

Statistics

Literature Review on Control Charts for Autocorrelated Processes

Revisão de Literatura sobre cartas de controle para processos autocorrelacionados

**Valentina Wolff Lirio^I , Tailon Martins^I , Bianca Reichert^I ,
Gilvete Sylvania Wolff Lirio^{II} , Adriano Mendonça Souza^I ,
Wesley Vieira da Silva^{III} **

^I Universidade Federal de Santa Maria, Departamento de Estatística, Santa Maria, RS, Brasil

^{II} Universidade Regional Integrada do Alto Uruguai e das Missões (Campus Santo Ângelo),
Departamento de matemática, Santo Ângelo, RS, Brasil

^{III} Universidade Federal de Alagoas – Maceió, AL, Brasil

ABSTRACT

This research aims to present a literature review (LR) on control charts for autocorrelated processes, intending to contribute to the scientific knowledge of the process management area. The article was constructed having a research question defined ex-ante, elaborated based on the literature, where a research protocol adapted from Tranfield, Denyer, and Smart (2003) was systematized according to the methodological rigor demanded in the literature review, which resulted in the composition of the research corpus. The research corpus was evaluated in detail using the bibliometric packages HistCite, VOSviewer, package R and Iramuteq. It was evaluated the validity of the 3 bibliometric laws were verified: Lotka's Law, Bradford's Law, and Zipf's Law based on the authors' citation and co-citation techniques. This study is considered as relevant and original since it is the first literature review on control charts for autocorrelated processes. The results confirmed the 3 classical bibliometric laws for the researched corpus. As a practical implication, this review provides support for quality management scholars to unlock potential research gaps. The main limitation of the work refers to the composition of the textual corpus, given that databases of national journals were not consulted due to the difficulty in terms of compatibility of the software used in this research.

Keywords: Statistical process control; Autocorrelated data; Quality management

RESUMO

Esta pesquisa tem por objetivo apresentar uma revisão de literatura (RL) sobre cartas de controle para processos autocorrelacionados, com a pretensão de contribuir para o conhecimento científico da área de gestão de processos. O artigo foi construído tendo uma pergunta de pesquisa definida ex-ante, elaborada com base na literatura, onde foi sistematizado um protocolo de pesquisa adaptado de Tranfield, Denyer e Smart (2003) condizente com o rigor metodológico exigido na revisão de literatura, que resultou na composição do corpus da pesquisa. Avaliou-se o corpus da pesquisa de maneira detalhada valendo-se dos pacotes bibliométricos HistCite, VOSviewer, pacote R e Iramuteq. Avaliou-se a validade das 3 leis bibliométricas: Lei de Lotka, Lei de Bradford e Lei de Zipf com base nas técnicas de citação e cocitação de autores. Este estudo pode ser considerado como relevante e original em função de ser a primeira Revisão Sistemática de Literatura sobre cartas de controle para processos autocorrelacionados. Os resultados sugerem a confirmação das 3 leis clássicas bibliométricas para o corpus pesquisado. Como implicação prática, esta revisão serve de apoio para que os acadêmicos da área de gestão de qualidade possam desvendar potenciais lacunas de pesquisa. A principal limitação do trabalho refere-se a composição do corpus textual, dado que bases de periódicos nacionais não foram consultadas em função da dificuldade em termos de compatibilidade dos softwares usados nesta pesquisa.

Palavras-chave: Controle estatístico de processos; Dados autocorrelacionados; Gestão da Qualidade

1. INTRODUCTION

The production processes evaluation has created challenges for researchers and decision-makers, because production maintains an accelerated pace, to meet the needs and demands of internal and external customers. Quality measurement makes it possible to use the most diverse types of tools to make products and processes more and more efficient.

Many problems generated are due to intense production and small batches, in which changes in machinery, projects, and specialized people must accompany the request, demand, and requirements imposed by the project. Although Shewhart's control charts, developed in 1924 and disseminated in 1931, are widely used today, they are useful for the evaluation and monitoring of mean and variability, requiring the assumptions for their application, as Benneyan (1998) points out.

According to Montgomery (2004), Marcondes Filho, Fogliatto, and Oliveira (2011), and Faisal *et al.* (2018) control charts are a useful technique for improving productivity, being the main tools for performing statistical process control. They

are effective at preventing defects, prevent necessary process adjustment, and provide diagnostic information and process capability.

On Osei-Aning, Abbasi, and Riaz's (2017) view the modified control charts are adjusted to compensate for the influence due to autocorrelation. Modified moving averages and cumulative sums charts, through exponential ponderation, are two types of charts widely used for autocorrelated data monitoring. As those charts have design parameters, their formulation and choice play a significant role in detecting out-of-control situations.

Autocorrelated processes can be treated with the use of control charts based on exponentially weighted moving average statistics (EWMA), initially presented in the 1950's by Roberts (1959), as a generalization of arithmetic moving average graphs. The accumulated sum graphs (CUSUM) were initially proposed in England by Page (1954) and have been studied by many authors. Ewan (1963) outlined many control chart schemes and the type of process in which CUSUM charts are most used.

Leoni, Costa, and Machado (2017), evaluated the effect of autocorrelation through the T^2 control chart of Hotelling, verifying the effect of an anomaly in processes. They concluded that the autocorrelation violation hypothesis must be verified before the control chart is used since the anomaly presence affects the performance of traditional charts, which reduces the ability to detect deviations apart from the average.

This research aims to present a literature review on the control charts for autocorrelated processes, to contribute to scientific knowledge in the area of quality management. The main motivation for conducting this study lies in the need to consolidate this type of control chart used in a more relevant way in the evaluation of production processes susceptible to autocorrelated data. Another motivating is to seek to understand the particularities of this tool that generates diagnoses of process analysis, considerably with greater assertiveness, when compared to case letters without the presence of autocorrelation.

The rigorous research protocol for the literature review was systematized, which resulted in the composition of the research corpus. The knowledge dissemination was analyzed in two phases, one with a detailed synthesis of the corpus general characteristics and the other and the other one focused on the content analysis based on clusters that emerged in the study themes.

This study is relevant and original as it is the first literature review on control charts for autocorrelated processes. When searching the web of science journals database, there were no reviews that deal with production processes where the data are autocorrelated. Therefore, this is a theme where it demonstrates the need to prepare this review, confirming its originality.

The theoretical contribution consists in mapping the field characteristics and analyzing the studies based on a specific research problem. The contribution consists of offering evidence on the results of the studies so that they can help process managers in decision-making, creating strategies for cost optimization oriented to competitiveness, by the findings in the studies that make up the research corpus.

Based on the research purpose, as well as on the need to elaborate a literature review, the following central research question was asked: "In the past twenty years where statistical process control charts for autocorrelated process been applied?"

Considerations made previously, the following secondary questions arise: Who are the most cited authors? Which countries are the authors from? What are the main journals? What are the main themes and keywords identified?

This study is divided into four sections, which can be summarized as follows: the first one refers to the introduction; the second one brings the methodological procedures adopted; the third one concerns the presentation and analysis of results; the fourth one deals with the final considerations, limitations, and recommendations for future research.

2. METHODOLOGY

The objective is to map the studies published on the Web of Science, to contribute to scientific knowledge from a research question that guides the analysis. The intention of carrying out this study is to search the literature on quality management based on the following research question: "To what extent are the statistical process control charts that have been applied in the last twenty-two years?"

The LR was conducted based on the proposition of a research protocol, as demonstrated in Figure 1, which is aligned with the methodological rigor of an LR, as a way to present legitimacy and offer a structured knowledge and base for decision-makers, Tranfield, Denyer & Smart (2003).

2.1 First stage: review planning

The first stage refers to the revision planning, which consists of the orientation of specialists on the main theme, as well as the research *corpus* construction process, using the concept of Aarts (1991) as a selected and organized textual set that expresses a certain sense of language. In this LR the *corpus* seeks to extract attributes quantitatively, as well as qualitative representations from the analyzed content, Bauer and Aarts (2010). This phase comprises the research protocol elaboration, as a way of assigning objectivity to the steps performed and their descriptions, Tranfield, Denyer, and Smart (2003). The protocol elements are research question, population, and sample, the strategy adopted for study inclusion and exclusion.

