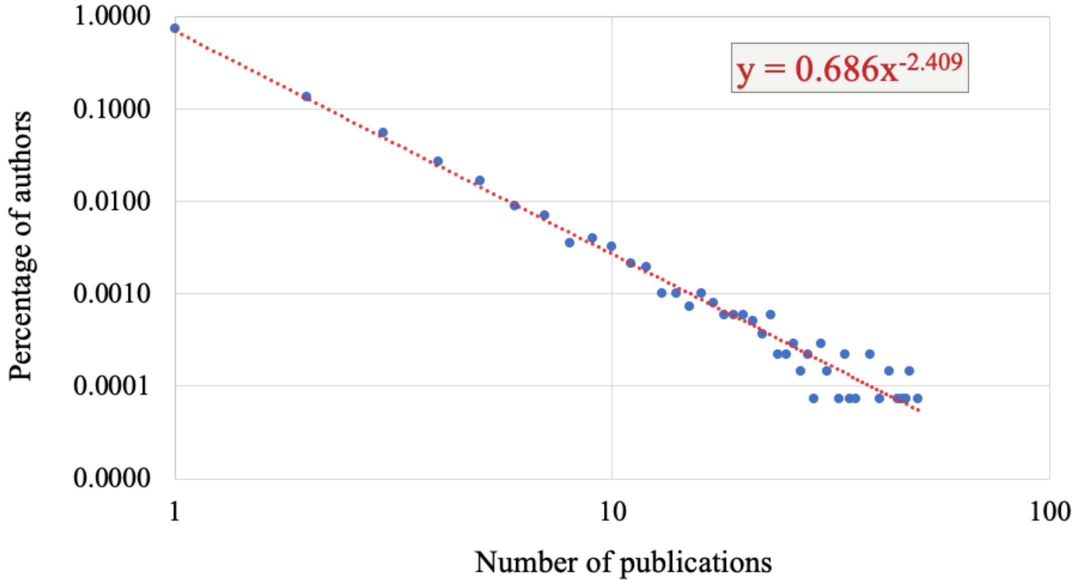
Source: the authors.

To identify the most relevant authors, we considered indicators mentioned in Table 3, as the number of articles by author and fractional production. The ten most writers in the area produced at least 39 articles, with author Li M in the first position with 50 works. Concerning fractional production, which indicates each of the authors and co-authors are credited with a fraction of the total contribution of the articles (Urbizagástegui Alvarado, 2002), author Yang C stands out with 14.13 works. This same author was not found among the top ten products, demonstrating his preference for working with few co-authors or individuals. On the other hand, the ten most productive researchers participate in works with an average of 4.7 authors per article, above the average considering all articles in this study (2.36 authors / article).

Using the bibliometrix, the frequency of the authors and the applicability of Lotka's Law (Lotka, 1926) were carried out. Lotka's law is a classical law of bibliometrics that states the number of authors who published n articles on a given topic; that is approximately $1/n^2$ compared to those published once. In their class, it is around 60% (Urbizagástegui Alvarado, 2008). In general, Lotka's Law can be expressed as $y = c * x^{-\alpha}$, where c and α are defined according to the observed data for each scientific field and y is the percentage of authors with x publications.

Figure 3

Lotka's Law of Scientific Productivity for Precision Agriculture



Source: the authors.

The data captured by bibliometrix were tabulated in Microsoft Excel, and the scatter plot built in logarithmic scale, insertion of the power trend line, and its equation (Fig. 3). The constant (c = 0.686, e.g., 68.6%) was estimated to represent the theoretical percentage of researchers who contributed with only one article. A = 2.4, that is, considering the total period of 28 years coverage of this study, as the community of researchers in Precision Agriculture (Potter, 1981). The researchers' productivity approaches the parameters observed by Lotka (c = 0.6 and α = 2).

Table 4

Most cited articles globally and locally

Classification	Article	Journal	TC	LC
1	Haboudane D (2004)	Remote Sens Environ	1,309	11
2	Viscarra Rossel R A (2006)	Geoderma	1,149	0
3	Haboudane D (2002)	Remote Sens Environ	1,034	9
4	Cassman K G (1999)	Proc Natl Acad Sci USA	862	0
5	Zhang C (2012)	Precis Agric	695	0
6	Di H J (2002)	Nutr Cycl Agroecosyst	652	0
7	Mulla D J (2013)	Biosyst Eng	582	13
8	Corwin D L (2005)	Comput Electron Agric	559	0
9	Gitelson A A (2004)	J Plant Physiol	528	1
10	Slaughter D C (2008)	Comput Electron Agric	443	0

Source: the authors.

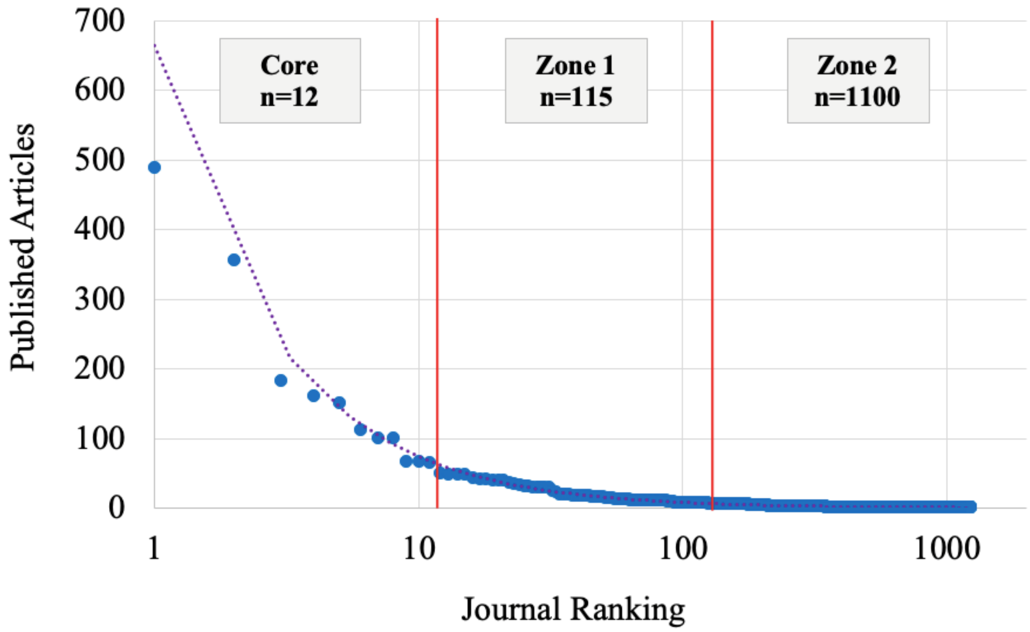
The most cited documents (TC) show an increased number of citations for each period (1, 2, and 3). The most cited (Haboudane et al., 2004) was published in Remote Sensing of Environment. The theme was new methods to improve hyperspectral remote sensing prediction of the Leaf Area Index (LAI) in Precision Farming. The author was the third most cited (Haboudane et al., 2002), published in the same journal in 2002, and studied crop chlorophyll indices' prediction through remote sensing.

