

The Impacts of Bibliometrics Measurement in the Scientific Community A Statistical Analysis of Multiple Case Studies

Vincenzo Basile¹, Massimiliano Giacalone², Paolo Carmelo Cozzucoli³

¹ Assistant Professor in Economics and Business Management, Federico II University of Naples, Italy

² Assistant Professor in Statistics, Federico II University of Naples, Italy

³ Associate Professor in Statistics, University of Calabria, Italy

Correspondence: Vincenzo Basile, Assistant Professor in Economics and Business Management, Federico II University of Naples, Italy

Received: April 14, 2022 Accepted: July 5, 2022 Online Published: July 12, 2022

doi:10.5539/res.v14n3p10

URL: <https://doi.org/10.5539/res.v14n3p10>

Abstract

In recent years, statistical methods such as bibliometrics have increasingly intensified to analyse books, articles, and other publications. Bibliometric methods, as techniques to measure the information distribution models, are frequently used in the field of information science and social research. The main purpose of this article is to offer scholars a general framework for the comparison between positive and negative aspects of bibliometrics, on the methods and tools used. Therefore, both the strengths and the critical points will be highlighted, to obtain a complete and detailed overview of the entire argument. In the methodological part, a bibliometric analysis will be applied to various case studies, such as with the Generalized Error Distribution, analysing and commenting on the data, and using the Bibliometrix software. The results suggest that in the future there will be greater consolidation of bibliometrics, as the introduction of increasingly advanced technologies will create new tools and methods characterized by a high degree of automation and speed.

Keywords: generalized error distribution, bibliometrics, Lp-norm, Skewed distributions, Ged-Garch

1. Introduction

The measurement as the quantification of attributes of an object or event is normally used to compare and evaluate a phenomenon. The scope and application of measurement are dependent on the context and discipline. In statistics as well as the social and behavioural sciences, measurements can have multiple levels, which would include nominal, ordinal, interval and ratio scales. Furthermore, measurement and metrics are fundamental activities for managerial sciences (such as in marketing and retailing) and in the creation and evaluation of value creation processes (Basile, 2019). The development of modern information technologies and the significant changes that are affecting the politics and organization of modern scientific research has led to a more decisive development of a quantitative approach to evaluative analysis. This approach is offered by Bibliometry, which in some way represents a guarantee of objectivity of the results, making the discipline itself an effective and ever-expanding tool in both the scientific evaluation and in that documentary policy in the library. The definition of bibliometrics physiologically approaches other methodologies such as Scientometry and Infometry which, while not presenting differences in methodology, are distinguished by having different purposes and areas of research. Therefore, Bibliometry is also evaluated about the other methodologies, making it a discipline as fascinating as it is complex. Although the coining of the term Bibliometry dates to 1969 by Alan Pritchard, its use and practice can be traced back to 1890. Bibliometry is a discipline that uses mathematical and statistical techniques to study distribution models of scientific productions and verify them. their impact and effect on the entire scientific community. This article after having dealt with all aspects of the discipline in detail highlights lights and shadows. Conceptually, this article will follow a very specific scheme of different parts, which refer to two different objectives. In the first part, we clarify the definitions of Bibliometry and its objectives, historical evolution of the discipline, bibliometric indexes and methods, Impact Factor and H-index and define the theoretical approach. In particular, the definition of the Bibliometry discipline and its objectives will be examined in-depth, dealing with it also about other methodologies. Our attention will focus on the historical and social evolution of the discipline, starting from its birth in the first decades of the twentieth century and reaching the role it covers today. Although the term was coined only in 1969, in this section reference will be made to clear examples that testify that long ago there were the first raw forms of this discipline. Moreover, we will deepen the discussion about bibliometric methods and the different differences between

traditional and modern methods, going to define what are the elements that have characterized this change. In addition, space will be given to the most important bibliometric methods, namely citation analysis and content analysis. Continuing, we find bibliometric indices, analysing first what they are and what they are for and highlighting the most important bibliometric indicators in terms of quality, namely the Impact Factor and the H-index. Furthermore, we can define the application approach with a brief introduction of the work carried out, which will indicate the tools used for the research, and an in-depth analysis of the Bibliometrix R package, which allowed us to analyse the data obtained. Subsequently, different case studies will be analysed on Lp-norm, on Ged-Garch models, on the Skewed Ged distribution, and finally on the inferential aspects of Generalized Error distribution. We decided to select a collection of data obtained from Scopus and analyse them using the Bibliometrix R package for each case study. The results obtained in the form of figures will be analysed and commented on, providing the reader with a complete reading key on the opportunities and possibilities offered by these tools. Furthermore, as evidence of the infinite breadth of possibilities, we have decided to use and comment on different graphics, to be able to obtain different and more significant information for each case study for that research. Finally, the last section of this article will be dedicated to our considerations of the role that Bibliometry can acquire in the future. Furthermore, there will also be an overall reflection of the various bibliometric analyses.

2. Definition of Bibliometry and Its Objectives

Bibliometry can be defined as the science that allows, using mathematical and statistical techniques, to investigate scientific production from a purely statistical point of view, analysing a very vast set of data, which can be modelled by referring to a given interval of time or referring to a specific sector that interests our research. Considering the wide range of applications of Bibliometry, it must also be analysed concerning other methodologies such as Scientometry and Infometry. Furthermore, the definition of Bibliometry physiologically approaches the methodologies just mentioned, which, while not presenting differences in methodology, are distinguished by having different purposes and areas of research. In particular, the main purpose of Bibliometry is to derive quantitative relationships between documents and between the elements that compose them, regardless of the area. Scientometry, whose founding father is the English information specialist Derek John De Solla Price, instead has as its main purpose the evaluation and measurement of the contribution of scientists, institutions, and nations in the advancement of knowledge, and to do this it can resort to a qualitative approach through, for example, peer review, the panel or the degree of internationalization, and a quantitative approach through the calculation of publications and citations. In this second aspect, Bibliometry and Scientometry become, in other words, two indistinguishable entities (De Bellis, 2005). Finally, Infometry, which derives from the term "informetric" and was first proposed in 1979 by Nacke to cover that part of information science that deals with the application of mathematical methods to the discipline, go to study information in any of its forms and areas. It is considered "the set of sets of all other metrics since they all count some kind of information" (De Bellis, 2005). What therefore distinguishes all the methodologies is the object of the analysis if the field of investigation concerns the problems relating to documents, we are in the field of Bibliometry if the investigation concerns information, we are in the field of Infometry if finally, the investigation concerns science we are in the field of Scientometry. Bibliometry helps to describe the history and general state of the art of a specific field or research topic, considering written production as the main channel of formal communication between scientists (Bellardo, 1980). Over time, different definitions have been developed that have made it possible to transform from a raw and primitive concept to a more worked and modelled one, which adapts more easily to the technological, political, scientific, statistical, and mathematical developments that have characterized the different phases of our company. Regarding the objectives, we can commonly delineate among the primary objectives, the search, and reception of useful information obtained only after analysing entire collections and bibliographic services and extracting quantitative relationships between documents and elements that compose them (for example words, citations, authors, institutions, etc.). The use of a bibliometric approach makes it possible to provide more objective and reliable analyses based on statistical techniques (Pritchard 1969; Broadus 1987; Diodato and Gellatly, 2013), having the ability to carry out both basic and advanced analyses of large volumes of documentation. The key procedures that allow this can be traced back to performance analysis (Peters and Van Raan, 1991; White and McCain, 1998) and scientific mapping (Börner et al, 2003; Noyons et al, 1999). The purposes of these procedures are different, on the one hand, the aim is to evaluate everything based on bibliographic data, for example by measuring the power or effect of certain actors on a specific area, on the other hand, on the other hand, it seeks to analyse cognitive models through a synchronic (Callon et al, 1983, Noyons and Van Raan, 1998) or diachronic (Cobo et al, 2011; Garfield, 1994) vision.