2.2 Second stage: driving LR

The second stage refers to a comprehensive and impartial search conducted through the keywords identification and searched terms related to the research

question raised in this study. The articles were searched in the *Web of Science* database, which covers the databases as *Science Citation Index Expanded*, *Social Sciences Citation*, *Arts & Humanities Citation Index*, *Conference Proceedings Citation Index (Science)*, *Conference Proceedings Citation Index*, *Social Science & Humanities*, *Emerging Sources Citation Index*, *Current Contents Connect*, *Derwent Innovations Index*SM, *KCI*, *Russian Science Citation Index*, *SciELO Citation Index*, *Cambridge University Press*, *Elsevier*, *Springer*, *Wiley-Blackwell*, and *Nature Publishing Group*.

The articles were located from the following keywords: *control charts and autocorrelated processes*. The syntax used during the search process in the *Web of Science* database was TS = ("control charts" AND "autocorrelated process").

It was defined as inclusion criteria: the presence of these words in article titles; the search for articles published in periodicals; English as the document language. This strategy was thus defined by extracting articles that effectively analyzed the proposed relations, the title being a support criterion, and the composition of articles in the industrial engineering area where 46 articles were located in the *Web of Science*.

From the previous criteria, articles were found linked to relevant sources such as the *Journal of Quality Technology*, *Quality and Reliability Engineering International*, *Computers & Industrial Engineering*, *Technometrics*, *Applied Mathematics and Computation*, *Asia-pacific Journal of Operational Research*, *Expert Systems with Applications and Information Sciences*. For the search, a time period was not defined for the choice of articles, which made it possible for all studies published within the investigated theme, to have the same chance of being located, composing a probabilistic stratum. After the inclusion of the criteria through the filters, the sample comprised a total of 18 articles from the *Web of Science*.

The abstracts and the introduction were read individually for each article and after finding that some studies were not adherent to the theme, they were

excluded. This procedure confirmed its adequacy to the research question in the study of the relationship between control charts and autocorrelated processes. Thus, the corpus of the research was 18 articles from the *Web of Science*. These studies were in agreement with the criteria previously defined in this LR, which were reviewed using the *face-validity* technique by 3 specialists in the area.

The research corpus was compiled in an electronic spreadsheet, highlighting the elements of each article individually that contribute to the realization of RSL. Citation and content indicators were extracted from the articles. The citation indicators were listed in 7 elements: the year in which the article was published, linked journal, the title of the article, number of citations of the article, name of the authors of the article, number of authors, and country of performance of the authors. The content indicators included 7 elements analyzed in the studies, which were listed in: keywords cited in the article, objective, contribution, theme related to letters of control of autocorrelated processes, collection procedure, main results, and limitations of the research.

2.3 Third stage: dissemination of knowledge

The third stage represents the LR results dissemination. These results represent the group synthesis, which consists of generating knowledge conducted in two stages, Tranfield, Denyer, and Smart (2003).

2.3.1 The first stage of the third stage: detailed analysis

The first stage of the Third Stage, referring to the dissemination of knowledge, provides a detailed analysis of the characteristics of the articles that make up the research corpus, elaborated using an electronic spreadsheet, HistCite, VOSviewer, Rstudio, and Iramuteq software.

In the first stage, a detailed analysis of the general characteristics of the 18 articles in the corpus was presented, containing: the relationship between the

authors that make up the corpus, the relationship between the most mentioned authors, the relationship between the most closely coupled authors, the most relevant words in the title, abstract and words of the corpus and the relationship between the keywords of the corpus.

2.3.2 The second stage of the third stage: an in-depth analysis

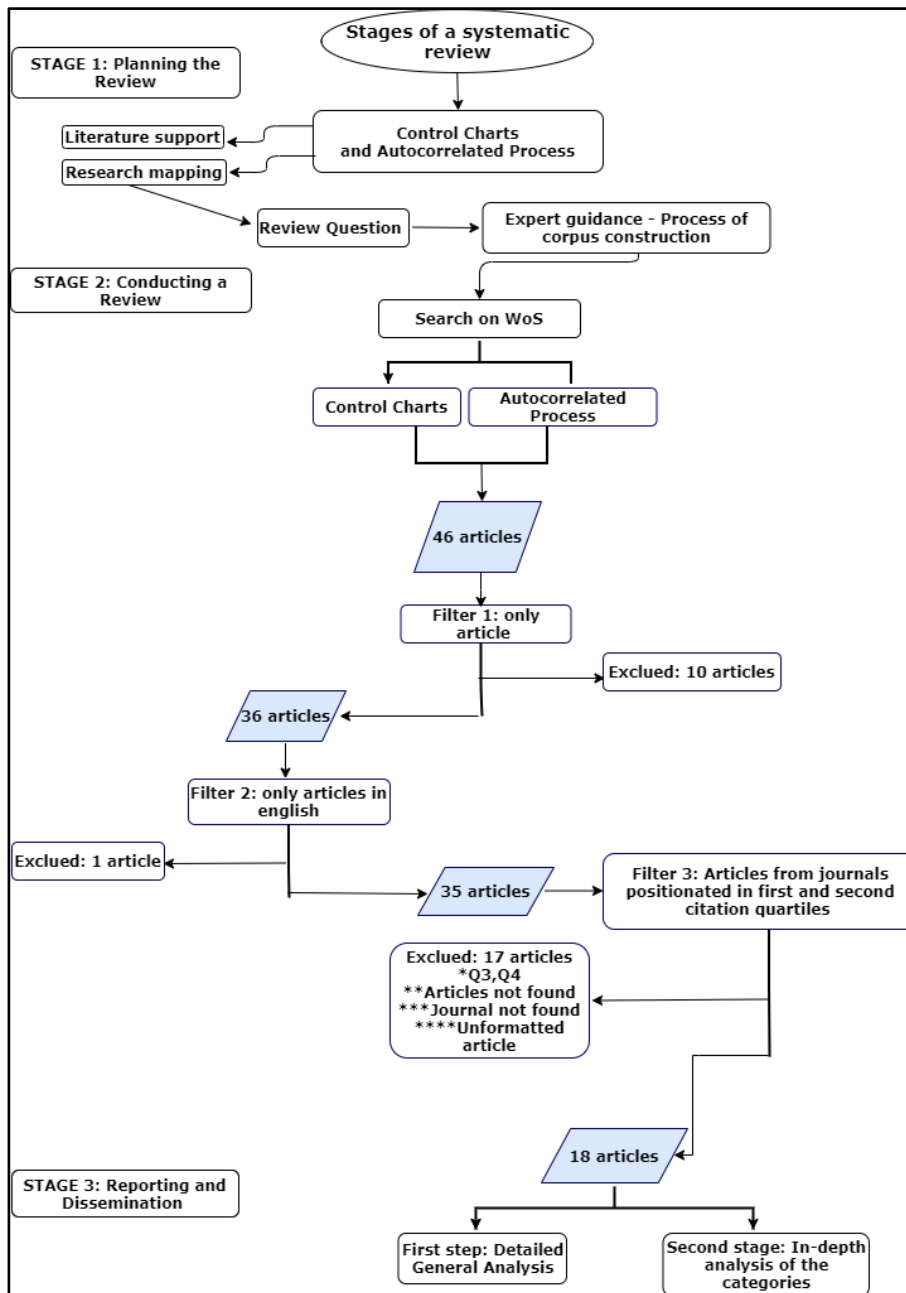
The second stage of the Third Stage, referring to the dissemination of knowledge, consists of conducting an in-depth analysis of the textual corpus, referring to the thematic analysis. This stage was constructed by identifying clusters or categories of analysis that represented similar research themes found in the 18 articles.

An analysis is carried out with the most frequent variables used in the studies to estimate the relationship proposed in the research question. This stage is driven by the concentration of a shared consensus between different themes that make up the textual corpus. The creation of clusters and categories encompasses a detailed description of their contributions, highlighting relevant parts of the corpus, based on the literature.

2.4 Graphic presentation of the research protocol

The construction of the LR research protocol and its generated results contribute to mapping the studies referring to control charts. Figure 1 presents the research protocol systematization.

Figure 1 - Systematization of the research protocol



Source: The authors (2020)

Figure 1 presents relevant evidence about control charts for autocorrelated processes, which is in line with the principles of literature review defined by Tranfield, Denyer, and Smart (2003) and seeks to analyze all the theoretical evidence from the *corpus* articles.