Likewise, the second most cited article (Viscarra Rossel et al., 2006), published in Geoderma, also working with remote sensing techniques to improve soil analysis using Diffuse Reflectance Spectroscopy (DRS). The three articles have more than 1,000 citations each. On the other hand, the difference between total citations (TC) and local citations (LC) is remarkable and reflects the number of citations cited by another article in the database. This fact demonstrates that the articles claim respect for other topics in addition to Precision Agriculture (PA) and, due to the difference in the values of the TC and LC indicators, they have a greater relationship with matters external to the PA.

Although the journals mentioned accumulate many citations, these are not among the ten most productive journals in the PA area. The number of journals published on the theme was in Table 2, and the annual productivity rate of the journals on PA. To measure the journal productivity and define their research for an area of Precision Agriculture, the second law of bibliometrics was used. The Bradford's Law or Law of Scientific Knowledge Dispersion (Bradford, 1934) consists of placing periodicals in descending order of productivity on a given subject, distinguishing them in a nucleus of the journals more dedicated to the theme and in different areas with the same number of articles in each nucleus (Pinheiro, 1983).

Figure 4

Bradford's Law of scientific productivity of journals



Source: the authors.

Figure 4 presents descriptive data on the production of journals extracted from the bibliometrix. For better visualization, the ranking was found on a logarithmic scale. The ratio in a zone, and the amount in the previous zone is called Bradford Multiplier (k), and the number of journals in the core is defined by r_0 (Figure 5). The value of k = 9.5 complete in the study described the core with the 12 most relevant journals for Precision Agriculture out of a total of 1227, totaling together a third of all production (1908 articles). Some of the top journals are Computers and Electronics in Agriculture (489 articles - 8.42%), Precision Agriculture (356 articles - 6.13%) and Chinese Agricultural Engineering Society Transactions (184 articles - 3.17%).

Figure 5

Mathematical formulation of Bradford's Law.

$$k = (1{,}781 * Y_m)^{\frac{1}{P}}$$

$$r_0 = \frac{T * (k - 1)}{(k^P - 1)}$$

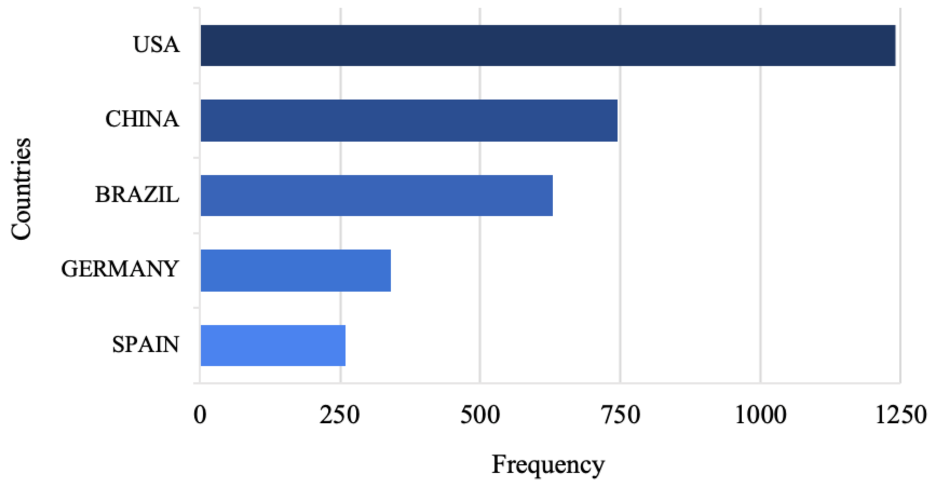
Y_m = Maximum productivity
P = Number of zones
T = Total number of journals

Note. Adapted from Egghe (1990).

One of the other essential areas of scientometric mapping is the countries where research was carried out. Figure 6 shows the five countries with the highest frequency in the publications involving Precision Agriculture. The United States of America stands out with 1240 articles (21.4%), China with 745 (12.8%), and Brazil with 630 (10.9%). The three countries concentrate about 45% of all documents on PA, while 55% are distributed in more than 90 countries.

Figure 6

Scientific production of the main countries (based on the affiliation of the first author)