3. Historical Evolution of the Discipline

The term Bibliometry was coined by Alan Pritchard in 1969 who proposed to replace the little-used and somewhat ambiguous term "statistical bibliography" with Bibliometry (bibliometrics) to indicate the application of mathematics and statistical methods to books and other forms of written communication. The phases that have characterized the historical evolution of Bibliometry are 4: a first phase characterized by the very first forms of statistical bibliography, which is the basis of current Bibliometry; a second phase characterized by the increasingly accentuated development of this statistical

approach, witnessed by the processing and publication of what is defined by the entire scientific criticism as the fundamental laws of Bibliometry; a third phase that refers to the post-war period and the enormous contribution of Garfield, and finally the last phase that coincides with the last decade, in which new indexes and data databases were developed. Historically, Bibliometry originated in the West and was derived from statistical studies of bibliographies (Egghe and Rousseau, 1990). Although the term Bibliometry was coined in 1969, its use and application date back to the 90s of the nineteenth century. Campbell's work (1896) which used statistical methods to study the scattering of subjects in publications is probably the first concrete attempt at bibliometric studies (Sengupta, 1992). Arriving at this point of historical analysis, it may be natural to ask whether there were primitive forms of Bibliometry before the twentieth century, the answer is yes as it is enough to think, for example, of the complex phases of counting, conservation, and archiving that has characterized the most ancient times. ancient, starting from the Romans up to the Royal Library of Paris, which can be received as the earliest forms of this science. Only at the end of the nineteenth century did we begin to witness a process that would have led to the application of biometric techniques to books and other products of written communication, to obtain useful information for the acquisition policy in the library (De Bellis 2005). Let's say that in some way it represented the first phase of statistical bibliographic, which over time would change into the definitive Bibliometry. There are several studies carried out in the various fields of knowledge that have made an evolution of this magnitude possible, just think for example Franck Campbell who is a publication 1896 used statistical calculation to summarize the thematic breakdown of the publications belonging to the most important bibliographies of the time, or the work of FJ Cole and Nellie who applied statistical analysis to the literature of the sector, broken down by country and referred to a time interval, to measure the degree of growth of the reference discipline. A real turning point in the discipline, however, is represented by three works that appeared between 1926 and 1935, which contributed to the formulation and publication of the empirical laws on the behaviour of literature which are the basis of Bibliometry, and which have in some way started the second phase. Among the main military milestones of its development (Tague-Sutcliffe, 1992) are Lotka's method of measuring the productivity of scientists (1926), Bradford's law of dispersion of scientific knowledge (1934), and the distribution and frequency model of words in a Zipf's law (Ferrer i Cancho and Solé 2001). To understand the concept behind these three fundamental laws we use table 1, which summarizes the key idea and relationships of Lotka, Bradford, and Zipf (Chen and Leimkuhler, 1986).

Table 1. The main contributions on the bibliometrics' bases

Bibliometric bases	Findings/Main contributions	Main authors and references
Lotka's law	Lotka's law is a statistical observation offered by Alfred Lotka, to whom a relatively small number of scientists would be responsible for a large part of the contributions produced by the entire scientific community.	Bookstein (1977); Nicholls (1989); Newby et al. (2003); Patra et al. (2006)
Bradford's law	Bradford's law is a model first described by Samuel c. Bradford in 1934 which estimates the exponentially diminishing search returns of references to scientific journals.	Patra et al. (2006); Su et al. (2020)
Zipf's law	Zipf's law states that, if you take a text, which can be more or less long, and proceed to draw up a ranking of the words based on the decreasing frequency of appearance, it is possible to obtain a constant by multiplying the position number of the word in the ranking by the number of times it appears in the text.	Bail ón-Moreno et al. (2005); Adamic et al. (2020)

Source: own elaboration.

The third historical phase of the discipline opens with the end of the Second World War, in the new not only political but also socio-economic structure and the organization of knowledge, especially of a scientific nature. In the immediate post-war period, however, the evaluation of the research was still conducted without the use of bibliometric indicators, using instead either the tested system of peer-reviewing or macro-level economic indicators. Basically, in the post-war period, the idea and the need that the object of scientific activity can be kept under control and planned towards specific objectives is recovered, especially when it has been shown that science could be a lever for economic growth. At this point, the figure of Eugene Garfield becomes very important in the history of Bibliometry, they began working at the John Hopkins Welch Medical Library Project as he realized the inadequacy of the tools to receive useful information. Garfield introduced an interdisciplinary index project in 1955 for the first time in the journal Science. Thanks to Lederberg, Garfield's project found full realization only in 1963 with the publication of the Science Citation Index (SCI). The citations were the fundamental element for Garfield as they created a network of connections between documents, which

became what he had called the association-of-ideas index, capable of linking together documents that can hardly be placed under the same heading. Garfield also unwittingly managed to develop citations as an evaluation tool, and this evolution was received not only by the United States but also in all other countries around the world. Finally, then we move on to the fourth and last phase of the evolution process, the one that coincides with the last decade. In this period there have been important innovations such as the creation of new citational indices, including open-access ones, the development of increasingly sophisticated bibliometric indicators, and the application of the methodology in increasingly vast areas such as the world wide web. Furthermore, in addition to this, we must remember that over time various data databases have been born, such as Scopus, Web of Science, and Google Scholar, which allow us to analyse bibliographic collections in a short time and in a completely automatic way to obtain a series of charts to use for our analyses. In other words, in the last decade, there has been a change that has made bibliometric analysis an easy-to-use tool characterized by a fair speed of results.

4. Bibliometric Methods

Bibliometry guarantees a wide application in the field of knowledge, where bibliometric methods are used to be able to verify the impact and effect of a group of researchers or a particular publication in the entire scientific community. The most important and well-known bibliometric methods are certainly citation analysis and content analysis, alternatively, there are also other bibliometric applications such as the measurement of the frequency of terms and the exploration of grammatical structures, which compared to those mentioned previously have a less significant impact. Citation analysis is one of the main bibliometric methods that use citations in scientific productions to create a series of connections to other works or other researchers. In addition, a series of specific models operate within the same citation analysis, the coupling of co-citations and bibliographic couplings. As for the citation analysis, over time there have been numerous changes although for many decades the Science Citation Index (now owned by Clarivate Analytics) has been considered the main tool for measuring citations, for a time now. web services are questioning its entire dominance. The reason for this change coincides with the fact that the studies carried out in recent years provide rather different results, so it seems essential to resort to different sources to evaluate the citational impact of scientific production. Therefore, the Web is decisive in this sense, which has also allowed the birth of various data databases, such as Scopus and Google Scholar, which have introduced new methods of evaluating citations. Content analysis, on the other hand, born as a textual analysis when conducted exclusively on texts, is a character analysis in the social sciences applied to the study of the content of the communication. In practice, through the analysis of content, researchers are provided with a method that allows them to analyse large texts, for example by analysing the most used keywords to identify the supporting structures of their content or for example to detect which are the most used words within of the publication, etc. We say that these are the most important bibliometric methods we must consider that compared to the past these methods have acquired a certain speed and automation that allows a researcher or a group of them to be able to use these methods with exceptional ease. Therefore, the key element that makes traditional bibliometric methods different from modern ones is automation, which has been permanently mobilized within today's society and has influenced all its aspects.

5. The Bibliometrics Indexes

The evaluation criteria of scientific research can be of different types, for example, those of a quantitative type, those of a qualitative type, and finally those of another type. The quantitative criteria, i.e., through bibliometric indices, consist of the extrapolation of quantitative relationships between the various documents, using a quantitative analysis of bibliographic citations. The qualitative criteria, through peer review, or a procedure for selecting a series of articles and projects carried out following an evaluation by specialists who verify the suitability and correctness of the information. And finally, the criteria of another type, using the various forms, for example, participation by invitation to conferences, awarding of prizes and awards, publications on Wikipedia, and open access systems. What interests us mainly are the quantitative criteria, and therefore basically the bibliometric indexes. Bibliometric indices thus allow to quantitatively assess the impact of research within the disciplinary community to which they belong. They are usually based on the weight of the citations, but some different natures are not based on citation algorithms. Bibliometric indicators can be applied to the production of a single researcher, a periodical, a working group, a scientific community, a university, an entire country, etc. The elements considered for scientific research are different, for example, the number of citations of an author's publications, the number of citations received by journals (Impact Factor), and the productivity and impact of a single author's publications (H-index). Furthermore, not all indexes are easily usable as some may be covered by copyright that limits their use only to those who request appropriate authorization, for example, the case of the Impact Factor, and others that can be used freely. For all those who need it using online programs and databases, a clear example could be H-Index.

5.1 The Impact Factor (IF)

The Impact Factor (IF) is considered one of the most important bibliometric indices, developed in 1955 by the American

Eugene Garfield. The IF is contained in the Journal Citation Reports (JCR) publication, the publication of which took place for the first time in 1975. In this sense it is appropriate to make an accurate clarification, that is, it must be remembered that the IF does not is the only index contained in the JCR, but there are other indicators such as for example, the immediacy index, created to detect the speed with which the article is cited, The Cited Half-Life, which somehow measures the stability over time of the citations received and the 5-Year Impact Factor. The IF is obtained by dividing the number of citations received each year by the articles published in the same period and in the same magazine. Over time, moreover, a heated discussion has developed around this index regarding its effectiveness in obtaining acceptable results, in particular, the main criticisms derive from the fact that the two-year time base proved to be penalizing and not very representative for some disciplines; the impossibility of checking the data used for the calculation of the indicator; risk of distortion of results due to uncertainty in the definition of the quantities; the controlled abuse of self-citations and the lack of cleanliness of all those articles that present clear errors, considered as such by both the authors themselves and the scientific community. We have mentioned only some of the criticalities that characterize this index, moreover, as in all things, in addition to the weak points there are the strengths, more precisely we refer to the ease of the index, the stability of the measure over time and to the constant updating of the data. After having given a more theoretical reading key, let's move on to the practical one with the help of table 2 (below).

Table 2. Comparison of the citation number and papers published

<i>Citations in 2020 of articles published in 2018</i>	300
<i>Citations in 2020 of articles published in 2019</i>	150
TOTAL CITATIONS	450
<i>Number of articles published in 2018</i>	80
<i>Numbers of articles published in 2019</i>	40
TOTAL ARTICLES	120

Source: own elaboration.

