Machine Learning-Based Analysis For The Characterization And Phenomenological Study Of Two-Phase Fluids: A Bibliometric Study Using Vosviewer And Scopus

Msc.July Andrea Gomez Camperos[1], Msc. Haidee Yulady Jaramillo[2], Phd. Sir Alexci Suárez Castrillón [3]

[1] Mechanical Engineering Department, Faculty of Engineering, University Francisco de Paula Santander Sectional Ocaña, Research Group on New Technologies, Sustainability and Innovation (GINSTI), Ocaña, Norte de Santander

[2] Department of Civil Engineering, Faculty of Engineering, University Francisco de Paula Santander Sectional Ocaña, Research Group on New Technologies, Sustainability and Innovation (GINSTI), Ocaña, Norte de Santander

[3] Department of Systems Engineering, Faculty of Engineering, Universidad Francisco de Paula Santander Seccional Ocaña, Grupo de Investigación en Ciencias y Tecnología (GRUCITE), Ocaña, Norte de Santander.

*Corresponding Author: Msc.July Andrea Gomez Camperos

Abstract:

Bibliometric analyses have been the primary way of examining and evaluating the literature within a field of study. By focusing on citation count and source, researchers have been able to identify journal articles that are considered high impact in scope and relevance, qualifying them as "citation classics" in a field. In this context, this article analyzed the characteristics of publications related to Deep Learning Based Analysis for the characterization and Phenomenological Study of multiphase fluids. The research was conducted in the Scopus database to identify the academic participation in this topic and the data were analyzed using the VOSviewer software, with a scientific mapping methodology.

Keywords: Two-phase Flow,Machine learning, Scopus; VOSviewe

1. Introduction

The role of the oil and gas industry in society is irrefutably vital, as it is the main driver of the global energy system that contributes significantly to economic and social development. [1] . In the oil industry, produced fluids (oil, salt water and gas) are often transported simultaneously in pipelines. This situation is defined as multiphase flow, which involves the transport of two or more fluids with different properties flowing together within a pipeline [2].

The problem of how to measure mixtures of oil, water and gas has been of interest to the petroleum industry since the early 1980s [3], Fluids with two or more phases in a piping system form flow patterns, which represent the spatial distribution of the phases involved when flowing in a pipe [4].

When two different phases flow simultaneously through any conduit, the two fluids can be distributed in a wide variety of patterns. Based on the inherent characteristics and hydrodynamic parameters, the flow distribution can be broadly classified into different flow regimes or patterns, namely, slow flow, bubble flow, annular flow, stratified flow, dispersed flow, and dispersed flow.[5] .

Recently, with the increase of computing power and the development of advanced computational techniques, machine learning approaches have been increasingly applied in two-phase flow metering[6] .

A powerful tool used for the identification of flow patterns that attracts many researchers is the machine learning technique [7]. Useful information or characters are extracted to train a network that represents the relationship between the measured data and the patterns to be identified.

The most widely used machine learning method for flow pattern recognition is the artificial neural network (ANN)[8]. To obtain effective recognition and prediction results, researchers try to employ different types of topology, such as Feed Forward Neural Networks [9], self-organizing neural network [10] , Multiple Layer Perceptron [11].

However, there are some challenges and bottlenecks in recognizing flow patterns and volume fractions through pipes, so it is necessary to perform a Machine Learning Based Analysis for the Characterization and Phenomenological Study of Two-Phase Fluids.

However, there are some challenges and bottlenecks in recognizing flow patterns and volume fractions through pipes, so it is necessary to perform a Machine Learning Based Analysis for the Characterization and Phenomenological Study of Two-Phase Fluids. [12].

Bibliometric analysis has at least two main objectives: 1) to quantitatively measure the quality of journals or authors using statistical indicators such as citation rates and 2) to analyze the structure of knowledge and the development of specific research fields.[13].

This paper aims to perform a bibliometric analysis of Machine Learning-Based Analysis for the Characterization and Phenomenological Study of Two-Phase Fluids research, explore existing publications and provide a comprehensive overview.

1. Methodology

Several literature review methodologies are available, including structured review, model/framework review, meta-analysis, theoretical examination, hybrid futures research, framework, bibliometric review, and systematic reviews.[14].