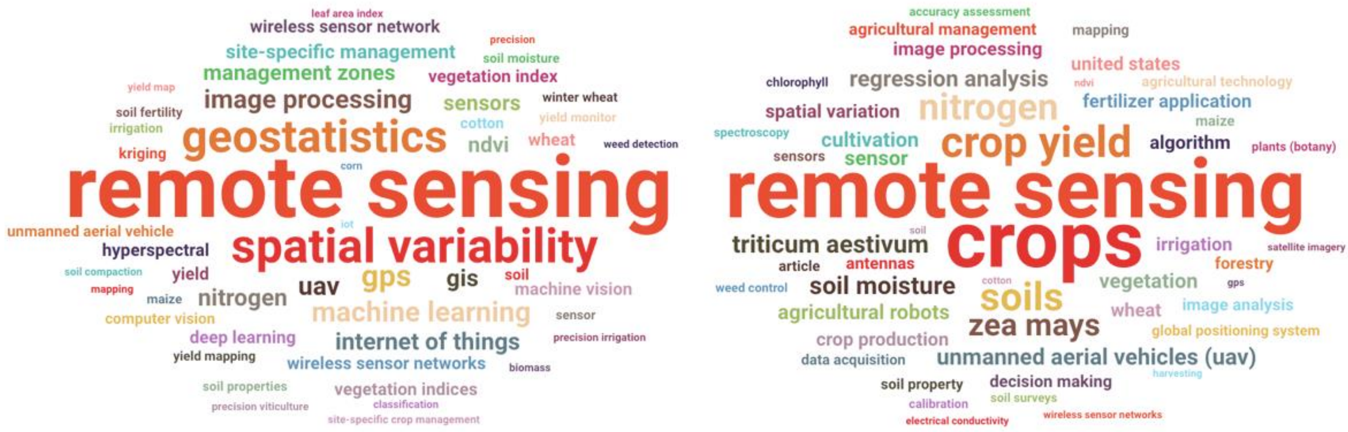
Source: the authors.

Figure 7 presents word clouds of the 50 most used keywords by authors (A) and attributed by the databases (B) (adapted from the descriptive analysis of bibliometrix R). The keywords "precision agriculture" and "agriculture" were excluded in the word clouds. According to Table 2, in the 5807 articles published, 11,920 different terms provided by the researchers were identified and 17,306 associated with the Scopus or WoS database, also defined as Keywords Plus.

In both groups of words, the term "remote sensing" was the most used. No wonder it was the main topic of interest of the most cited articles in the field of Precision Agriculture, as shown in Table 4. Although both groups reveal similar research trends, as the author's keywords tend to identify the content more specific, while Keywords Plus emphasize methods and techniques, as observed in other research fields (Xie et al., 2008; Zhang et al., 2016).

Figure 7

Most popular keywords used in studies associated with Precision Farming



(A) Author's keywords **(B) Keywords Plus**

Source: the authors.

The frequency distribution of keywords confirms the third law of bibliometrics, Zipf's Law (Zipf, 1949), which contacts a direct relationship between the frequency of occurrence y of a term in a text and its order in the order of frequencies as follows: $y = k / n$, where k is a constant (Cassettari et al., 2015). However, it is worth noting that the law became applicable to this bibliometric study when considering as Keywords Plus, for $n < 11$ and $k \cong 2509$.

4.2 Conceptual, intellectual and social structures

Strategic diagrams and thematic evolution

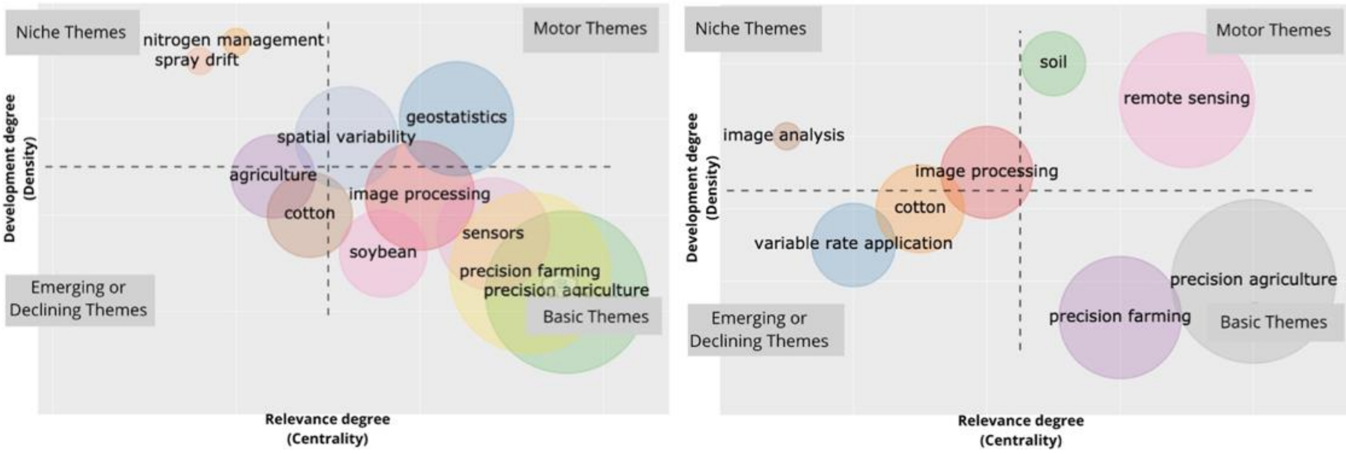
To study a conceptual structure of a research field, a bibliometric technique of keyword co-occurrence analysis allows identifying the main concepts dealt with by the research field, the evolution of these themes, and how trends, since this statistical method in accordance with the content of documents with greater precision and allows the use of different visualization techniques, including strategic diagrams (Aria & Cuccurullo, 2017; Cobo et al., 2011).

For this study, the author's keywords were adopted as the unit of analysis. They selected descriptors that the researchers believe in representing the content of their articles in a more specific way. The analysis also takes into account the different time intervals within the study to assess thematic evolution: period 1

(1993-2005), period 2 (2006-2015) and period 3 (2016-2020). The R Bibliometrix package changes the mapping of the conceptual structure of research in the Precision Agriculture (PA) area through the creation of thematic maps, defined by Callon, Courtial and Laville (1991) as strategic diagrams for each period analyzed, as shown in Figures 8 and 9.