The table above shows us the number of citations in 2020 of articles published in 2018, equal to 300, the number of citations in 2020 of articles published in 2019, equal to 150, and finally the total number of citations given by the sum of 300 +150 (450). It also gives us the total number of articles in 2018 and 2019, which we are going to add up to get the total number of articles. At this point, based on the information acquired, we just need to apply the formula provided for the IF:

$$IF = \frac{\text{Total citations}}{\text{Total Articles}}$$

In the case under consideration, we just need to divide the total number of citations, equal to 450, by the total number of articles, equal to 120. In conclusion, as we can see from the example just given, we are faced with an index that in its entirety it is characterized by a high degree of simplicity, which makes it one of the most used indices.

5.2 H-Index

The H-Index, the Hirsch index or also commonly called H-Index, is one of the most used indexes in the bibliometric field. This index, compared to the one discussed above, is born and works differently. The name derives from its founder, the physicist E. Hirsch of the University of California at San Diego, who invented this index to allow us to quantify the prolificacy and scientific impact of the author, based on both the number of publications and the number of citations received. According to Hirsch (2005), “a researcher has an index h if h of his published articles in n years (N_p) have obtained at least h citations each, and the remaining (N_p-h) articles have each received fewer than h citations”. So practically, a researcher who has an index equal to 5, basically means that he has made public 5 articles each of which has received 5 citations, while the remainder did not obtain any citation or in any case a small number. The H-Index was designed to make comparable the production of less prolific authors, but with a high number of citations, and that of very prolific authors but with a low number of citations (Piazzini, 2010). What perhaps makes this index highly appreciated by scientific critics is the fact that it is a free access index, unlike IF, and therefore usable by all those who need it without further authorizations. both an indicator that at the same time measures the quantity of the scientific production (i.e., the number of articles published in each period) of a researcher, and the quality/impact of the same (citational impact totalled in each period) (Hirsch, 2005). In addition to the advantages offered by bibliometric indices in general, such as the extreme objectivity of the results and the ease of calculation, this index, albeit a summary index of the quality and quantity of scientific production, is better suited than other indices (Hirsch 2005). Furthermore, being an objective information provider, it could be used to allocate funds to universities or researchers' promotions (Costas and Bordons 2007). Another important benefit is its robustness (Vanclay, 2008), given its insensitivity to the little-mentioned articles and the possibility

of calculating it easily with simple internet databases, from studies carried out it has been shown that its effectiveness is shown above all for comparisons that occur between scientists working in the same field, particularly physicists and mathematicians. Having said that, now let's move on to indicate which are the phases that allow us to derive the H-Index, and to do this we use one of the many databases, namely Scopus. Once you have logged into Scopus through your browser, you must select the search by authors by typing in "Author Search" and indicating the name parameters of the author in question. Once the research has started, it is necessary to select the documents relating to our survey and click on "View Citation Overview". Once this is done, the h-index will be shown on the right of the screen with the relative link for the figure. This procedure can also be repeated using other databases, as previously mentioned, or using the bibliometric R package, whereby going to the "Authors" tab we can open a figure that shows us, based on the selected data, the authors who have a value of the index h - upper index, as shown in the figure relating to the Ged-Garch case study.

6. Bibliometrics Analysis

The use of bibliometrics is gradually spreading to all disciplines. It is mainly adopted for science mapping due to the emphasis on empirical contributions in producing voluminous, fragmented, and controversial research streams. After the study design phase, analysed in the previous paragraphs, the second step includes the use of open-source statistics. In the data collection phase, we used the Scopus database to create the .bib file, ready for the third phase of data analysis. Our choice fell precisely on the Scopus database, founded by the Elsevier Science publishing house in 2004, as it is one of the most widespread databases on various scientific fields that are frequently used for research in the literature (Guz and Rushchitsky, 2009). Although alternative databases can be considered to retrieve bibliometric data sets, such as Web of Science or Google Scholar, our choice reflects our needs as we have found it appropriate to rely on the largest searchable and continuously updated citational source (Rew, 2009). About this, however, an appropriate clarification must be made, i.e., it must be said that a heated debate has arisen on which database is the best (Bar-Ilan et al, 2007; Vieira and Gomes 2009; De Winter et al. 2014; Har- zing and Alakangas, 2016), a very difficult choice as all databases lend themselves optimally to our work, what fundamentally changes are some small things that can make one more suitable instead of another. Perhaps what we have been able to appreciate about Scopus is the frequent periodic updating, which allows us to obtain very recent data, the wide range of articles from more than about 5000 publishers, and above all the presence of open access journals. That said, the case studies examined will focus on the Generalized Error Distribution, the Lp-norm estimators, the asymmetric Generalized Error distribution, and finally the Ged-Garch models.

6.1 Bibliometrix R-tool

Science mapping is manifold and bulky because it is multi-step and frequently requires numerous and several software tools, which are not all necessarily freeware (or freemium). Although automated workflows that integrate these software tools into an organized data flow are emerging. After selecting the publications relevant to our research, we used the Bibliometrix R package (Aria and Cuccurullo, 2016). Bibliometrix is an open-source tool for quantitative research in scientometry and Bibliometry that includes all major bibliometric methods. Bibliometrix as a package for bibliometric analysis written in R works with ecosystem software and operates in an integrated environment consisting of open libraries, the open algorithm and open-graphical software. The Bibliometrix package allows you to import data from major databases, perform bibliometric analyses and build data matrices for co-citation, coupling, scientific collaboration analysis and co-word analysis. Furthermore, *Biblioshiny*, a web interface for Bibliometrix, was used for the creation of a concept map and a co-citation network (Aria and Cuccurullo, 2016). Bibliographic data are processed through a workflow: study design, data collection, data analysis, data visualization and interpretation. Aria and Cuccurullo (2017) stated that bibliometric analysis is a cumbersome activity, which contains many producers. However, according to Guler et al. (2016), there are automated software tools that are used by information scientists or practitioners. By extracting descriptive and network data within bibliographic literature, one can perform citation analysis. Citation analyses are used to reveal the scientific growth in a specific field at three levels: micro, macro and meso. The conventional method used in citation analysis is bibliographic coupling, co-citation, co-author and co-word (Derviş, 2019). The analysis of the results then continued with their visualization using the data reduction technique. Bibliometrix supports a recommended workflow to perform bibliometric analyses. The proposed tool is flexible and can be rapidly upgraded and integrated with other statistical R packages as it is programmed in R. It is therefore useful in a constantly changing science such as bibliometrics. Data importing and converting in Bibliometrix follow a common workflow to search and export bibliographic documents, mainly based on three steps¹.

1. *Write and submit a query*: A query is usually based on a set of terms linked by Boolean operators. The search engine will query the DB to identify records matching the query.

((i.e., $TI = (\text{bibliometric AND analysis})$, the search engine will search for all the records in which the title will contain the words 'bibliometric' and 'analysis' simultaneously.))

2. *Refine search results*: Search results can be refined by applying some filtering criteria for additional fields.

(i.e., selecting Document Type = 'Journal Article' AND Language = 'English' AND Timespan = 1990:2020 AND Subject Category = 'Management')

3. *Export search results*: In this step, the user must choose which metadata he wants to download and the export file format to save the results. To be able to work, *bibliometrix* requires a minimum set of mandatory metadata.

(i.e., Authors' name, Title, Journal Title, Affiliation, Publication year, etc.). Our advice is to always select all the metadata fields to be sure you can perform all the analyses implemented in bibliometrix. Many databases support a variety of different export formats, some commercial (i.e., EndNote, Mendeley, etc.) and some standard (i.e., HTML, plaintext, BibTeX, etc.).

6.2 Generalized Error Distribution

The Generalized Error Distribution is a family of statistical distributions which constitutes a generalization in the Gauss curve. It contains a series of infinite probability distributions all symmetric (Giacalone, 2021). The first phase of our analysis coincides with researching the topic in question on the Scopus website and selecting the publications relevant to our research. After a series of unsuccessful attempts, due to the lack of relevance of the topics sought, and after the appropriate modifications, we managed to obtain an acceptable result in terms of quality and quantity. We decided to search for the more complex "Generalized Error Distribution parameter estimation" instead of the simple "Generalized Error Distribution", which allowed us to obtain a total number of documents equal to 1104 and several selected documents equal to 160. Then we downloaded the .bib file and using R-studio and the Bibliometrix R package we were able to obtain an enormous amount of information to analyse. In the "Dataset" tab, for example, it might be interesting to analyse what are the main information, discovering for example what is the reference range taken into consideration or what type of documents we have selected. All this information is summarized in the table below, which shows us that in our research we have selected several documents equal to 160 referring to a period ranging from 1983 to 2021 and with a total number of authors equal to 350.

Table 3. Descriptions and results of data set

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	1983:2021
Sources (Journals, Books, etc)	112
Documents	160
Average years from publication	7.38
Average citations per documents	10.28
Average citations per year per doc	1.099
References	3997
DOCUMENT TYPES	
article	138
book chapter	2
conference paper	18
review	2
DOCUMENT CONTENTS	
Keywords Plus (ID)	756
Author's Keywords (DE)	571
AUTHORS	
Authors	350
Author Appearances	393
Authors of single-authored documents	27
Authors of multi-authored documents	323
AUTHORS COLLABORATION	
Single-authored documents	29
Documents per Author	0.457
Authors per Document	2.19
Co-Authors per Documents	2.46
Collaboration Index	2.47

Source: own elaboration.