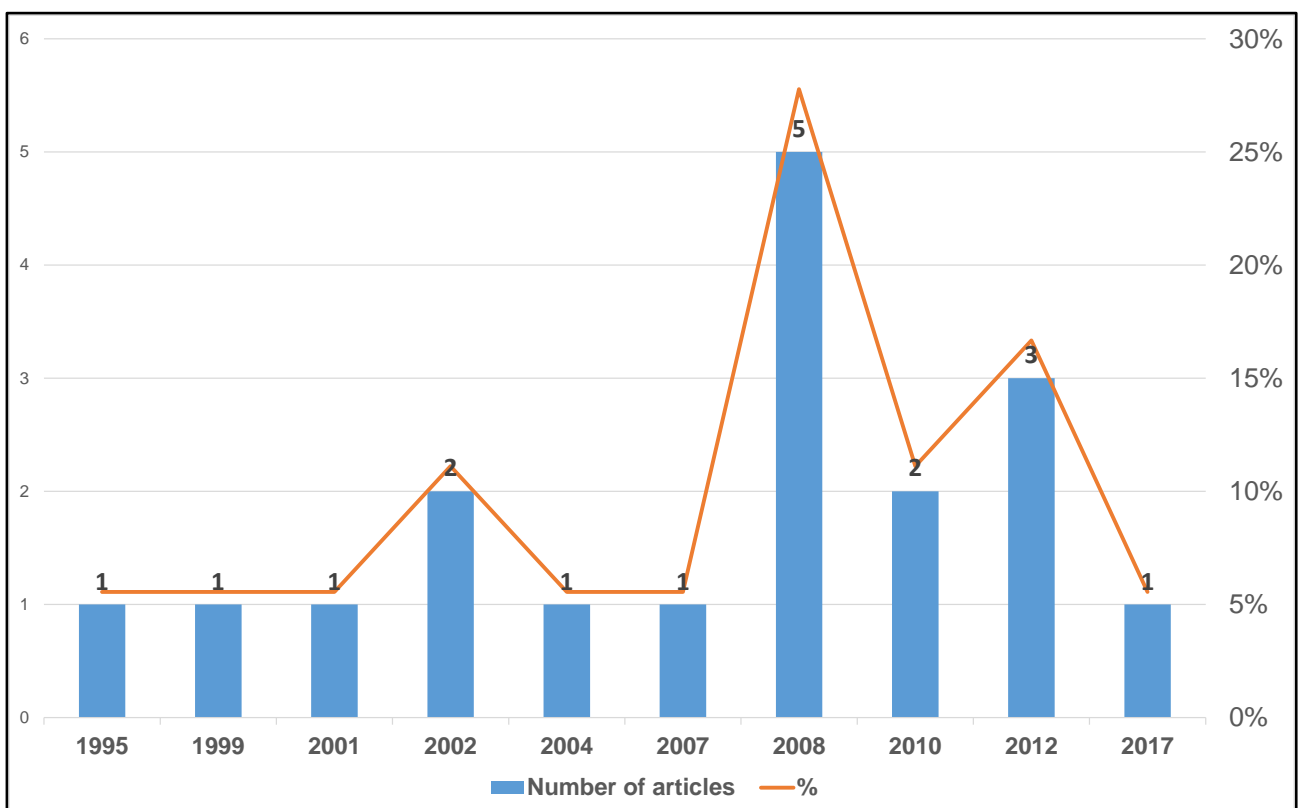
3. DATA PRESENTATION AND ANALYSIS

This section aims to present the articles' analysis results of the *corpus*, based on 8 periodicals and 39 authors and co-authors, selected in the *Web of Science* database. The authors' productivity and organizations were evaluated, as well as the journals' productivity and the word occurrence frequency or word co-occurrence in a text.

3.1 Number of articles published per year

Figure 2 shows the behavior over the years of the number of articles published, as well as the relative participation of the *corpus* articles.

Figure 2 – Quantity of documents published per year



Source: The authors (2020)

It can be seen in Figure 2 that the period between 1995 and 2002 does not show the absence of variability in the number of articles published, remaining

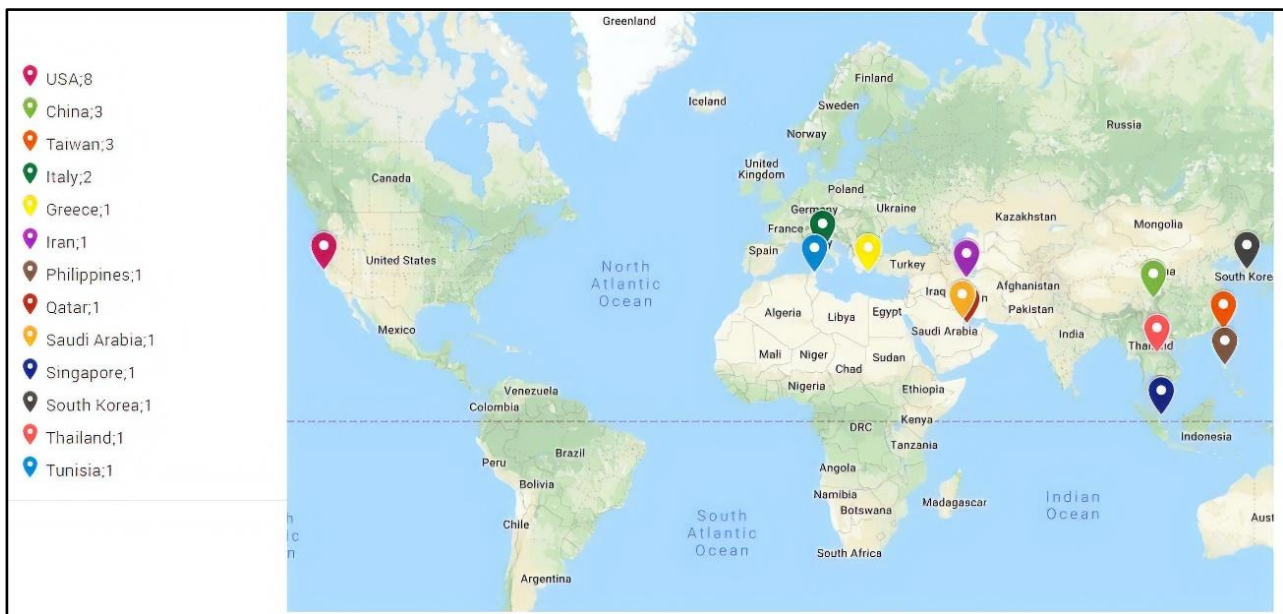
between 1 and 2 publications over the period, which corresponds to 28% of the period's production. In 2008, there was a sudden increase in the number of articles published, from 1 to 5 documents, which corresponds to the accumulation of 67% of works published in the period between 1995 and 2008. Between 2010 and 2017, there was a drop in the number of documents published, corresponding to 6 documents, corresponding to 33% of the total number of articles in this period analyzed.

It stands out in the period 1995 to 2017 the works developed by the author's Lu and Reynolds, which was published in 2001, entitled Cusum charts for monitoring an autocorrelated process, in the Journal of quality technology, where it obtained a total of 68 citations received, in addition of the work developed by Runger and Willemain, published in 1995, entitled as Model-based and model-free control of autocorrelated processes, in the Journal of Quality technology, where it obtained a total of 66 citations received.

In this sense, it is possible to confirm the validity of Lotka's Law that measures the productivity of authors, since 97% of the authors, corresponding to 38 authors publish only 1 article while 1 author corresponding to 3% publishes 3 articles.

3.2 Geographical distribution of authors

The geographic distribution of the 18 articles comprising the textual corpus in the period considered was analyzed. A total of 25 citations were found from authors who had their affiliations in institutions in the 13 countries evaluated geographically, as can be seen in Figure 3.

Figure 3 - Geographical distribution of the authors cited in the textual *corpus*

Source: The authors (2020)

Based on Figure 3, it can be seen that the United States has 8 authors who declared they have ties to American institutions, which corresponds to 32% of the total number of authors and co-authors. Then there are China and Taiwan who came in second place with 3 authors in each linked country, which corresponds to 12% of the corpus and Italy which obtained 2 authors linked to Italian institutions, which corresponds to 8%.