Figure 8

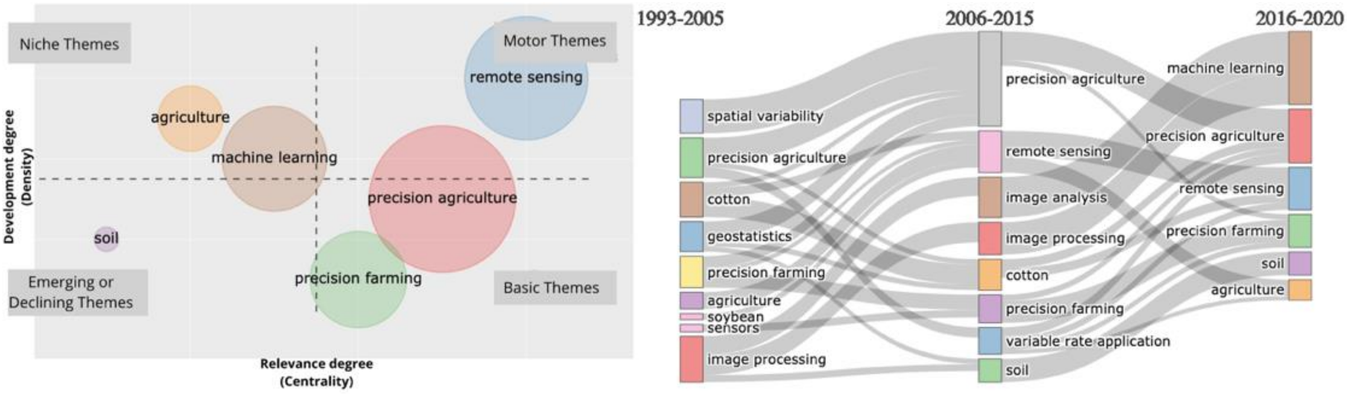
Strategy diagram for period 1 (left) and 2 (right)



Source: the authors.

Figure 9

Strategic diagram for period 3 (left) and thematic evolution (right)



Source: the authors.

The development of the three strategic diagrams and the common thematic evolution map a dynamic analysis to understand the evolution of the research field in this study. The themes represented in the images represent the keyword clusters of the co-occurrence analysis. Thus, Precision Agriculture is understood as a set of research themes mapped in a two-dimensional space. The cluster label identifies the most needed keyword in the corresponding theme, and the cluster size is proportional to the words it contains. Given that a conceptual structure of Precision Agriculture is investigated, the themes "precision agriculture" and "precision farming" have greater centrality, that is, they are general and transversal to the different areas of investigation in the research field.

Regarding period 1, from 1993 to 2005 (Figure 8, left), there are other transversal themes such as “image processing”, which became more developed and specialized in the following period. In turn, geostatistics was the most relevant driving theme of the literature at this stage. No wonder, it allowed the structuring of Precision Agriculture (PA), connecting to different concepts applicable to other themes. The disappearance of the same following period is due to its evolution, since it became an intrinsic concept to PA. Some specialized themes were well developed in the first period, such as the case of “nitrogen management” and “spray drift”, both related to the rationing of agricultural resources and minimization of environmental damage.

In the second period, from 2006 to 2015 (Figure 8, right), “remote sensing” becomes the most important topic for Precision Agriculture (PA). This reinforces the data on the most cited articles (Table 4) published in the previous period (Haboudane et al., 2002, 2004; Viscarra Rossel et al., 2006) studies on different remote sensing techniques applied to PA. An emerging theme that was later incorporated into the concepts of precision in agriculture for the “variable rate application” that relates to the application of agricultural inputs (water, fertilizers, pesticides, among others), reinforcing the soil theme that appears in this period with a lot of density, but in the last period analyzed, from 2016 to 2020 (Figure 9, on the left), it becomes a theme in decline (marginalized).

In this last phase, there is a reduction in the number of themes. The concept of “image analysis,” which emerged in the earlier period as a specialized theme of “image processing”, was absorbed along with other concepts (“machine vision”, “artificial intelligence”, “unmanned aerial vehicle”, “IoT”, etc.) for the “machine learning” cluster, a theme associated with Industry 4.0. As it is a theme in the field of Information Technology (IT) research, it has a low centrality in this study, but with a degree of development (density).

Finally, the strategic diagrams and the conceptual evolution map (Figure 9, right) reveal that the main themes are different over time. Except for the concepts that describe Precision Farming own research field, most terms appear only once as the most significant keyword in the cluster which evolve to new associates. The most solid themes identified in the evolution were those with high inclusion rates among the associated themes (“machine learning”, “precision farming” and “remote sensing”), which may indicate research trends in the short term.

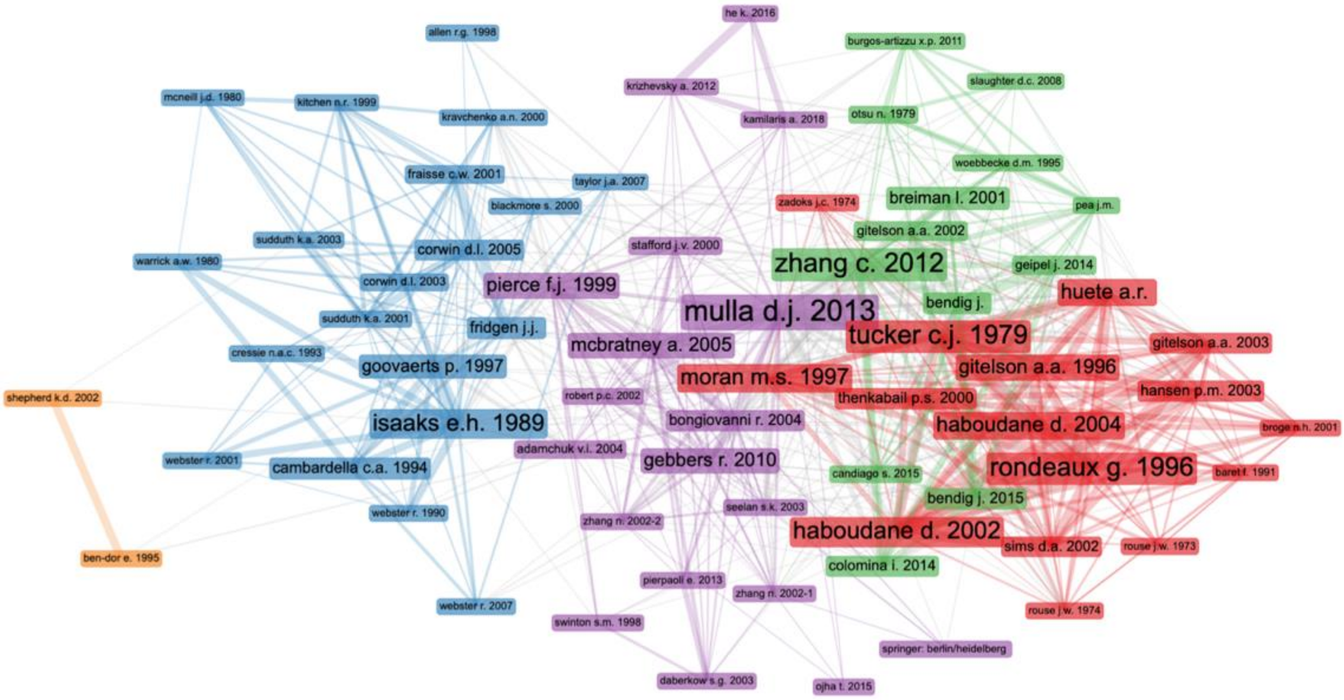
Intellectual Structure of Precision Agriculture