In the source tab, we can determine the most relevant source using a specific chart entitled "Most Relevant Sources". Concerning our research, we can note, in figure 1, that about the number of documents the most relevant source is Communications in statistics simulation and computer with several documents that goes beyond 8, while the position of the least relevant source is disputed between different sources that hold less than 2.5 documents.

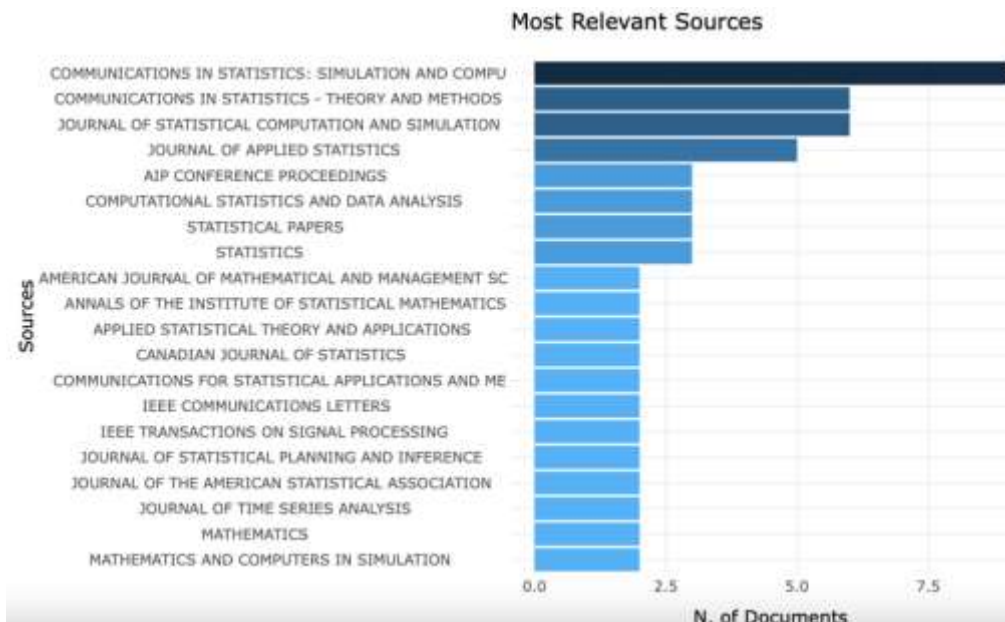


Figure 1. Statistic simulation of most relevant sources

Source: own elaboration.



Figure 2. Country collaboration map

Source: own elaboration.

In figure 2, the Country Collaboration Map is shown instead, that is the collaboration map. This map allows us to analyse the collaboration network using a geographic map. As we can see, the world countries take on a more or less bright blue

colour depending on the number of publications, for example, the brightest blue can be seen in the United States while a less bright blue is noted in Europe. Furthermore, it is important in this sense to follow the red lines that are precisely the links of collaborations, for example, the most important one is the one that connects the part of the United States with China. This figure is important because it allows us to understand which links are underlying the individual publications within the data collection. As regards the analysis of the content, we can draw a lot of other information from the R-Bibliometrix package in the form of figures. A concrete example is given in figure 3 below, which shows us the so-called “TreeMap”, which is a tree-lined map that allows us to show hierarchical data using nested rectangles. This map can be obtained for 4 types of parameters: title, abstract, keywords and authors keywords. In this case, we take into consideration the one related to the title and we put the maximum number of words or 200. As we can see from the map, the percentage and the size of the squares, the term Generalized has a value equal to 84 and a percentage equal to 8% on the entire set of titles, this means that generalized has been used in the title 84 times out of 100. Below we have for example estimation, parameter, model, regression, distribution etc. with the relative values and percentages, which can be useful for identifying which are the keywords used by the authors in the title. In general, if we analyse the trend of scientific production in this argument, we can see how their significant growth has been, due to the speed of obtaining these data. With this speed compared to the past, not only has there been certain automation, but the same speed has made this type of information more attractive to all those who have scientific production as their objective.



Figure 3. Treemap hierarchical data

Source: own elaboration.

6.3 Lp-Norm

Lp-Norm estimators are characterized by an exponent p which generalizes the method of minima, when p is a 2 the estimators Lp-Norm coincide with the method of least squares (Giacalone, 2020). The first necessary step is to search for Lp-Norm on the Scopus website, selecting the articles and publications that interest us and are relevant to our research. We initially searched only for the term without inserting any time limitation or limitation concerning the area to which it belongs, obtaining a result of 2,369 documents (see Figure 4).

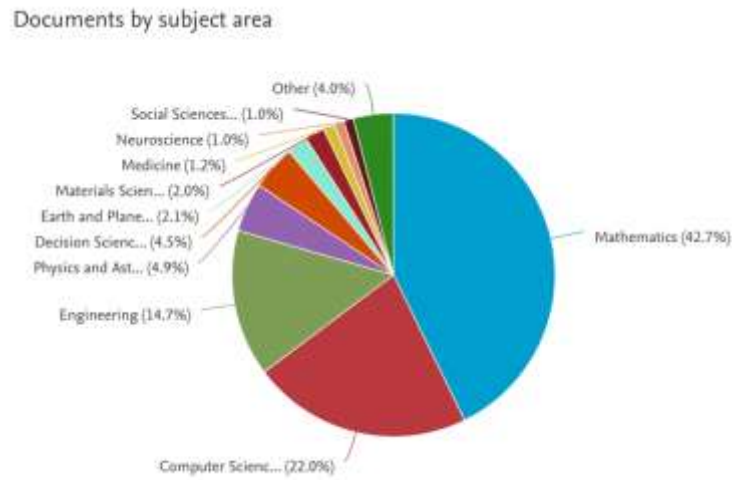


Figure 4. Lp-Norm on the Scopus website

Source: own elaboration.

However, analysing a collection of 2,369 documents, in which some are fundamentally not relevant to our research, appears to work as complex as it is useless. So, we worked to make some changes to the research, searching for “LP-NORM REGRESSION” instead of the cruder “Lp-NORM” and manually entering the relevant fields (Giacalone et al., 2018). This resulted in a result of 150 documents, which represents an acceptable number to start phase 2 of this analysis, which consists in selecting the items and exporting them. Subsequently, using RStudio and the R Bibliometrix package (Aria and Cuccurullo, 2017) and loading the previously obtained file, we were able to obtain a series of information that is filtered and catalogued according to different criteria. In the “Dataset” tab, for example, we can obtain what are the “Main Information” (Main information, see table 4).

Table 4. Main dataset information

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	1964:2019
Sources (Journals, Books, etc)	29
Documents	37
Average years from publication	20.9
Average citations per documents	18.51
Average citations per year per doc	1.058
References	1005
DOCUMENT TYPES	
article	33
conference paper	3
review	1
DOCUMENT CONTENTS	
Keywords Plus (ID)	150
Author's Keywords (DE)	125
AUTHORS	
Authors	63
Author Appearances	72
Authors of single-authored documents	15
Authors of multi-authored documents	48
AUTHORS COLLABORATION	
Single-authored documents	16
Documents per Author	0.587
Authors per Document	1.7
Co-Authors per Documents	1.95
Collaboration Index	2.29

Source: own elaboration.

This table is divided into 5 parts: a first part that contains all the information relating to the selected data, such as the time interval taken into consideration (1964: 2019), the sources (newspapers, books etc., 29), the number of documents (37), average years from publication (20.9), average citations per document (18.51), average citations per year per document (1,058), references (1005); a second part that contains information relating to the type of documents, in this case, we have 33 articles, 3 conference papers and 1 review; a third party concerning the contents of the documents; a fourth part that summarizes the information relating to the authors, for example, the total number (63), the appearances of the author (72) etc.; and finally the last part concerning the collaboration of the authors. Continuing our analysis, in the “Authors” tab, for example, we may be interested in seeing the most relevant authors (in the Most Relevant Authors section, see figure 5).