Bibliometrics belongs to the discipline of scientometrics, which can be applied to evaluate the scientific activities of a field of research or of a specific journal. [15].

This study applies bibliometric methodology to analyze the literature on Machine Learning Based Analysis for the Characterization and Phenomenological Study of Two-Phase Fluids. Bibliometric analysis is a well-accepted scientific method used worldwide. It uses mathematical and statistical

techniques for its evaluation. Based on citation mapping, the bibliometric technique can quantitatively synthesize a research topic and provide information on the main research streams.

A scientific mapping methodology was used in this research, which examines the relationships between research components. The analysis refers to the intellectual interactions and structural connections between research components. Techniques for scientific mapping include citation analysis, co-citation analysis, bibliographic linkage, co-word analysis, and co-authorship analysis. Such techniques, when combined with network analysis, are fundamental to present the bibliometric structure and intellectual structure of the research field. Table 1 presents a summary of the different techniques for scientific mapping with emphasis on their use and data considerations.

Table 1. Techniques for scientific mapping

Technique	Use	Analysis Unit	Data requirements
Citation analysis	Analyze the relationships between publications by identifying the most influential publications in a field of research.	Documents	Author's name Citation Title Journals DOI References
Analysis of joint appointments	Analyze the relationships between cited publications to understand the development of foundational topics in a field of research.	Documents	References
Co-Word Analysis	Explore existing or future relationships between topics in a field of research by focusing on the written content of the publication itself.	Keywords	Title Abstract Keywords of the author Keywords of the index Full text
Co-authorship analysis	Examine the interactions or social relationships between authors and their affiliations and equivalent impacts on the development of the research field.	Authors' affiliations	Author's affiliation (institution and country)

Results and discusión

The SLR search protocol was developed considering the Scopus databases where the following keywords were entered:

- Two-phase Flow -Machine learning
- Two-phase Flow -Tansformer neuronal networks

1.1 Keywords Two-phase Flow -Machine learning
1.1.1   Production indicators

The temporal search period was from the first article on Two-phase Flow -Machine learning, which was published in 2005 in the Scopus database, to the year 2023. In 2008 there were 61 articles published. The number of publications has increased significantly since 2018, with the highest number (74) in 2021 and 2022, as seen in Figure 1.
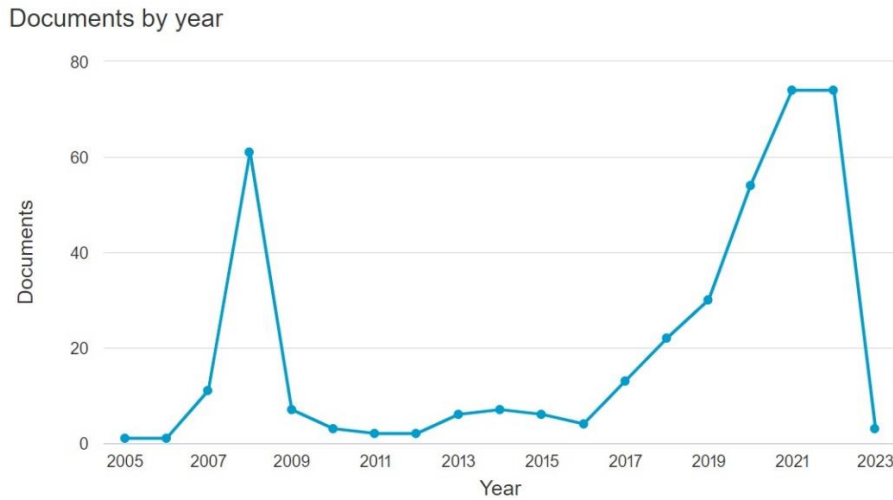

Figure 1. Evolution of scientific production registered in Scopus

3.1.2 Citation análisis

The analysis of citations is performed for this research by author and by countries of the scopus database, it can be observed which are the authors with the highest number of citations in this case in figure 2 it can be analyzed that the author Zhang y, has 8 documents with 151 citations.

Figure 2. Analysis of citations by Author, VOSviewer software.