The intellectual structure of a research field allows us to investigate how a scientific community influences an author's work, as evidenced by the creation of co-citation networks. This is due to the relationships of content and associated meanings between articles published in a specific scientific domain and the frequency with which other authors simultaneously cite them through the co-occurrence of references to authors, documents, journals, etc. This knowledge framework strengthens the results of the conceptual framework and is important not only for research, but also for policy and practice creation (Aria & Cuccurullo, 2017; Grácio, 2016).

For this study, an analysis of document co-citation was adopted to determine which articles are more directed to the research field of Precision Agriculture since this is an effective unit of analysis to understand the intellectual structure in science and other areas (Small, 1973; White & Griffith, 1981).

Figure 10

Co-citation network for articles related to Precision Agriculture



Source: the authors.

In Figure 10, an intellectual structure of a network for the co-citation of the most relevant research in the field of Precision Agriculture is presented, obtained through the biblioshiny application, a web interface for the bibliometrix R package, based on the collected data. The network is formed by 5 clusters numbered from left to right: cluster 1 (orange); set 2 (blue); cluster 3 (violet); cluster 4 (green); cluster 5 (red).

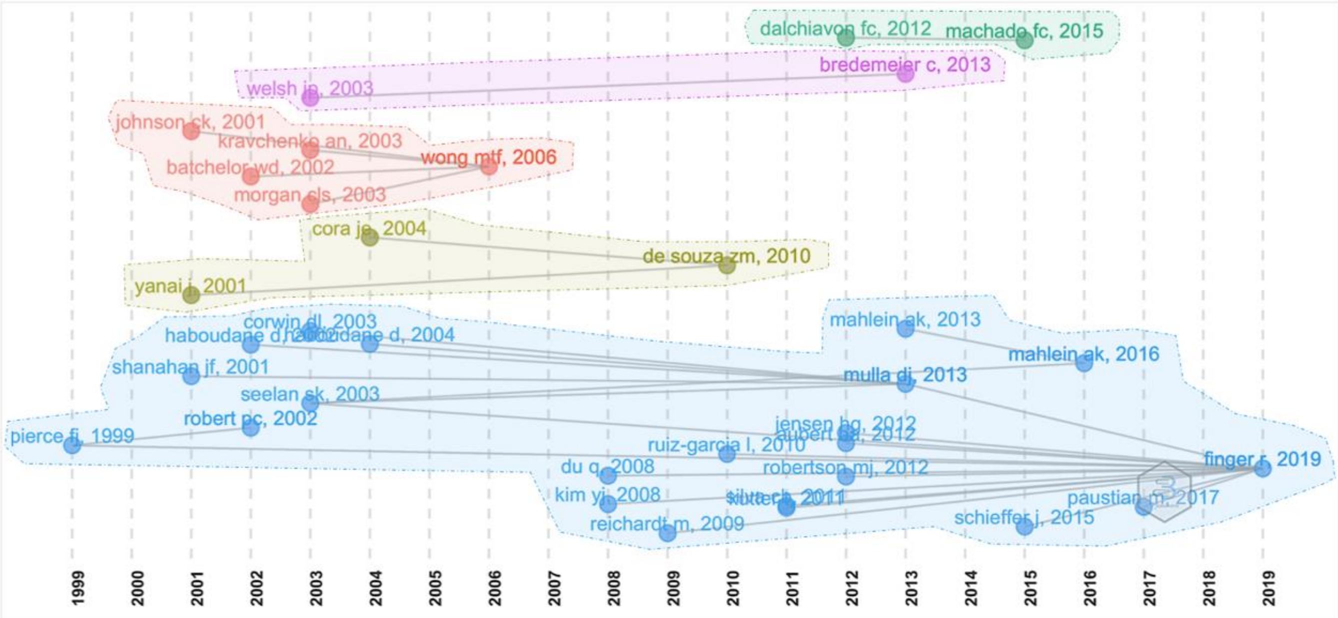
Cluster 1 is one of the most influential workgroups in the area. That is, the most cited by document and is associated with the characterization and analysis of the spectral properties of the soil. Cluster 2 comprises theory and application of geostatistics, scale distributions, spatial trends and field zone management.

Cluster 3 refers directly to Precision Agriculture, relating to general and specific aspects, such as sustainability and food safety. Cluster 4 studies on machine learning, image analysis and processing, with data captured from unmanned aerial vehicles (UAVs). This is a theme that has emerged in Precision Agriculture over the last decade, as no conceptual evolution map has ever seen (Figure 9, right). For example, cluster 5 encompasses remote sensing based on aircraft or satellites for optimizing vegetation indices, analysis of reflectance spectra, among others used in agricultural crops to monitor the crop.

The intellectual structure can also be represented by a chronological map of the most relevant direct citations within the study's data collection (Garfield, 2004). Thus, historiographical mapping (Figure 11) has 33 of the most cited articles in 5807 documents.

Figure 11

Network of direct historical quotes on Precision Agriculture



Source: the authors.

Since 1993 have been published articles on Precision Agriculture. However, the relevant majority was published from 1999 to 2019, with a predominant sub-network of 21 articles. It is important to emphasize that the focus of the historiographical mapping is not on the authors, but in the topics discussed. The strength and influence of this dominant sub-network are due to the authors' commitment to approach Precision Agriculture in its essence, discussing concepts, aspects, sustainability, and advances. The other subnets are more focused on specific themes and applied research.