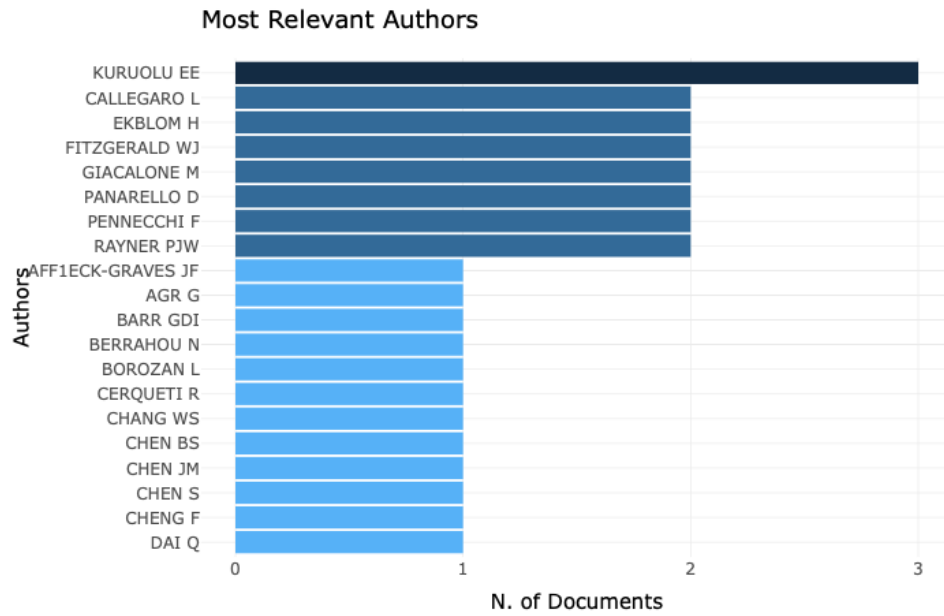


Figure 5. Author’s relevance

Source: own elaboration.

As we can see from the figure, in this selected range the most relevant author based on the number of documents is Kuruoglu EE. (3), closely followed by Callegaro L. (2), Giacalone M. (2), Panarello D. (2), Penneccchi F. (2) etc. Furthermore, in the data collection we have selected, it could also be interesting to note which keywords are most used by the authors, by inserting the "author's keywords" field in the parameter and entering the value 5 as the number of words to obtain a more precise result (see the following figure 6).

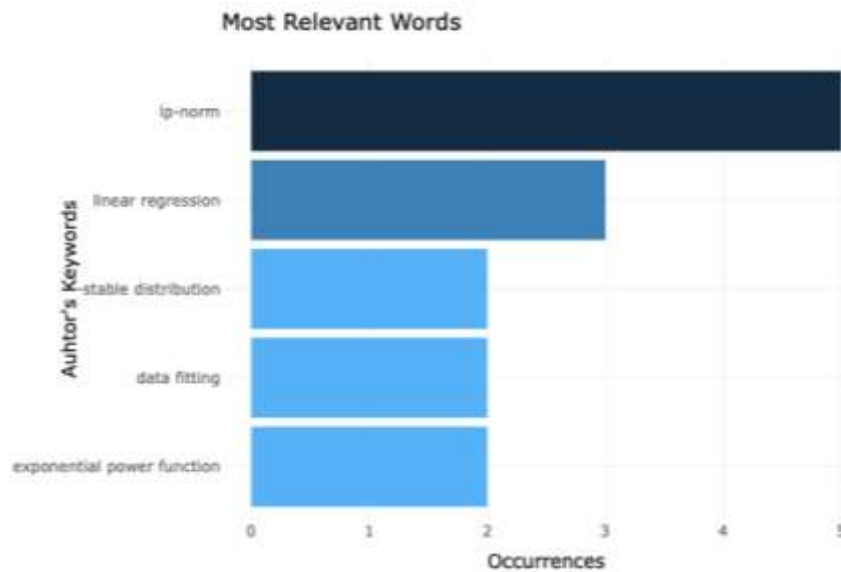


Figure 6. Author’s keywords

Source: own elaboration.

The result obtained (figure 7 below) shows what we mentioned previously, with the keyword "LP-NORM" occupying the first position and the keyword "Exponential Power Function" the last. Finally, figure 7 (historical direct citation network) shows us the network of historical citations from 1973 to 2019. This allows us to understand the passages over time that led to the citation network using a grey line. as a whole.

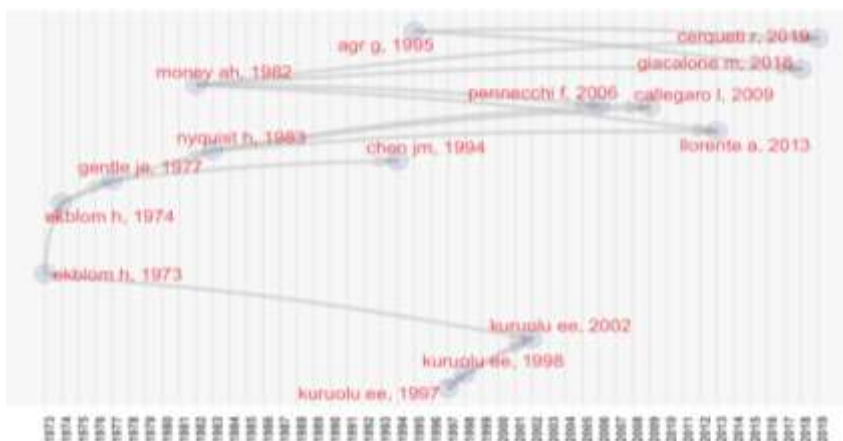


Figure 7. Historical direct citation network

Source: own elaboration.

6.4 GED-GARCH Model

This paragraph will analyse the data obtained by querying the Scopus database on GED-GARCH Model on 5 November 2020. First, it is important to define the concept of GED-GARCH to better understand what we will go into later. The Garch are time series models created to predict the volatility of time series, while GED-GARCH is attributable to the factual situation that occurs when the time series models have residues coming from the Generalized Error Distribution. After following all the usual steps to the letter, we were able to obtain a file of 59 documents (Cerqueti et al., 2020); information obtainable in the “Main Information” table present in the Dataset sheet.

Table 5. Main dataset information

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2000:2020
Sources (Journals, Books, etc)	56
Documents	59
Average years from publication	7.03
Average citations per documents	8.186
Average citations per year per doc	0.8442
References	1376
DOCUMENT TYPES	
article	39
book chapter	2
conference paper	18
DOCUMENT CONTENTS	
Keywords Plus (ID)	268
Author's Keywords (DE)	193
AUTHORS	
Authors	118
Author Appearances	138
Authors of single-authored documents	9
Authors of multi-authored documents	109
AUTHORS COLLABORATION	
Single-authored documents	10
Documents per Author	0.5
Authors per Document	2
Co-Authors per Documents	2.34
Collaboration Index	2.22

Source: own elaboration.

In addition to the same information seen previously (see Lp-Norm case study), it is important to note that the time interval analysed is 2000: 2020, i.e., only publications and articles from 2000 to 2020 were taken into consideration. The reasons for this choice derive from the fact that selecting data that are too old would have somehow altered the results of the research.

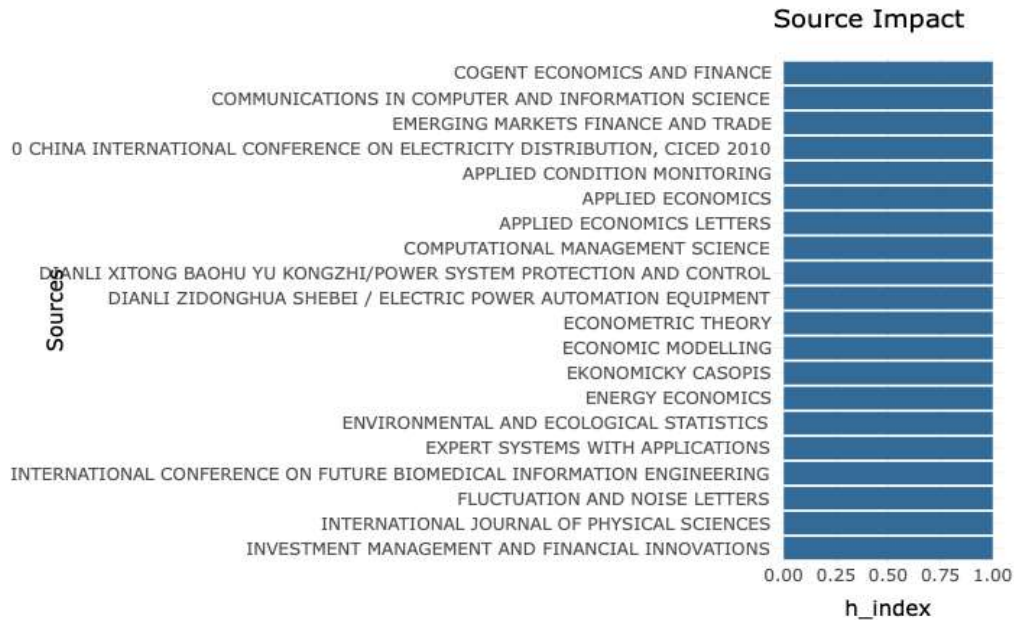


Figure 8. Source impact on I-Index

Source: own elaboration.