Figure 3 shows that the countries with the highest number of citations are China and the United States, and that in South America the subject is little explored in Brazil, with 12 documents with 67 citations, and in Colombia, with 5 documents with 5 citations.
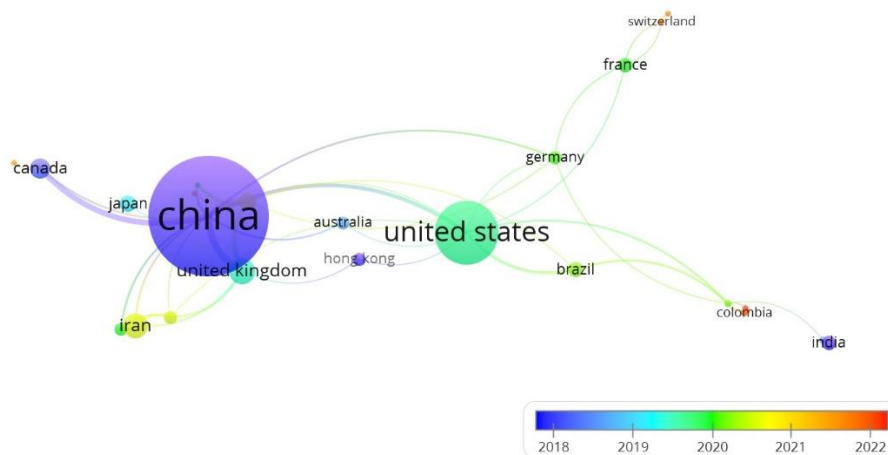
Figure 3. Citation analysis by country, VOSviewer software.

3.1.3 Co-Word Analysis

Co-word analysis refers to the study of the co-occurrences, or joint occurrences, of two terms in a given text, with the purpose of identifying the conceptual and thematic structure of a scientific domain.
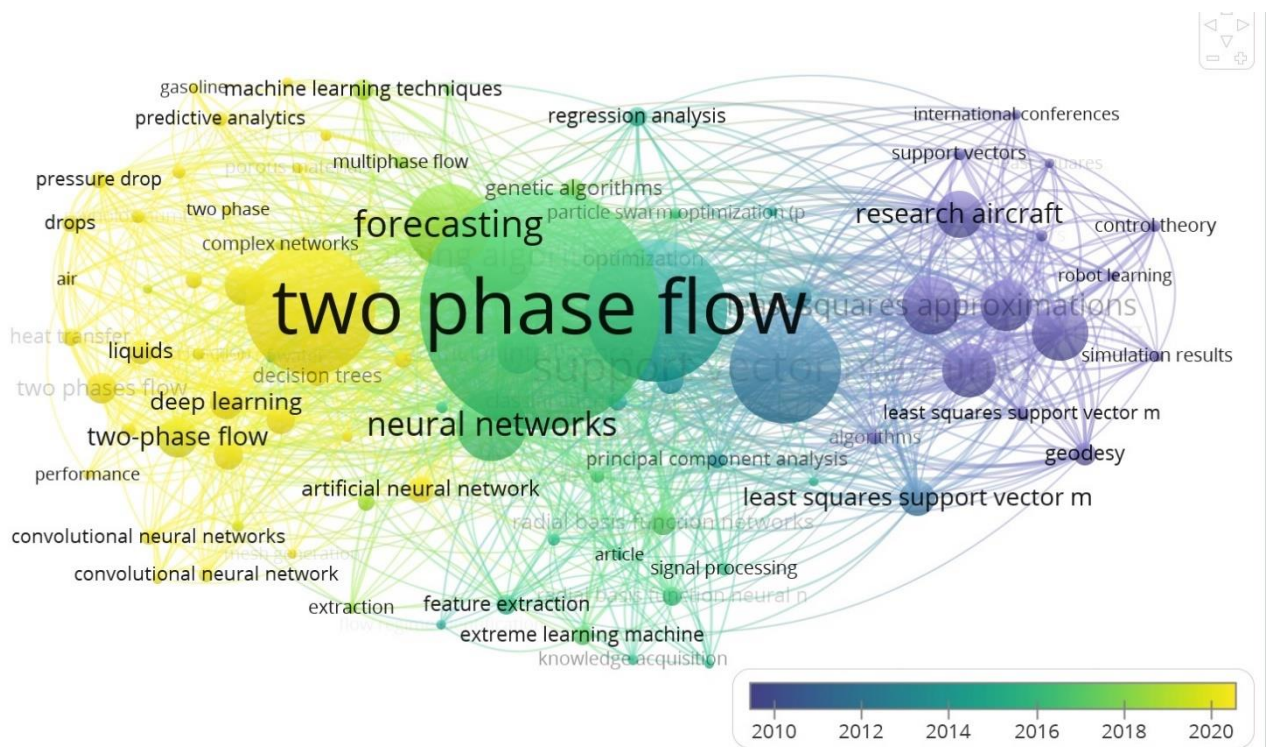