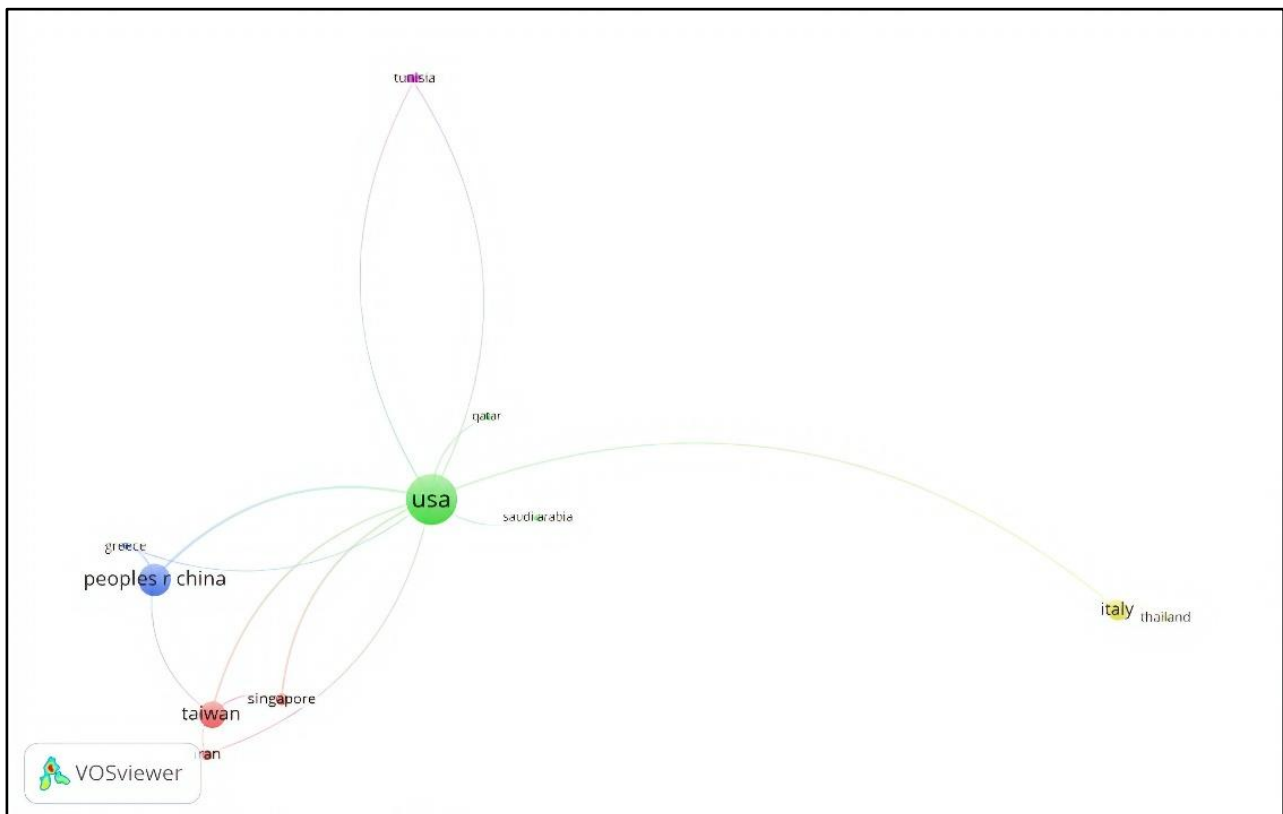
The remaining countries, Greece, Iran, the Philippines, Qatar, Saudi Arabia, Singapore, South Korea, Thailand, and Tunisia, show only 1 author in each country with institutional links, corresponding to 4% institutions related to the total of the institutional link shown.

It should be noted that the four most-cited countries through their institutional links (USA, China, Taiwan, and Italy) correspond to 16 citations of the total of the 25 citations of the countries in the most diverse articles, identified by their links, which corresponds to 64% of the total citations of the countries while making this survey. Some authors have different institutional links from different countries, which means that one author can represent more than one country. But,

in this study, only the first author of each paper on the textual *corpus* was considered.

It can be seen in Figure 4 the main collaborations between countries, formulated based on the VOSviewer software and, considering as an ex-ante parameter, the formation of clusters, the fact that the country appears in the document once and that it has been mentioned at least 2 times. In all analyzes of clusters, throughout this work the clusters will be identified according to the colors as follows: Cluster 1: red color; Cluster 2: green color; Cluster 3: dark blue color; Cluster 4: yellow and Cluster 5: purple.

Figure 4 - Map of citations between countries in the textual *corpus*



Source: The authors (2020)

The formulation of the 5 *clusters* in terms of citation between countries is perceived based on Figure 4. The green *cluster* stands out as the one that has more intensity in terms of scientific collaboration that is the United States. The first red *cluster's* network refers to the following countries: Iran, Singapore, and Taiwan,

which obtained a total of 113 citations and 10 relational links with the other countries.

The second *cluster* of green color includes the following countries: United States, Qatar, and Saudi Arabia, which had a total of 285 citations among all its members and 17 relational ties, with the United States with 46.91% of the total citations received.

The third dark blue *cluster* consists of the countries: Greece and China with 120 citations and 7 relational links.

The fourth yellow *cluster* is composed of Italy and Thailand, which had a total of 55 citations and 4 relational links with the other countries of the network.

Finally, the fifth and last *cluster* of purple color comprehends the countries: Tunisia and South Korea, with 23 citations and 4 relational links with the other connected countries. Among the 13 countries in the sample, only the Philippines has no link with the other countries, so it does not appear in the above *cluster*, the country has only 3 citations.

3.3 Authors' productivity Analysis

This section aims to show the results obtained from the citation indicators extracted from the *corpus* articles to evaluate the authors' productivity based on Lotka's bibliometric law.

3.3.1 Evaluation of authorship composition

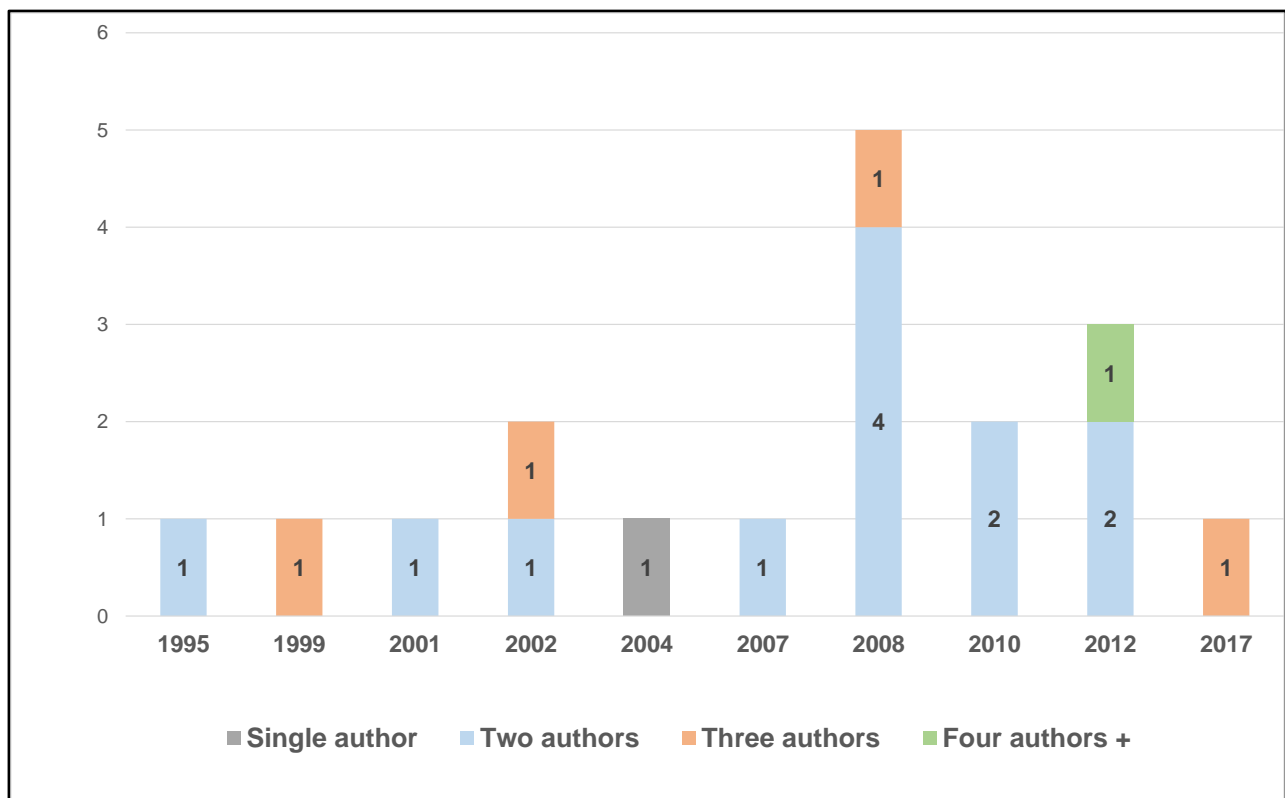
To evaluate the authorship composition of the *corpus* articles, the number of authors of each of the articles was extracted individually, aiming at quantifying the tendency of publications made in partnerships.