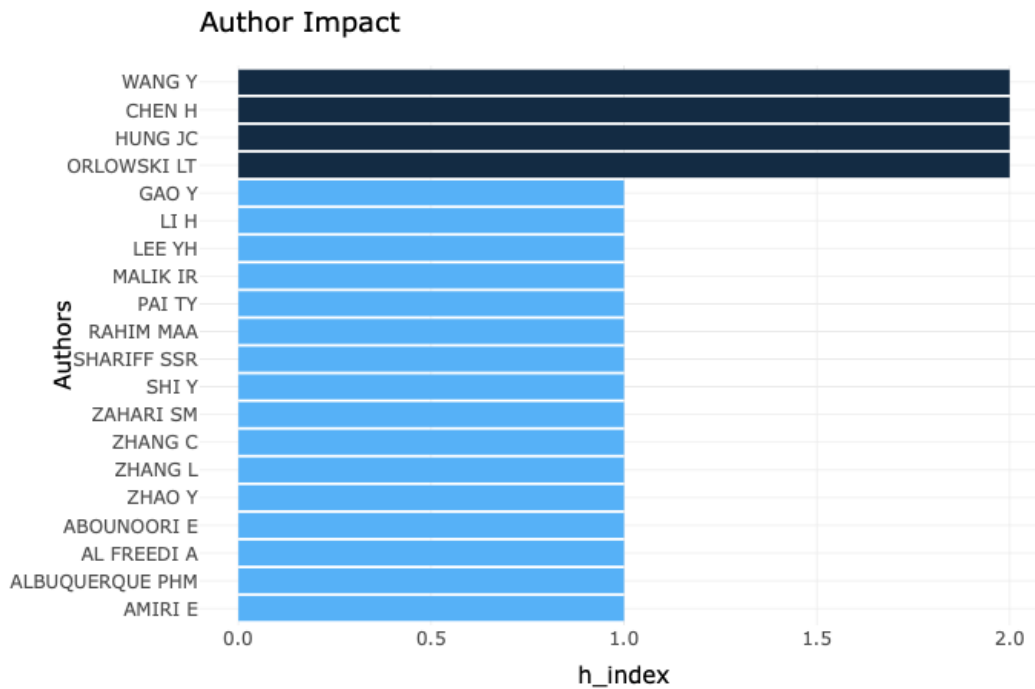


Figure 9. Author's impact on I-Index

Source: own elaboration.

H-index referred to the sources, as we can see from figure 8, is constant for all sources. A sort of objective balance of sources and sources can therefore be noted. Different speech regarding H-Index referred to the authors (see figure 9), in which there is no clear equilibrium but there are authors who have a higher index than others. Obviously, in these

figures we have inserted a maximum number of sources and authors, to avoid obtaining a figure that is too dispersive and not very precise. Although the two figures have common elements, we cannot fail to highlight the differences in the results obtained while the H-Index referred to the sources is constant due to the more or less equal importance of the same, the H-Index referred to the individual authors tends to be more differentiated due to its high subjective character. Usually, the comparison between two figures that refer to two different objects is fruitless, in this case, however, our main goal is yes to comment and analyse the data obtained but one of our goals is also to allow the reader to fully understand what the opportunities and possibilities are offered. For example, these figures can be useful to the scientific researcher if you want to carry out a cross-analysis of the authors and sources about a specific topic, taking the H-index value. In bibliometric analysis, even elements that are apparently of little value may become the subject of investigation. To explain what has just been said, let's look at figure 9, which examines the scientific production of the different countries, and table 6, which summarizes all the frequencies relating to the countries in question.

Table 6. Frequency and relative region

region	Freq
CHINA	60
IRAN	5
MALAYSIA	5
USA	5
PAKISTAN	4
CANADA	3
AUSTRALIA	2
ETHIOPIA	2
POLAND	2
SOUTH AFRICA	2
TUNISIA	2
UK	2
BELGIUM	1
BRAZIL	1
CZECH REPUBLIC	1
GREECE	1
JAPAN	1
MAURITIUS	1
MOROCCO	1
NETHERLANDS	1

Source: own elaboration.

Table 6 clearly shows us that China holds the highest value of scientific documents produced, equal to 60 (as can be seen in the map of figure 9 below), in the time interval from 2000 to 2020, this information is reproduced in figure 9 in the form of a geographical map using a more marked blue. The most relevant information is undoubtedly the enormous gap/gap that exists between the leader's China and the following countries, including the USA, which is still at value 5. Furthermore, the quantitative analysis of scientific production can also be analysed by taking as a reference a figure showing the progressive growth of the different countries. After comparing the different figures, to be able to understand, for example, what progress has been made within the entire scientific community or even analyse how China's growth has occurred from 2000 to today.

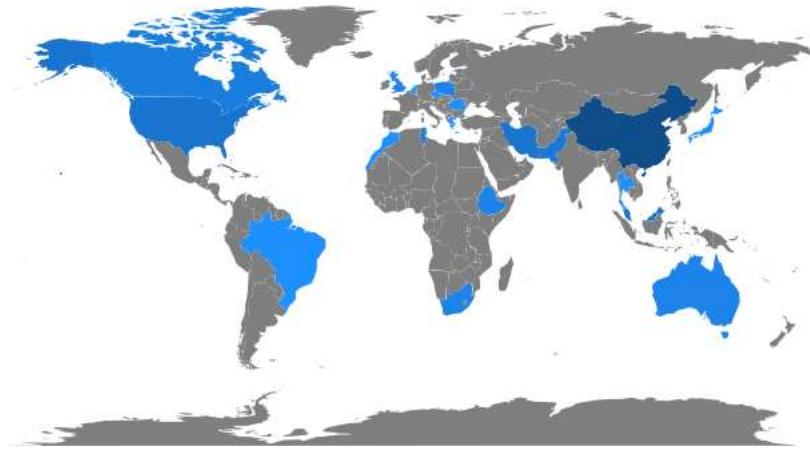


Figure 9. Country scientific production

Source: own elaboration.

6.5 Skewed GED

The latest bibliometric analysis of this article has as its research object Skewed Ged, that is, a family of an asymmetric probability distribution. Among the special cases, we remember the asymmetric Gauss distribution, the asymmetric uniform distribution and the asymmetric Laplace distribution. For Skewed Ged we obtained a data file consisting of 75 documents from 65 different sources (Cerqueti, 2021). Furthermore, in this research, we refer exclusively to 3 types of documents: 68 articles, 2 book chapters and 5 conference papers. Based on these data and using the R-Bibliometrix package we have obtained several interesting figures, including figure 10 below, which shows us an overview of the average citations of the article per year.

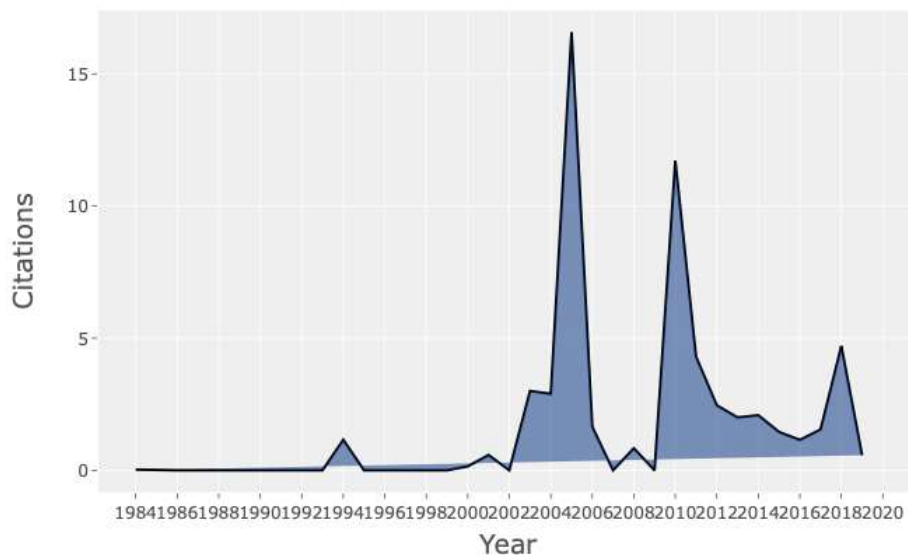


Figure 10. Average article citations per year

Source: own elaboration.

Figure 10 shows us the years on the x-axis, referring to the time interval taken into consideration, and the number of citations on the y-axis. We can see how from 1988 to the beginning of the 2000s there is a constant balance in the number of citations. Since the 2000s there has been a noticeable improvement, we can note the presence of impressive growth in correspondence with the two years 2004-2006. Further growth of less significant significance compared to the two years 2004-2006, occurred in 2010, where the number of citations reached a quota higher than 10. This figure offers us various food for thought, for example, it could be interesting to understand for what reasons there was a growth in the two years 2004-2006 and 2010 or understand why from 2000 onwards the situation has improved in terms of citations. Another important figure that could help us understand the differences between the period before the 2000s and the period from

2000 to today is the one that shows the annual scientific production.