Social structure and collaboration networks

A social structure of how authors, institutions, and countries interact within scientific research generates an extensive collection of bibliographic data. It is possible to apply a co-authorship analysis technique to identify research groups, authors, institutions, unknown communities of researchers, peers from collaborating countries, besides social and intellectual research groups (Aria & Cuccurullo, 2017; Peters & Van Raan, 1991).

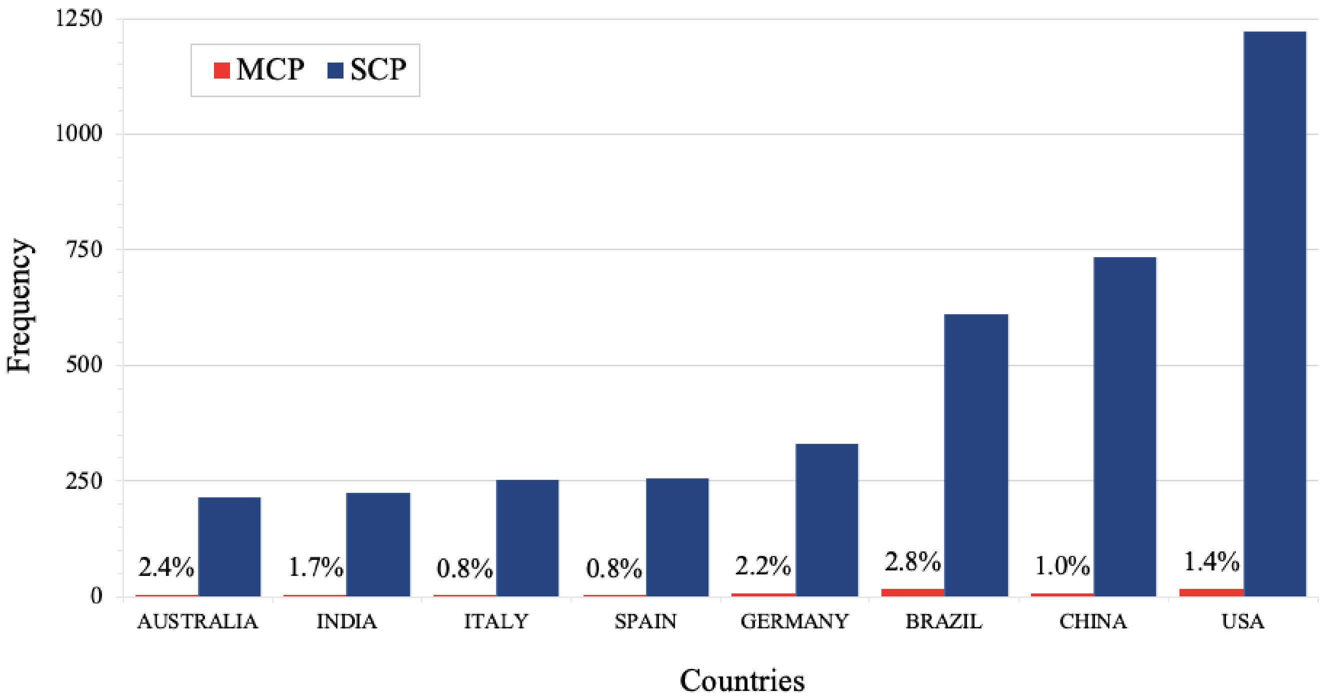
The descriptive analysis of the bibliographic data presented at the beginning of this study showed that 13705 authors from 98 countries contributed to the publication of 5807 articles associated with Precision Agriculture in the 28-year interval (1993 to 2020). Also, noteworthy is the large percentage of automatic collaboration, collaboration with other researchers (97.7%), and collaboration rate increased with each period analyzed (Table 2). However, this collaboration is predominantly done within national borders.

Two metrics are analyzed in this cross-country collaboration analysis: Multi-Country Publications (MCP) and Single-Country Publications (SCP). The MCP measures the number of publications for which at least one co-author is from a different country (international collaboration). On the other hand, the SCP quantifies publications written by authors from the same country (national collaboration). The collaboration data in the Precision Agriculture and countries were presented in Figure 12.

The countries with the most significant scientific production associated with Precision Agriculture (PA) do not necessarily carry out the most international collaborations. Figure 12 indicates the levels of openness in values, estimated by the ratio between the MCP and the number of publications in each country. Researchers from the United States, for example, have publications with several countries and present an absolute number of collaborations greater than the others, however their proportional values of partnerships are smaller. On the other hand, Brazil stands out in the absolute and relative values of partnerships, but it is worth noting that the percentage values of international collaboration in PA as a whole are low compared to other fields of research.

Figure 12

Type of collaboration of countries that publish the most on Precision Agriculture



Source: the authors.