In the views of Katz and Martin (1997) and Olmeda-Gómez *et al.* (2009), the composition of authorship refers to how contemporary researchers share data, equipment or ideas in a particular project and brings benefits such as specialization

or division of tasks among the participants of the work, such as reducing the time of involvement, improvement of the quality of the article, besides the reduction of the uncertainties regarding the acceptance of the article due to the diversity among the authors.

Figure 5 shows the authorship composition, taking into account the authorship decomposition with 1, 2, 3, and 4 or more authors.

Figure 5 - Composition of authorship of the textual *corpus*



Source: The authors (2020)

In Figure 5 the composition of authorship of articles in the corpus over time. We observed that 94.44% of the evaluated works, corresponding to 17 articles, have more than 1 author, which denotes an intensity of partnerships in terms of research over time with other co-authors.

In this sample, 1 work was obtained with only 1 author, which is equivalent to 5.56% of all evaluated works. Therefore, it is assumed that several studies analyzed from the textual corpus, come from reflections resulting from the maturity of researchers about the practice of doing research.

It was also noticed that 12 evaluated works were published with 2 authors, which represents 66.67% of the corpus, while the co-authorship with 3 authors includes 4 articles, which corresponds to 22.22% of the entire corpus, while the co-authorship with 4 authors obtained only 1 article, representing 5.56% of the sampled articles.

For the elaboration of the previous figure, the first author is considered to be the one who has the greatest contribution in terms of having devoted more time to work and so on.

3.3.2 Authors' citation analysis

According to Vanz (2004) and Romancini's (2010) view, the citations allow the investigation and visualization of the relationship between citing document and cited document, providing transparency in the subjects addressed in the different areas of knowledge. The analysis of citations can still be seen as a tool to estimate the contribution of the results of the scientific production that they had for a certain area of knowledge. When performing the citation analysis, it is sought to find out which researchers are most cited or more productive.

Through the articles in the *corpus*, it is clear that the most cited article was that of Tsung (1999), entitled as Join monitoring of PID - controlled processes, published in the Journal of Quality Technology, where it obtained a total of 114 citations, which corresponds to 23.75% of the 480 citations received by the 18 articles highlighted in the corpus.

In his work, the author proposed a new methodology aimed at measuring performance and robustness based on a statistical process control algorithm, using joint monitoring schemes based on Hotelling's T2 charts and Bonferroni's approach, that is, refers to a joint monitoring strategy based on controlled PID (Proportional-Integral-Derivative) processes to monitor the process input and output simultaneously. The results aim to overcome potential deficiencies of the

traditional statistical process control charts, by demonstrating that the joint monitoring of the two approaches mentioned above brings effectiveness in terms of the control charts.

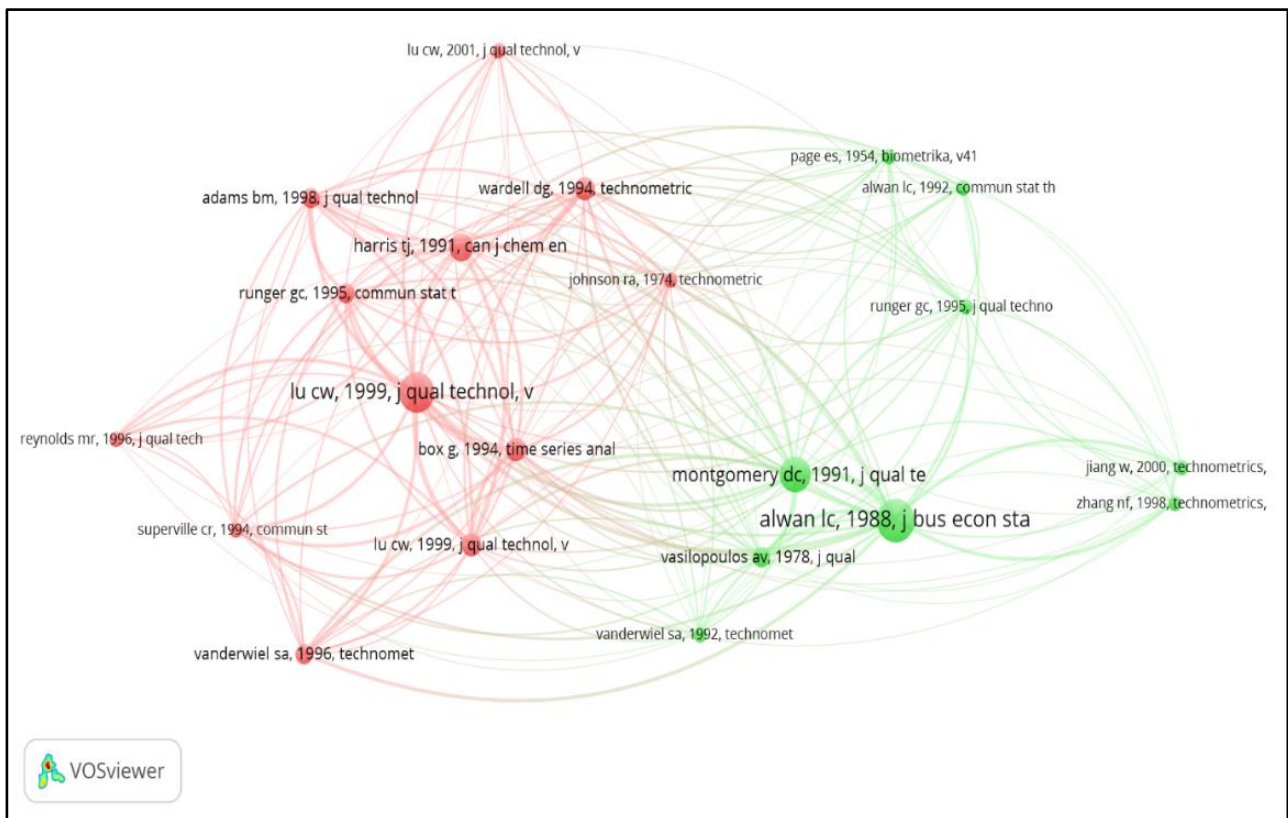
We can also highlight the work developed by Lu (2001), entitled as Cusum charts for monitoring an autocorrelated process, published by the Journal of Quality Technology, in 2001, where he obtained a total of 68 citations, corresponding to 14.17% of the total citations received from the corpus.

In this article the author investigates the Cumulative Sum (CUSUM) graphs based on observations and residues, showing that, for moderate levels of autocorrelation, both types of CUSUM graphs require approximately the same time to detect changes in the process average. For higher levels of autocorrelation, the two types of CUSUM plots (observations and residuals) work similarly for small changes, but the residuals CUSUM is slightly better than the large changes CUSUM.

3.3.3 Analysis of authors' co-citation

For Small (2004) and Grácio and Oliveira (2013), the co-citation analysis can be seen as the frequency with which two or more documents, authors, or areas are jointly quoted in later scientific productions, which demonstrates how a certain knowledge area is structured. In Figure 6, a covariance map is presented, based on 348 cited references in 18 *corpus* articles and elaborated by VOSviewer.