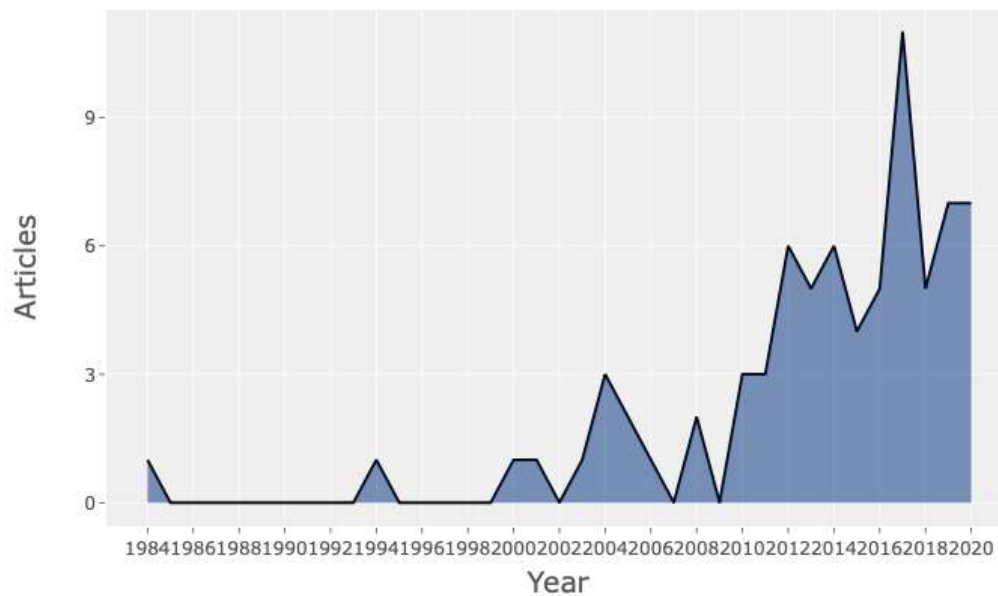


Figure 11. Annual scientific production

Source: own elaboration.

As shown in figure 11, over time there has been a progressive growth in scientific production what stands out most is the difference between the 2000s and the previous years. This gap could be justified by the introduction of new technologies that have made it possible to develop new methods and tools, characterized by a certain basic speed and a certain degree of automation. In the "Conceptual Structure" tab it is interesting to study the information contained in figure 12. This figure shows us the thematic evolution that has affected the data selected from 1984 to today.

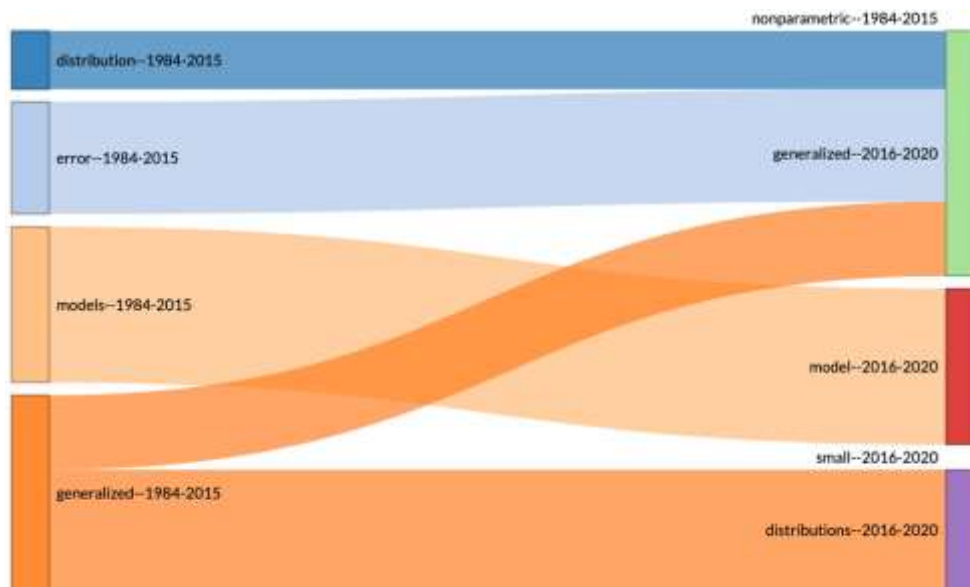


Figure 12. Annual scientific production

Source: own elaboration.

7. Limitations

The present study has some limitations. First, the application is based only on the Scopus database excluding other useful important databases such as Google Scholar, Research Gate and Academia. However, the thematic analysis performed by authors is qualitative and, consequently, subjective other researchers' analyses might result in indifferent themes and areas (Kokol et al., 2021). We must note that using only the Scopus database we have excluded other important papers for the

research topics considered in our application study. We have not considered other papers that are not indexed in this database. For example, in our application, we have not included papers such as conference proceedings and this is an important limitation. Therefore, we advise that the interpretation of the results take into consideration these limitations. For this reason, we suggest that it is necessary to consider data from a different database to confirm our application and compare the results obtained on the research topics considered.

8. Conclusions

The need for an objective approach to research evaluation, much desired by the entire scientific community to be able to count on tools that would allow for impartial evaluations, is the same as that of librarians who have introduced Bibliometry among the most used methodologies to reduce difficulties inherent to the subjectivity of assessments. This need for this type of approach will allow bibliometrics to be able to enjoy ample room for growth and redefine its role within society in the future. In our opinion, there will be greater consolidation of Bibliometry in the future, as the introduction of increasingly advanced technologies will create new tools and methods characterized by a high degree of automation and speed. Furthermore, I believe that it is extremely important, to favour this type of approach, to introduce over time new open-access systems or in any case new tools that are at the same time easy to use and above all free. As for the results obtained from the figures, it is important to underline the fact that they cannot be generalized, this is because we must consider that we have used Scopus as a reference database compared to others. What we are sure of, or rather we expect, is that over time there will be a progressive growth in both entire scientific productions, because the community will physiologically tend to progress, and a growth that will lead to an improvement in the results obtained. In conclusion, I think we are dealing with a discipline capable of guaranteeing enormous satisfaction over time both from the point of view of objective evaluation and from the point of view of research, representing one more piece of the complicated mosaic of scientific literature.

References

- Adamic, L. A., & Huberman, B. A. (2002). Zipf's law and the Internet. *Glottometrics*, 3(1), 143-150.
- Aghaei Chadegani, A., Salehi, H., Yunus, M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. *Asian social science*, 9(5), 18-26. <https://doi.org/10.5539/ass.v9n5p18>.
- Anzillero, T., & RIDI, R. (2012). *La bibliometria: metodi e strumenti per la misurazione della ricerca scientifica e delle raccolte bibliotecarie* (Bachelor's thesis, Università Ca'Foscari Venezia). <https://doi.org/10.4000/books.ledizioni.1334>.
- Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em questão*, 12(1), 11-32.
- Aria, M., & Cuccurullo, C. (2016). A brief introduction to bibliometrix. *Scientometrics*, 1, 1-17.
- 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>.
- Aria, M., Misuraca, M., & Spano, M. (2020). Mapping the evolution of social research and data science on 30 years of Social Indicators Research. *Social Indicators Research*, 1-29. <https://doi.org/10.1007/s11205-020-02281-3>.
- Baccini, A. (2010) *Valutare la ricerca scientifica. Uso e abuso degli indicatori bibliometrici*, Bologna, Il Mulino.
- Bailón-Moreno, R., Jurado-Alameda, E., Ruiz-Baños, R., & Courtial, J. P. (2005). Bibliometric laws: Empirical flaws of fit. *Scientometrics*, 63(2), 209-229. <https://doi.org/10.1007/s11192-005-0211-5>.
- Bar-Ilan, J., Levene, M., & Lin, A. (2007). Some measures for comparing citation databases. *Journal of Informetrics*, 1(1), 26-34. <https://doi.org/10.1016/j.joi.2006.08.001>.
- Basile, V. (2019). Marketing Performance Measurement in FMCG: Share of Wallet in Retailing Industry. *Editoriale Scientifica Naples*. October 2019. <https://doi.org/10.19245/25.05.bs.063>.
- Bellardo, T. (1980). The use of co-citations to study science.
- Bookstein, A. (1977). Patterns of scientific productivity and social change: A discussion of Lotka's law and bibliometric symmetry. *Journal of the American Society for Information Science*, 28(4), 206-210.
- Börner, K., Chen, C., & Boyack, K. (2003). Visualizing knowledge domains. In *Annual Review of Information Science and Technology*, 37, 179-255. <https://doi.org/10.1002/aris.1440370106>.
- Bornmann, L., & Daniel, H. D. (2007). What do we know about the h index? *Journal of the American Society for Information Science and Technology*, 58(9), 1381-1385. <https://doi.org/10.1002/asi.20609>.
- Brescia, V. (2020). Bibliometrix analisi: volontariato e community-based.