Figure 4. Keyword co-occurrence using VOSviewer network visualization.

Figure 4 shows the network of connecting terms, which were applied with at least 10 occurrences within the keywords. A total of 4228 keywords were identified, which allowed the discovery of the hotspot in the scientific papers analyzed in this research. However, only 89 words met the determined threshold (10 occurrences). The most frequent were:

Two pase Flow with 283 co-occurrences, learning systems with 163, support vector machines with 126, machine learning with 149, forecasting with 93, 6, these being the most important terms in the study.

The similarity measures served as input for the clustering analysis, forming 4 groups, one of each color (Figure 4). The words "Two pase Flow" were the most representative in terms of citations, assigned to group 1 (red color). The words "support vector machines" for group 2 (green color), "machine learning" for group 3 (blue color), and "learning algorithms" for group 4 (yellow color). Each node in the network represents a keyword, and the size of the node indicates the occurrence of the keyword. The link between the nodes represents the cooccurrence between the codewords. The thickness of the linking line indicates the occurrence or cooccurrences between the keywords.

3.1.4 Analysis of joint appointments

Bibliographic coupling The thematic groups created through the analysis of quotations and bibliographic coupling shed light on the main themes underpinning the intellectual structure and its development over time in the field of research.

Figure 5 shows the thematic similarity by bibliographic linkage.

Figure 5. Joint appointment analysis using VOSviewer network visualization.

3.1.5 Coauthorship Analysis

Scientific collaboration among researchers (co-authorship) is one of the most important aspects in the evolution of science, which presupposes an increase in the quality of the studies carried out. Hence, the importance of studying a) the visualization of co-authorship citations, b) networks of co-authored organizations/institutions, and, c) networks of co-authored countries (Figures 6, 7 and 8 respectively). The number and thickness of the lines indicate the links between elements, the distance between elements shows their relationship and the different colors indicate groups.
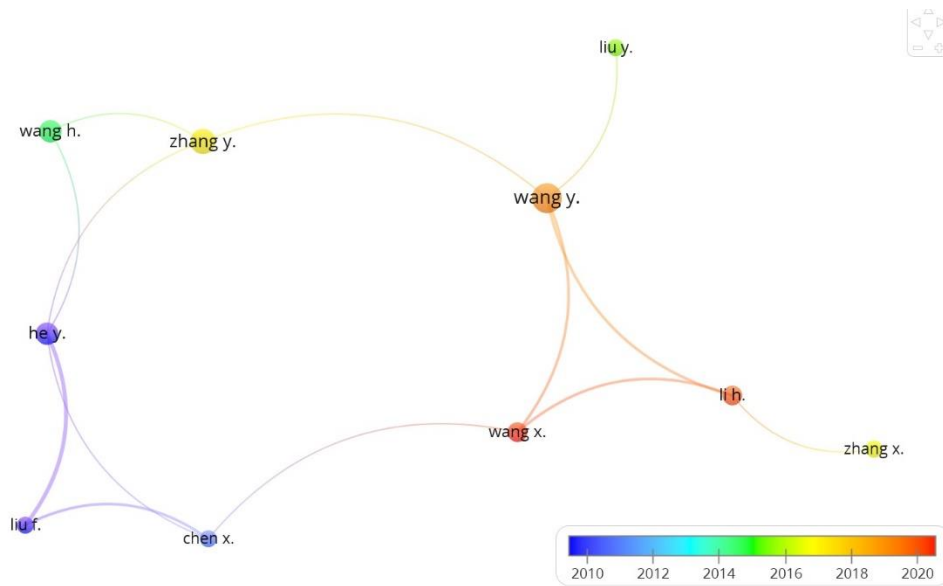
Figure 6. Viewing co-authorship citations using the VOSviewer network viewer.