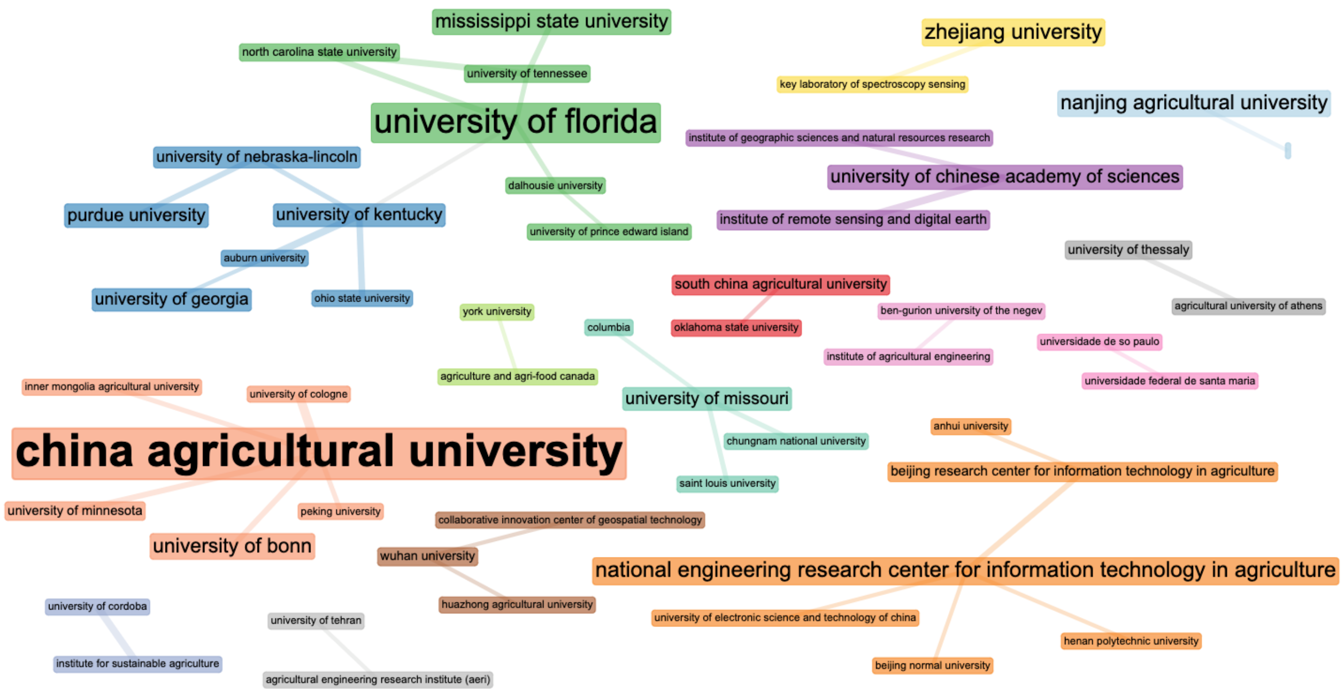
Concerning the affiliations researchers, the co-occurrence of institutions identified subnetworks of collaboration, revealing inter-institutional research groups. The purpose of the analysis is to focus on the relationships between the main institutions. They were eliminated as affiliations with less than five collaborative publications, resulting in 52 institutions in 16 research clusters in Precision Agriculture (Figure 13).

A fact that stands out when creating a collaboration network with at least five interactions between the affiliations is the formation of well-defined clusters and almost no peer interaction with other clusters. In the descriptive analysis, the institutions are among those that more published Precision Agriculture.

It was also found that some institutions maintain partnerships for research in the area with institutions from another country. This has become a bridge for knowledge traffic or even technology transfer to other institutions in their country, creating new sub-networks for internal collaboration.

Figure 13

Precision Agriculture social structure (network of collaboration between institutions)



Source: the authors.

China Agricultural University is the most presented institution (165) and has a robust collaborative network inter China, the United States, and Germany. The second institution that publishes the most in the area is the University of Florida with 129 articles and has a collaboration network with American and Canadian universities. In addition to universities, affiliations of government laboratories, research centers and institutes also make up the study groups. These partnerships are essential, especially for applied research in Precision Agriculture that demand the use of both natural and technological resources.

5 Conclusion

This study adopted several bibliometric techniques applied at different levels of analysis to investigate the performance of scientific production in Precision Agriculture (PA) research. Also, available is to examine the evolution of their cognitive structure, enabling the discovery of essential knowledge related to their thematic areas and behavior of collaboration networks through the mapping of their conceptual, intellectual and social structures, offering an overview of the literature.

The R Studio software and an R bibliometrix tool were used to load, analyze and visualize the bibliographic data merged from information retrieved from two databases commonly used by researchers: Scopus and Web of Science.

In response to the research questions in the introduction to this study (section 1), it was determined in section 4 that: (1) the topic of Precision Farming has increasingly attracted the attention of researchers in recent years, especially between 2016 and 2020, with a significant increase in the number of authors, articles, journals and citations. Throughout the analyzed period (28 years), publications grew at an annual rate of 27.44%, totaling 5807 documents. Also, noteworthy is the growth in the research collaboration index, that is, the average number of authorships per published article.

Responding to the second question (2), among the five authors who published the most in Precision Agriculture, most are of Chinese nationality and institutions, considering the amount needed by the author. On the other hand, when considering a fractional production, that is, when each of the authors and co-authors are credited with a fraction of the total contribution of the articles, other authors gain prominent spaces. By the applicability of Lotka's Law to the study, considering the total period of 28 years, a research community in PA is defined. With respect to countries, China, the United States and Brazil account for 45% of all articles published in the area. In terms of periodicals, Bradford's bibliometric law was revealed as the 12 most relevant sources, concentrating a third of the total number of publications, among which the journal with the title of Precision Agriculture.

To answer question (3), a survey of the author's keywords and the most frequent Keywords Plus in the researches was formulated, revealing the same “remote sensing”, “harvest”, “spatial variability,” and “geostatistics” among the most used. To complement the study, a co-occurrence analysis of author keywords was performed to create strategic diagrams and thematic evolution map, introducing the use of machine learning techniques combined with remote sensing as a growing trend in recent years and with enough density studies.

The last research question (4) concerns the intellectual and social structures of Precision Agriculture. Thus, an analysis of the co-citation of documents and co-authorship was carried out for the creation of co-citation and collaboration networks, respectively. Additionally, other views were created to better understand this information. There was an intense network of co-citation of articles that provided intellectual support for the development of the research field. This structure confirms the PA thematic map and how new themes emerged. The social structure, on the other hand, identifies strongly defined inter-institutional and international research groups, despite the predominance of collaborations in national research.

Finally, it is concluded that a qualitative review throughout the analysis is consistent with the quantitative results found and that among the main contributions of this study is the identification of consistent research axes in Precision Agriculture in the long term and the main research trends which can be useful for future studies and development of new technologies. Therefore, the result of this work can help both academics and professionals and companies in the sector.

However, this study is not without limitations. Although most of the data provided by the databases are, some records have empty, incomplete, or inconsistent fields. This situation is aggravated by merging of two bases that need to be consolidated into common fields. The tools used streamlined this process, but they do not have specific functions dedicated to data cleaning, which caused manual work and, in the case of large volumes of data, can be an expensive process. Another observed fact is that in Precision Agriculture, emerging themes have emerged quickly. This fact demands, in a few years, new bibliometric studies.