Figure 6 – Author's co-citation Map

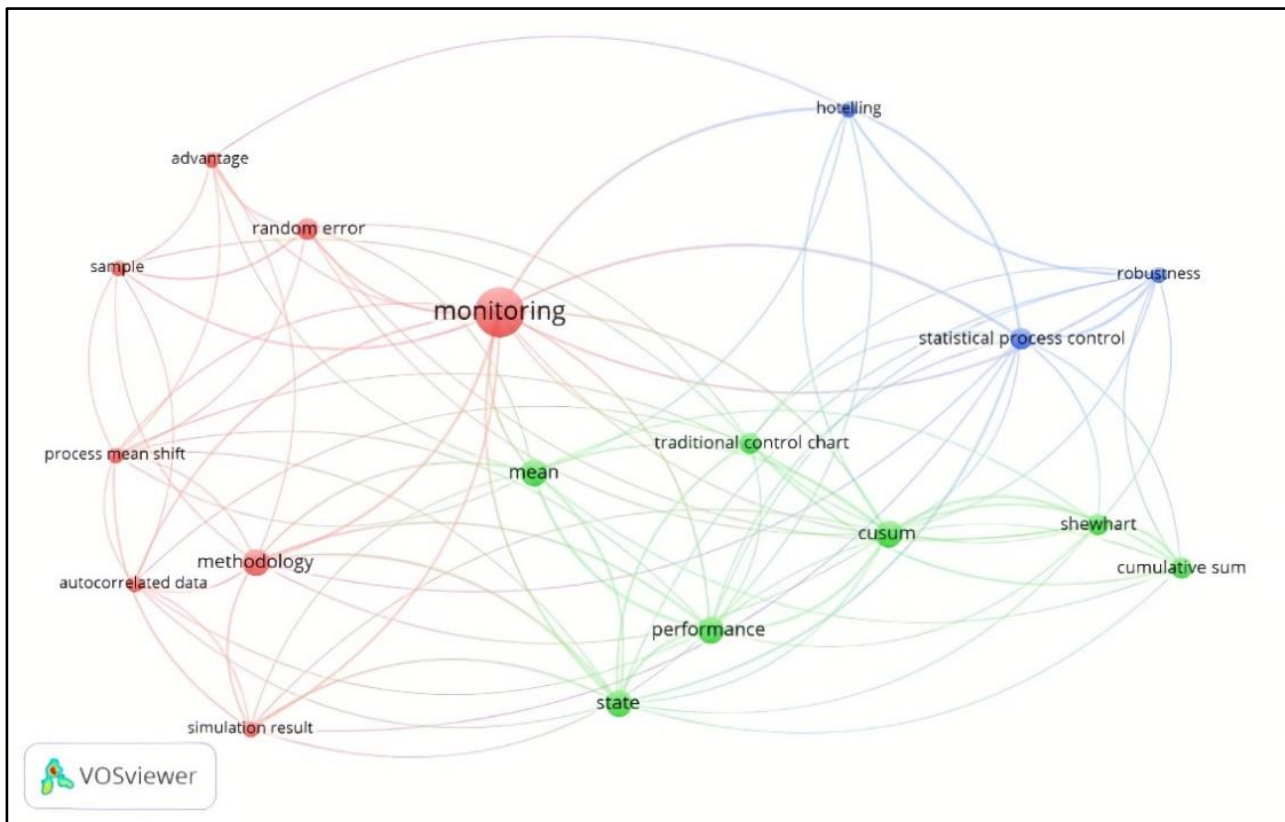


Source: The authors (2020)

It was defined as a unit of analysis the "cited references" and a counting method "full counting", where the same weight is attributed to each of the links between the references co-cited and the formation of the clusters about the size of the circles, colors, and thickness of the line segments is proportional to the frequency of co-citation between the connected authors (Figure 6). The co-citation map evidenced 2 clusters: the first one denoted by the red color (*cluster 1*) that contains 12 elements, while the second one (*cluster 2*) disposed of by the green color where it contains 9 elements.

Figure 7 shows the network of words evidences the relationship between them, made using VOSviewer software.

Figure 7 - Network of corpus words



Source: The authors (2020)

The network of words aims to map the possible themes related to the central theme, that is, a graphical abstract about theme: "control charts for autocorrelated processes" (Figure 7). The strength of the relationship between the nodes verified in the network can be evaluated by the proximity between them. It is possible to formulate the network of words by observing their groupings and the relational links with other words.

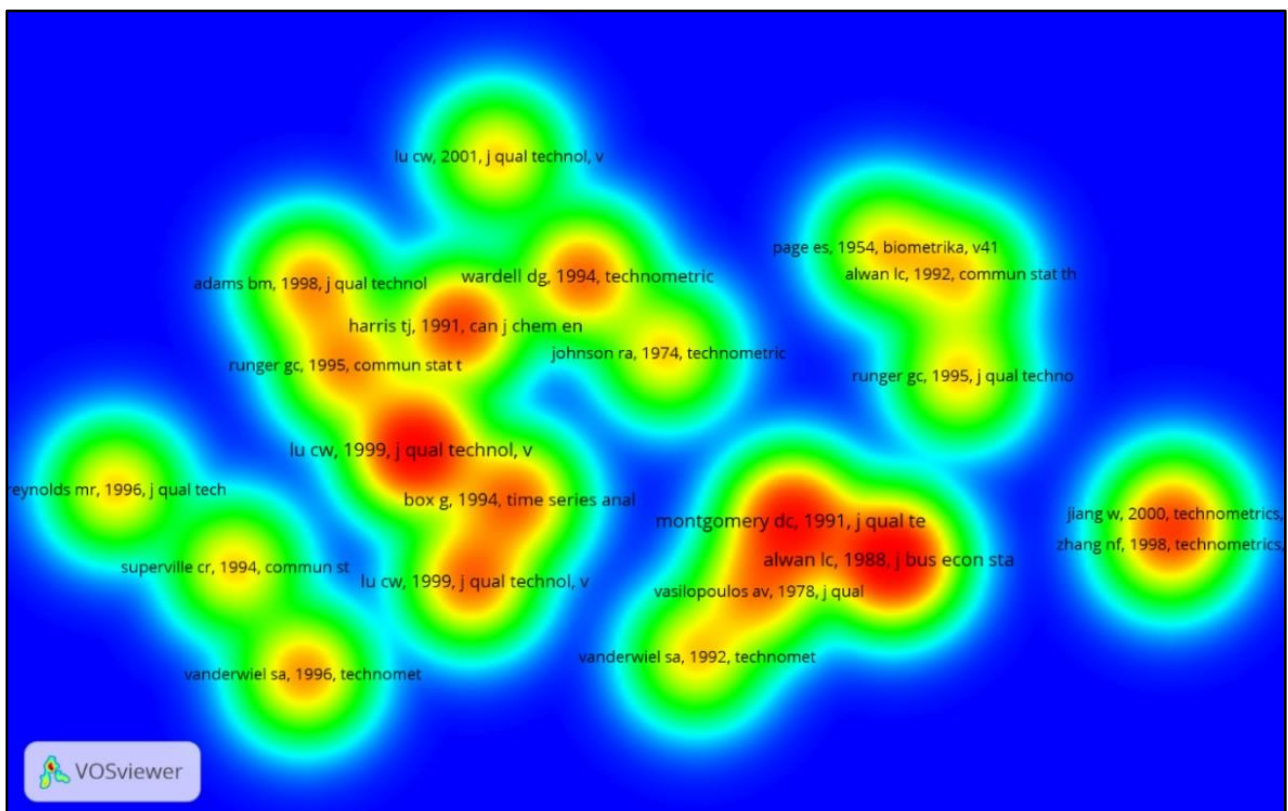
To do this, use the creation of the network as a unit of analysis or the VOSviewer software. Consider the title, keywords, and summary for analyzing the articles in the corpus and building the word network, in addition to the "binary count" counting method. In this same context, consider a minimum frequency of 3 occurrences out of a total of 428 terms, with a reduction of 39 terms, which support the construction of the association network between the terms searched that came from the theme of control of autocorrelated processes. As a consequence, 60% of

the previous 39 terms were selected, a priori, to constitute the network, resulting in a total of 23 terms that will form part of the word network as shown in Figure 7.

The network of words is divided into 3 *clusters*: *cluster 1* of the red color was denominated "Methodology of process monitoring" with 8 items; *cluster 2* of the green color was titled "Letter CUSUM" with 7 items; *cluster 3* of the blue color entitled "Multivariate Process" with 3 items.

Figure 8 shows the density map, highlighting the most prolific authors in this area. This map shows clear and isolated spots whether the authors do not work together or the works in sets are recent with few quotations. On the other hand, whether noticed dark reddish spots represented strong co-citation among the evaluated authors.

Figure 8 - Density map of co-cited references



Source: The authors (2020)

In Figure 8, the minimum number of 4 references was considered, extracting the 21 most cited authors simultaneously and their relational ties, which allows a

better visualization in the density map of the network. It is noteworthy that in the density map extracted from the networks, the analysis of co-citation considered the first author to be the main author of the work.

3.4 Journal productivity and reputation analysis

The journal productivity and reputation evaluation, verifies Bradford's law, for the collected sample. These analyzes are important to show both the functioning and the volume of articles published in the sampled period, revealing the dynamic behavior in terms of the theme contributions density.

In this study, we tried to answer the secondary question of research, based on the results of Table 3: *which are the main journals in the corpus in the researched area?* Based on Alabi's (1979) view, Bradford's Law states that the decreasing order of productivity of articles on a given subject in scientific journals makes it possible to establish exponentially divided groupings. In this sense, the journal number in each grouping that is proportional to $1:n:n^2$. Thus, by measuring the journals' productivity, it is possible to establish the nucleus and areas of dispersion on a given subject in the same set of journals.

In Table 1, several bibliometric indicators related to the productivity and reputation of the *corpus* journals are journal-title, the quartet of Scimago citations, the number of articles published in the textual *corpus*, H-index score, Journal Citation Report (JCR), the country and Publisher corresponding to each *corpus* journal, based on the Bradford bibliometric law that estimates the degree of journals relevance for a certain area of knowledge.