Thus, Figure 6 shows a visualization of 13 most relevant co-authors during the period from 2010 to 2020, as well as the occurrences and links between the different authors. The most relevant co-author in the network was Liu y with 5 papers and with a total of 176 citations.
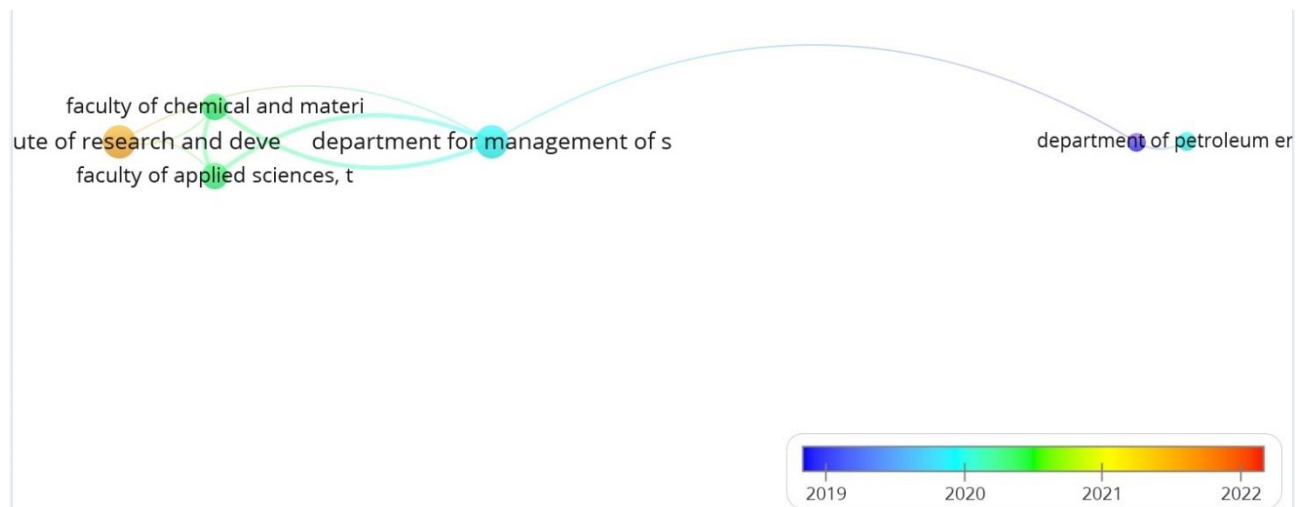

Figure 7. Co-authored organization/institution networks using VOSviewer network visualization.