- Broadus, R. N. (1987a). *Early approaches to bibliometrics*, «Journal of the American Society for Information Science», 38(2), 127-129. [https://doi.org/10.1002/\(SICI\)1097-4571\(198703\)38:2<127::AID-ASI6>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1097-4571(198703)38:2<127::AID-ASI6>3.0.CO;2-K).
- Broadus, R. N. (1987). Toward a definition of “bibliometrics”. *Scientometrics*, 12(5), 373-379. <https://doi.org/10.1007/BF02016680>.
- Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. In *Social Science Information*, 22(2), 191–235. <https://doi.org/10.1177/053901883022002003>.
- Campbell, F. (1896). *The theory of national and international bibliography: With special reference to the introduction of the system in the record of modern literature*. Library Bureau.
- Cerqueti, R., Giacalone, M., & Mattera, R. (2020). Skewed non-Gaussian GARCH models for cryptocurrency volatility modelling. *Information Sciences*. <https://doi.org/10.1016/j.ins.2020.03.075>.
- Cerqueti, R., Giacalone, M., & Mattera, R. (2021). Model-based fuzzy time series clustering of conditional higher moments. *International Journal of Approximate Reasoning*, 134, 34-52. <https://doi.org/10.1016/j.ijar.2021.03.011>.
- Cerqueti, R., Giacalone, M., & Panarello, D. (2019). A Generalized Error Distribution Copula-based method for portfolios risk assessment. *Physica A: Statistical Mechanics and its Applications*, 524, 687-695. <https://doi.org/10.1016/j.physa.2019.04.077>.
- Chen, Y. S., & Leimkuhler, F. F. (1986). A relationship between Lotka's law, Bradford's law, and Zipf's law. *Journal of the American Society for Information Science*, 37(5), 307-314. [https://doi.org/10.1002/\(SICI\)1097-4571\(198609\)37:5<307::AID-ASI5>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1097-4571(198609)37:5<307::AID-ASI5>3.0.CO;2-8).
- 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. In *Journal of Infometrics*, 5(1), 146–166. <https://doi.org/10.1016/j.joi.2010.10.002>.
- Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro-level. *Journal of infometrics*, 1(3), 193- 203. <https://doi.org/10.1016/j.joi.2007.02.001>.
- De Bellis, N. (2005). La citazione bibliografica nell'epoca della sua riproducibilità tecnica: bibliometria e analisi delle citazioni dallo Science Citation Index alla Cybermetrica.
- De Bellis, N. (2014) *Introduzione alla bibliometria: dalla teoria alla pratica*, Roma, AIB.
- De Robbio, A. (2007). Analisi citazionale e indicatori bibliometrici nel modello Open Access. *Bollettino aib*, 2007(3), 257-288.
- de Winter, J. C. F., Zadpoor, A. A., & Dodou, D. (2014). The expansion of google scholar versus web of science: A longitudinal study. *Scientometrics*, 98(2), 1547–1565. <https://doi.org/10.1007/s11192-013-1089-2>.
- Derviş, 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>.
- Diaby, V., Campbell, K., & Goeree, R. (2013). Multi-criteria decision analysis (MCDA) in health care: a bibliometric analysis. *Operations Research for Health Care*, 2(1-2), 20-24. <https://doi.org/10.1016/j.orhc.2013.03.001>
- Diodato, V. P., & Gellatly, P. (2013). *Dictionary of bibliometrics*. New York: Routledge. <https://doi.org/10.4324/9780203714133>.
- Egghe, L., & Rousseau, R. (1990). *Introduction to infometrics: Quantitative methods in library, documentation and information science*. Elsevier Science Publishers.
- Ferrer i Cancho, R., & Solé R. V. (2001). Two regimes in the frequency of words and the origins of complex Lexicons: Zipf's law revisited*. *Journal of Quantitative Linguistics*, 8(3), 165-173. <https://doi.org/10.1076/jql.8.3.165.4101>.
- Garfield, E. (1994). Scientography: Mapping the tracks of science. In *Current Contents: Social & Behavioural Sciences*, 7, 5–10.
- Giacalone, M. (2020). A combined method based on kurtosis indexes for estimating p in non-linear Lp-norm regression. *Sustainable Futures*, 100008. <https://doi.org/10.1016/j.sftr.2020.100008>.
- Giacalone, M. (2021). Optimal forecasting accuracy using Lp-norm combination. *Metron*, 1-44. <https://doi.org/10.1007/s40300-021-00218-5>.
- Giacalone, M., Panarello, D., & Mattera, R. (2018). Multicollinearity in regression: an efficiency comparison between Lp-norm and least-squares estimators. *Quality & Quantity*, 52(4), 1831-1859. <https://doi.org/10.1007/s11135-017-0571-y>.

- Guler, A. T., Waaijer, C. J., & Palmblad, M. (2016). Scientific workflows for bibliometrics. *Scientometrics*, 107(2), 385-398. <https://doi.org/10.1007/s11192-016-1885-6>.
- Guz, A. N., & Rushchitsky, J. J. (2009). Scopus: A system for the evaluation of scientific journals. *International Applied Mechanics*, 45(4), 351. <https://doi.org/10.1007/s10778-009-0189-4>.
- Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the web of science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, 106(2), 787-804. <https://doi.org/10.1007/s11192-015-1798-9>.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of sciences*, 102(46), (p. 16569-16572). <https://doi.org/10.1073/pnas.0507655102>.
- Kokol, P., Blažun Vošner, H., & Završnik, J. (2021). Application of bibliometrics in medicine: a historical bibliometrics analysis. *Health Information & Libraries Journal*, 38(2), 125-138. <https://doi.org/10.1111/hir.12295>.
- La Torre, G., Sciarra, I., Chiappetta, M., & Monteduro, A. (2017). Nuovi indicatori bibliometrici nella letteratura scientifica: un panorama in continua evoluzione. *Clin Ter*, 168(2), e65-71.
- Misuraca, M., Scepi, G., & Spano, M. (2020). A bibliometric study on the evolution of conjoint analysis in 1998-2017. *Statistica Applicata-Italian Journal of Applied Statistics*, (1), 9-27.
- Nacke, O. (1979). Informetrie: Ein neuer Name für eine neue Disziplin, *Nachrichten für Dokumentation*.
- Newby, G. B., Greenberg, J., & Jones, P. (2003). Open-source software development and Lotka's law: bibliometric patterns in programming. *Journal of the American Society for information science and technology*, 54(2), 169-178. <https://doi.org/10.1002/asi.10177>.
- Nicholls, P. T. (1989). Bibliometric modelling processes and the empirical validity of Lotka's law. *Journal of the American Society for Information Science*, 40(6), 379-385. [https://doi.org/10.1002/\(SICI\)1097-4571\(198911\)40:6<379::AID-AS11>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-4571(198911)40:6<379::AID-AS11>3.0.CO;2-Q).
- Noyons, E. C. M., & van Raan, A. F. J. (1998). Advanced mapping of science and technology. In *Scientometrics*, 41(1-2), 61-67. <https://doi.org/10.1007/BF02457967>.
- Noyons, E. C. M., Moed, H. F., & van Raan, A. F. J. (1999). Integrating research performance analysis and science mapping. In *Scientometrics*, 46(3), 591-604. <https://doi.org/10.1007/BF02459614>.
- Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature I. *Libri*, 46(3), 149-158. <https://doi.org/10.1515/libr.1996.46.3.149>.
- Paolini, A., & Quagli, A. (2013). Una riflessione sugli strumenti bibliometrici per la valutazione della ricerca e una proposta: il real impact factor. *Management Control*. <https://doi.org/10.3280/MACO2013-003007>.
- Patra, S. K., & Mishra, S. (2006). Bibliometric study of bioinformatics literature. *Scientometrics*, 67(3), 477-489. <https://doi.org/10.1556/Scient.67.2006.3.9>.
- Patra, S. K., Bhattacharya, P., & Verma, N. (2006). Bibliometric study of literature on bibliometrics. *DESIDOC Journal of Library & Information Technology*, 26(1). <https://doi.org/10.14429/dbit.26.1.3672>.
- Peters, H., & van Raan, A. F. J. (1991). Structuring scientific activities by co-author analysis: An exercise on a university faculty level. In *Scientometrics*, 20(1), 235-255. <https://doi.org/10.1007/BF02018157>.
- Piazzini T. (2010), Gli indicatori bibliometrici riflessioni sparse per un uso attento e consapevole, *JLIS.it*, 1.
- Pritchard, A. (1969). *Statistical bibliography or bibliometrics?* «Journal of Documentation», 25(4), 348-49. <https://doi.org/10.1108/eb026482>.
- Sengupta, J. K. (1992). Fuzzy systems approach in data envelopment analysis. *Computers & mathematics with applications*, 24(8-9), 259-266. [https://doi.org/10.1016/0898-1221\(92\)90203-T](https://doi.org/10.1016/0898-1221(92)90203-T).
- Su, Y. S., Lin, C. L., Chen, S. Y., & Lai, C. F. (2020). Bibliometric study of social network analysis literature. *Library Hi-Tech*. <https://doi.org/10.1108/LHT-01-2019-0028>.
- Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information processing & management*, 28(1), 1-3. [https://doi.org/10.1016/0306-4573\(92\)90087-G](https://doi.org/10.1016/0306-4573(92)90087-G).
- Vanclay, J. K. (2008). Ranking forestry journals using the h-index. *Journal of informetrics*, 2(4), 326-334. <https://doi.org/10.1016/j.joi.2008.07.002>.
- Vieira, E. S., & Gomes, J. A. N. F. (2009). A comparison of Scopus and web of science for a typical university. *Scientometrics*, 81(2), 587-600. <https://doi.org/10.1007/s11192-009-2178-0>.

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for information science*, 49(4), 327-355. [https://doi.org/10.1002/\(SICI\)1097-4571\(19980401\)49:4<327::AID-ASI4>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-4571(19980401)49:4<327::AID-ASI4>3.0.CO;2-4).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

ⁱ <https://www.bibliometrix.org>, © 2022 K-Synth Srl, Academic Spin-Off of the University of Naples Federico II. All Rights Reserved.