Table 1 - Analysis of Productivity and Reputation of the Journal

n	Title	SJR Quartile	Zone	Quantity Articles	SJR Index	H Index	JCR (2017)	Country	Publisher
1	Journal of Quality Technology	Q1	Zone 1	6	1,814	73	2.306	United States	American Society of Quality
2	Computers & Industrial Engineering	Q1	Zone 2	2	1,463	103	3.195	United Kingdom	Elsevier
3	Technometrics	Q1	Zone 2	2	1,546	69	1.568	United Kingdom	United Kingdom
4	Quality and Reliability Engineering International	Q2	Zone 2	4	0,955	46	1.604	United Kingdom	John Wiley & Sons
5	Expert Systems with Applications	Q1	Zone 3	1	1,271	145	3.768	United Kingdom	Elsevier
6	Information Sciences	Q1	Zone 3	1	1,635	142	4.305	Netherlands	Elsevier
7	Applied Mathematics and Computation	Q1	Zone 3	1	1,065	117	2.300	United States	Elsevier
8	Asia-Pacific Journal of Operational Research	Q2	Zone 3	1	0,477	26	0.558	Singapore	World Scientific Publishing

Source: The authors (2020)

Related to the reputation of the selected journals set out in Table 1, we can see that the 13 articles were published in journals located in the first quartile (Q1) of quotations from the Scimago Journal Ranking, according to the SJR (2018) that is the 25% of the most cited articles, measuring the visibility and prestige of journals. The other 5 articles were published in journals in the second quartile (Q2) of quotations from the Scimago Journal Ranking (SJR), which corresponds to 50% of the total articles.

Taking into account the 18 journals of the textual corpus, these were distributed in 3 zones (Zone 1, Zone 2, and Zone 3). The first zone corresponds to

33.3% of the *corpus* articles, with 6 articles published in only 1 journal that corresponds to 12.5% of the 8 journals present in the work. In the second zone, 8 articles represent 44.4% of the *corpus* of the research, divided into 3 journals equivalent to 37.5% of the total. Finally, the third zone with 4 works, which represent 22.2% of the *corpus*, divided into 4 journals, corresponding to 50.0% of the total.

About H-index, created by Hirsch (2005), in Table 1, show the six journals with the highest indicators were Journal of Quality Technology, Computers, and Industrial Engineering, Technometrics, Applied Mathematics and Computation, Expert Systems with Applications and Information Sciences. To interpret the described H-index results, it was taken as an example the case of the journal Expert Systems with Applications (H-index = 145) as this journal obtained 145 of its articles cited at least 145 times, a high reputation for such a scientific journal. Although it is a measure widely used to measure the author and journal quality, there is no consensus around its use leading to questions from the scientific community. Thus, the questionnaire on the productivity of the evaluated journals was answered.

Regarding the distribution by the country where the investigated journal is hosted, it was possible to verify that the 8 journals are distributed in 4 countries: United States (2 journals), United Kingdom (4 journals), Netherlands (1 journal) and Singapore (1 journal).

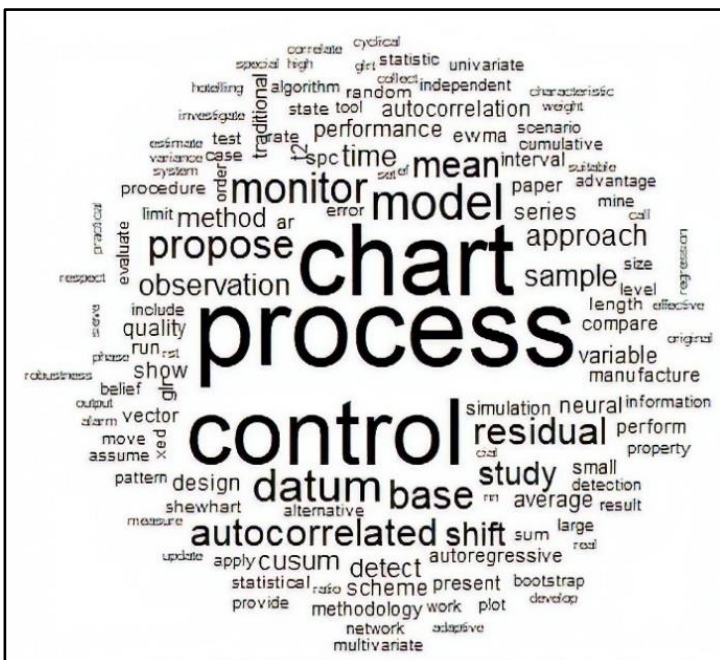
When evaluating the publishers contained in Table 3, it was observed that there is a sprinkling in terms of the number of publishers, given the presence of 5 publishers in the *corpus*: Elsevier (4 journals) corresponding to 50% of the market, American Society of Quality, Taylor & Francis, John Wiley & Sons and World Scientific Publishing with 1 journal each, corresponding to 12.5% of the journals total periodicals evaluated.

3.5 Word Cloud and Word Network

Clouds of words describe representations of a visually hierarchical list for classification purposes, where the greater or lesser number of appearances of a given term is given proportionally by the size of the source. The words are randomly positioned so that the most frequent words appear larger than the others, thus evidencing their prominence in the textual *corpus*. Besides, it makes it possible to identify the co-occurrences among the words and their result gives indications of the connection between them, aiding in the identification of the representation structure based on Zipf's law (ZIPF, 1945).

Figure 9 shows the abstract word cloud formulated through Iramuteq software.

Figure 9 - *Corpus* abstracts word cloud



Source: The authors (2020)

It is possible to highlight the words "process", "chart" and "control" in Figure 9. The word "process" has 84 occurrences of the total of 1127 occurrences, representing 7.45% of total words which most appear in the *corpus* abstracts, "chart" obtained 75 occurrences, which represents 6.65% of the total and "control" had 69 occurrences, corresponding to 6.12% of the total.

It is worth emphasizing that the greater and more centralized a particular word in the cloud, the greater the degree of its evocation by the subjects. Conversely, the further apart and smaller the size the less will be, Camargo and Justo (2013).

In Table 2 we present the Top 10 keywords occurrences for the word network seen in Figure 7.

Table 2 - Top 10 of Keyword Count

Terms	Ocorrences	%	Relevance
Comulative Sum	4	11,76%	2,45
Robustness	3	8,82%	1,61
Shewhart	4	11,76%	1,60
Sample	3	8,82%	1,36
Hotelling	3	8,82%	1,21
Random Error	4	11,76%	1,10
Autocorrelated data	3	8,82%	1,09
Statistical Process Control	4	11,76%	1,05
Advantage	3	8,82%	1,03
Process Mean Shift	3	8,82%	1,02

Source: The authors (2020)

In table 2 shows the occurrence and relevance of the keywords used to formulate the *clusters*. Also, the 3 words with the greatest appeal in terms of relevance were "Cumulative Sum", "Robustness" and "Shewhart", which means that these words have a higher relevance score about the other words present in the *corpus*. These words obtained 4 occurrences, representing 11.76% of the total of the 10 main words found in the *clusters*, and the word "Robustness" obtained 3 occurrences corresponding to 8.82% of the total.

4. FINAL CONSIDERATIONS

This work aimed at performing a bibliometric analysis of control charts for autocorrelated processes. For this, we sought support in the process of literature review, where the primary articles were mapped in quality management, specifically in the area of statistical process control.

The analysis was based on the three classical laws of bibliometrics: Lotka's Law, which evaluates authors' productivity; Bradford's Law, which measures the productivity of periodicals; and Zipf's Law, which measures the frequency of occurrence and co-occurrence of certain words in a text.

The *corpus* descriptive analysis revealed that the authors and co-authors are distributed in 13 countries, with emphasis on the United States. The period of analysis comprises the years 1995 to 2017, highlighting the year 2008 with 5 articles published.

The basic assumption of Lotka's Law - that there is a small elite of more proficient researchers while the vast majority publish little - in this sense, one can confirm the validity of Lotka's Law for a selected sample of authors, since 97.4 % of authors (38 authors) published only 1 document while 1 author published 3 documents. Regarding the co-citation, the following authors stand out: Lu (1999), Box (1994), Alwan (1988), and Montgomery (1991) with 10, 6, 11, and 9 citations, respectively.

About the journal distribution, the results confirm the Bradford's Law, given that the Journal of Quality Technology was positioned in zone 1, corresponding to 12.5% of the total journal sampled.