A network of 747 organizations is identified in co-authorships, only 7 organizations have more than 4 papers, the one that stands out is the department for magagement of scien that has 4 papers with 113 citations. As can be seen in Figure 7.
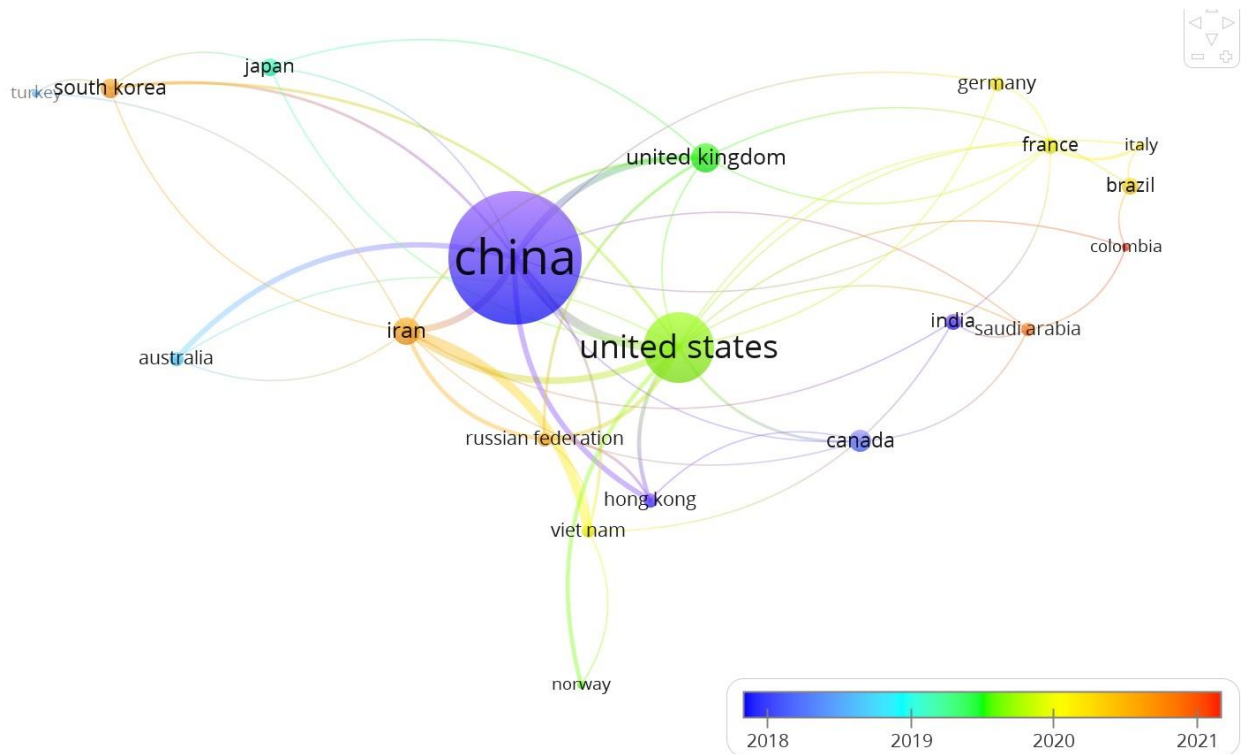
Figure 8. Co-authored country networks using VOSviewer network visualization.

In this same sense, Figure 8 shows the countries in co-authorship referred to this research. Figure 8 shows that China and the USA are the countries with the most co-authored documents; among the South American countries, Brazil has 12 co-authored documents and Colombia has 5 documents and 5 co-author citations, showing that the subject is not very developed in the South American countries.

3.1.6 Co- Citations

La co-citación de documentos se centra en como documentos primarios citan juntos a pares documentos secundarios indicando similitud semántica.
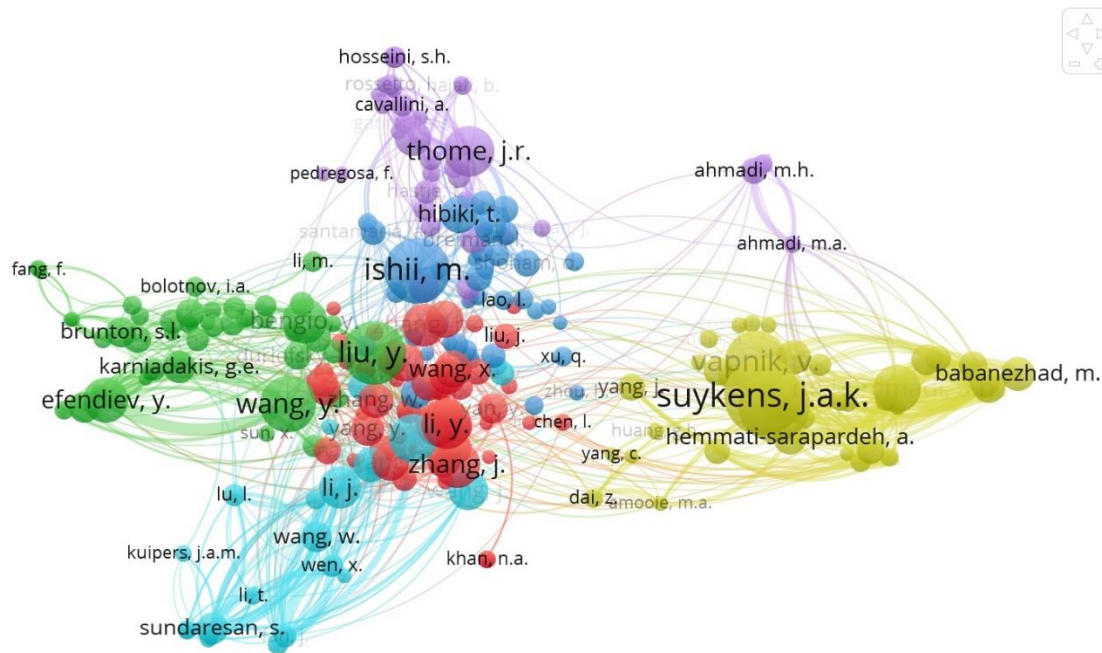
3.1.7