On the other hand, the limitations listed here are opportunities for further studies, specifying the results of this research by adding new databases, combining the use of complementary tools or the development of this to circumvent limitations of the tools adopted here, further optimizing the Scientific mapping Precision Agriculture and other areas.

References

Alkhateeb, F. (2010). BibTeX document generating using semantic web technologies. *Proceedings of the 1st International Conference on Intelligent Semantic Web-Services and Applications - ISWSA '10*, 1–6. https://doi.org/10.1145/1874590.1874592

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Bassoi, L. H., Inamasu, R. Y., Bernardi, A. C. de C., Vaz, C. M. P., Speranza, E. A., & Cruvinel, P. E. (2020). Agricultura de precisão e agricultura digital. *TECCOGS: Revista Digital de Tecnologias Cognitivas*, *20*, 17–36. https://doi.org/10.23925/1984-3585.2019i20p17-36

Bradford, C. S. (1934). Sources of information on specific subjects. *Engineering*, *137*, 85–86. https://doi.org/10.1177/016555158501000407

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155–205. https://doi.org/10.1007/BF02019280

Cassettari, R. R. B., Pinto, A.-L., Rodrigues, R. S., & Santos, L. S. dos. (2015). Comparação da Lei de Zipf em conteúdos textuais e discursos orais. *El Profesional de la Información*, *24*(2), 157. https://doi.org/10.3145/epi.2015.mar.09

Cisternas, I., Velásquez, I., Caro, A., & Rodríguez, A. (2020). Systematic literature review of implementations of precision agriculture. *Computers and Electronics in Agriculture*, *176*(July), 105626. https://doi.org/10.1016/j.compag.2020.105626

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, *5*(1), 146–166. https://doi.org/10.1016/j.joi.2010.10.002

Coelho, J. P. C., & Silva, J. R. M. (2009). *Agricultura de Precisão*. Associação dos Jovens Agricultores de Portugal.

Dervis, H. (2019). Bibliometric Analysis using Bibliometrix an R Package. *Journal of Scientometric Research*, *8*(3), 156–160. https://doi.org/10.5530/jscires.8.3.32

Diodato, V. P., & Gellatly, P. (2013). *Dictionary of Bibliometrics*. Routledge. https://doi.org/10.4324/9780203714133

Esfahani, H. J., Tavasoli, K., & Jabbarzadeh, A. (2019). Big data and social media: A scientometrics analysis. *International Journal of Data and Network Science*, *3*(3), 145–164. https://doi.org/10.5267/j.ijdns.2019.2.007

Garfield, E. (2004). Historiographic Mapping of Knowledge Domains Literature. *Journal of Information Science*, *30*(2), 119–145. https://doi.org/10.1177/0165551504042802

Grácio, M. C. C. (2016). Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, *21*(47), 82. https://doi.org/10.5007/1518-2924.2016v21n47p82

Haboudane, D., Miller, J. R., Elizabeth, P., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337–352. https://doi.org/10.1016/j.rse.2003.12.013

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, *81*(2–3), 416–426. https://doi.org/10.1016/S0034-4257(02)00018-4

Han, S., Goering, C. E., Cahn, M. D., & Hummel, J. W. (1993). A Robust Method for Estimating Soil Properties in Unsampled Cells. *Transactions of the ASAE*, *36*(5), 1363–1368. https://doi.org/10.13031/2013.28471

International Society for Precision Agriculture. (2021). *Precision Ag Definition*. ISPA. https://www.ispag.org

Kent Shannon, D., Clay, D. E., & Sudduth, K. A. (2018). *An Introduction to Precision Agriculture* (p. 1–12). https://doi.org/10.2134/precisionagbasics.2016.0084

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

Molin, J. P., Amaral, L. R., & Colaço, A. F. (2015). *Agricultura de Precisão* (1º ed). Oficina de Textos.

Peters, H. P. F., & Van Raan, A. F. J. (1991). Structuring scientific activities by co-author analysis. *Scientometrics*, *20*(1), 235–255. https://doi.org/10.1007/BF02018157

Pinheiro, L. V. R. (1983). Lei de Bradford: uma reformulação conceitual. *Ciência da Informação*, *12*(2), 59–80.

Potter, W. G. (1981). Lotka´s law revisited. *Library Trends*, *30*(1), 21–39.

Prado, H. (2018). Precisão na agricultura. *Revista Fonte: Tecnologia da Informação na Gestão Pública*, *15*(20), 48–46.

Precision Agriculture in the 21st Century. (1997). *Precision Agriculture in the 21st Century*. National Academies Press. https://doi.org/10.17226/5491

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation*, *25*(4), 348.

Rodríguez-Soler, R., Uribe-Toril, J., & De Pablo Valenciano, J. (2020). Worldwide trends in the scientific production on rural depopulation, a bibliometric analysis using bibliometrix R-tool. *Land Use Policy*, *97*, 104787. https://doi.org/10.1016/j.landusepol.2020.104787

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, *24*(4), 265–269. https://doi.org/10.1002/asi.4630240406

Urbizagástegui Alvarado, R. (2002). A Lei de Lotka na bibliometria brasileira. *Ciência da Informação*, *31*(2), 14–20. https://doi.org/10.1590/S0100-19652002000200002

Urbizagástegui Alvarado, R. (2008). A produtividade dos autores sobre a Lei de Lotka. *Ciência da Informação*, *37*(2), 87–102.

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, *131*(1–2), 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007

White, H. D., & Griffith, B. C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, *32*(3), 163–171. https://doi.org/10.1002/asi.4630320302

Xie, S., Zhang, J., & Ho, Y.-S. (2008). Assessment of world aerosol research trends by bibliometric analysis. *Scientometrics*, *77*(1), 113–130. https://doi.org/10.1007/s11192-007-1928-0

Zhang, J., Yu, Q., Zheng, F., Long, C., Lu, Z., & Duan, Z. (2016). Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. *Journal of the Association for Information Science and Technology*, *67*(4), 967–972. https://doi.org/10.1002/asi.23437

Zipf, G.-K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.