According to the assumptions of the Zipf Law, of a total of 1127 occurrences, the words with the highest frequency of occurrence were "process", "chart" and "control", which suggests their measurement and consequent confirmation through the elaboration of the word cloud and building your network.

Regarding the grouping of keywords, the co-occurrence network formulated contemplated 3 clusters. The three words with the greatest appeal in terms of relevance were "Cumulative Sum", "Robustness" and "Shewhart", which

represented that these words had a score of superior relevance about the other words present in the *corpus*.

The keywords *clusters* used as a basis for establishing the relationship among the terms referring to control charts for autocorrelated processes, 3 *clusters* were formed according to subjects that are interconnected. The subjects were *Methodology of process monitoring* - (*cluster 1*), which involved procedures that are related to process monitoring; *CUSUM Chart* - (*cluster 2*), in which it contained the types of process control *charts*; (*cluster 3*), which indicated the improvement of the process when observing the performance, as well as the state of the process.

The study main limitations were the *corpus* composition and the keyword selection for the word network *clusters*.

This study contributed to identifying the research *corpus* on control charts for autocorrelation processes and offered the major study in this theme and the three most prominent investigation areas.

ACKNOWLEDGMENTS

We thank the Coordination for the Improvement of Higher Education Personnel (CAPES, Brazil) and Research Incentive Fund (FAPESP - Senior) for financial support, the Statistical Analysis and Modeling Laboratory (LAME) of the Federal University of Santa Maria for technical support and the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

- AARTS, J. Intuition-based and observation-based grammars. In: **English corpus linguistics**. Routledge, 1991. p. 44-62.
- ALABI, G. Bradford's law and its application. **International Library Review**, v. 11, n. 1, p. 151-158, 1979.

ALWAN, L. C.; ROBERTS, H. V. Time-series modeling for statistical process control. **Journal of Business & Economic Statistics**, v. 6, n. 1, p. 87-95, 1988.

BAUER, M. W.; AARTS, Bauer. A Construção do Corpus: Um Princípio para a Coleta de Dados Qualitativos in Pesquisa Qualitativa com Texto, Imagem e Som: um manual prático. **Martin W. Bauer, George Gaskell (orgs)**, v. 10, 2010.

BENNEYAN, J. C. Use and interpretation of statistical quality control charts. **International Journal for Quality in Health Care**, v. 10, n. 1, p. 69-73, 1998.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. Time series analysis, forecasting and control. Englewood Cliffs. 1994.

CAMARGO, B. V.; JUSTO, A. M. IRAMUTEQ: um software gratuito para análise de dados textuais. **Temas em psicologia**, v. 21, n. 2, p. 513-518, 2013.

EWAN, W. D. When and how to use cu-sum charts. **Technometrics**, v. 5, n. 1, p. 1-22, 1963.

FAISAL, M. *et al.* A modified CUSUM control chart for monitoring industrial processes. **Quality and Reliability Engineering International**, v. 34, n. 6, p. 1045-1058, 2018.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. de. Análise de cocitação de autores: um estudo teórico-metodológico dos indicadores de proximidade, aplicados ao GT7 da ANCIB. **LIINC em Revista**, p. 196-213, 2013.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National academy of Sciences**, v. 102, n. 46, p. 16569-16572, 2005.

KATZ, J. S.; MARTIN, B. R. What is research collaboration?. **Research policy**, v. 26, n. 1, p. 1-18, 1997.

LEONI, R. C.; COSTA, A. F. B.; MACHADO, M. A. G. A presença da autocorrelação no gráfico de controle T2 de Hotelling. **Sistemas & Gestão**, v. 12, n. 3, p. 252-9, 2017.

LU, C.; REYNOLDS JR, M. R. EWMA control charts for monitoring the mean of autocorrelated processes. **Journal of Quality technology**, v. 31, n. 2, p. 166-188, 1999.

LU, C.; REYNOLDS JR, M. R. CUSUM charts for monitoring an autocorrelated process. **Journal of Quality Technology**, v. 33, n. 3, p. 316-334, 2001.

MARCONDES FILHO, D.; FOGLIATTO, F. S.; OLIVEIRA, L. P. L. de. Multivariate control charts for monitoring non-linear batch processes. **Production**, v. 21, n. 1, p. 132-148, 2011.

MONTGOMERY, D. C.; MASTRANGELO, Christina M. Some statistical process control methods for autocorrelated data. **Journal of Quality Technology**, v. 23, n. 3, p. 179-193, 1991.

MONTGOMERY, D. C. Introdução Ao Controle Estatístico Da Qualidade. Grupo Gen-LTC, 2004.

OLMEDA-GÓMEZ, C.; PERIANES-RODRÍGUEZ, A.; OVALLE-PERANDONES, M. A. Estructura de las redes de colaboración científica entre las universidades españolas. **Revista de información e comunicación**, p. 129-140, 2009.

OSEI-ANING, R.; ABBASI, S. A.; RIAZ, M. Mixed EWMA-CUSUM and mixed CUSUM-EWMA modified control charts for monitoring first order autoregressive processes. **Quality Technology & Quantitative Management**, v. 14, n. 4, p. 429-453, 2017.

PAGE, E S. Continuous inspection schemes. **Biometrika**, v. 41, n. 1/2, p. 100-115, 1954.

ROBERTS, S. W. Control chart tests based on geometric moving averages. **Technometrics**, n. 1, p. 239-250, 1959.

ROMANCINI, R. O que é uma citação? A análise de citações na ciência. **Intexto**, n. 23, p. 5-17, 2010.

SMALL, H. On the shoulders of Robert Merton: Towards a normative theory of citation. **Scientometrics**, v. 60, n. 1, p. 71-79, 2004.

TRANFIELD, D.; DENYER, D.; SMART, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. **British journal of management**, v. 14, n. 3, p. 207-222, 2003.

TSUNG, F. Joint monitoring of the PID-Controlled process. **Journal of Quality Technology**, v. 31, n. 3, p. 275-285, 1999.

VANZ, S. A. de S. A Produção Discente em Comunicação: análise das citações das dissertações defendidas nos programas de pós-graduação do Rio Grande do Sul. 2004.

ZIPF, G. K. The meaning-frequency relationship of words. **The Journal of general psychology**, v. 33, n. 2, p. 251-256, 1945.

Contribuições de autoria

1 – Valentina Wolff Lirio

Graduanda em Estatística pela Universidade Federal de Santa Maria – UFSM.

<https://orcid.org/0000-0002-9868-541X> – valentinawlirio@gmail.com

Contribution: Software, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing

2 – Tailon Martins

Doutorando do Programa de Pós-Graduação em Engenharia de Produção pela Universidade Federal de Santa Maria – UFSM.

<https://orcid.org/0000-0002-5709-6635> - martins.tailon@gmail.com

Contribution: Writing – review & editing

3 – Bianca Reichert

Doutoranda do Programa de Pós-Graduação em Engenharia de Produção pela Universidade Federal de Santa Maria – UFSM.

<https://orcid.org/0000-0003-1741-5440> - bianca.reichert@hotmail.com

Contribution: Writing – review & editing

4 – Gilvete Sylvania Wolff Lirio

Mestra em Engenharia de Produção, Departamento de matemática, Professora na Universidade Regional Integrada do Alto Uruguai e das Missões (URI – Campus Santo Ângelo).

<https://orcid.org/0000-0002-7901-3052> - gil.lirio@san.uri.br

Contribution: Writing – review & editing

5 – Adriano Mendonça Souza

Doutor em Engenharia de Produção, Departamento de Estatística, Professor Titular do Departamento de Estatística, Universidade Federal de Santa Maria – UFSM.

<https://orcid.org/0000-0002-1562-2246> - amsouza.sm@gmail.com

Contribution: Writing – review & editing, Supervision

6 – Wesley Vieira da Silva

Doutor em Engenharia de Produção, Professor permanente do programa de pós-graduação em Administração, Universidade Federal Rural do Semi-Árido (PPGA- UFRSA) e Professor e vice-coordenador do Mestrado em Administração Pública em Redes, Universidade Federal de Alagoas – UFAL.

<https://orcid.org/0000-0001-5354-8676> - wesley.silva@feac.ufal.br

Contribution: Conceptualization, Data curation, Methodology, Software, Supervision, Writing – original draft, Writing – review & editing

How to quote this article

LIRIO, V. W. *et al.* Literature Review on Control Charts for Autocorrelated Processes. **Ciência e Natura**, Santa Maria, v. 44, e2, 2022. DOI 10.5902/2179460X63150. Available in: <https://doi.org/10.5902/2179460X63150>.