Figure 9. Conscitation analysis using VOSviewer network visualization.

For this research, Figure 9 shows the degree of co-citation strength for secondary documents that are co-cited (i.e., referenced in the same primary document), showing that they share content similarities.

It should be recalled that, document co-citation is a dynamic measure that changes over time, as older documents accumulate more citations

3.2   Two-phase Flow -Tansformer neuronal networks

When a more specific search is made on the subject with the Tansformer neuronal networks technique in Figure 10, it can be seen that the scopus database does not yield any results. This new technique of neural networks has not yet been exploited in engineering.
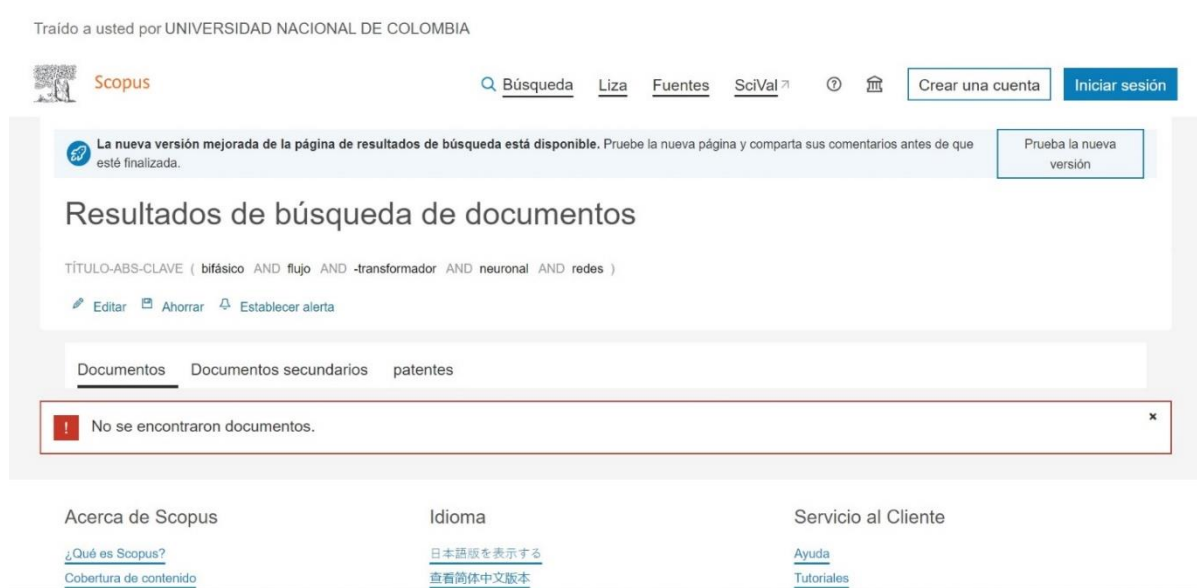

Figure 10. Scopus Databases.

Conclusions

Using a comprehensive bibliometric analysis, the current study has highlighted the research trends, languages, notable researchers and countries of origin of research on the topic of Machine Learning-Based Analysis for the Characterization and Phenomenological Study of Two-Phase Fluids.

China and the United States were found to be the countries with the most publications on the topic.

The bibliometric methodologies allowed discerning the relative influence of 383 scientific publications indexed in Scopus, and how these documents are grouped in networks. The conclusion is the need to create more research networks between countries, institutions, authors and co-authors.

Bibliography

[1]   J. K. Tamala, E. I. Maramag, K. A. Simeon, and J. J. Ignacio, "A bibliometric analysis of sustainable oil and gas production research using VOSviewer," Clean. Eng. Technol., vol. 7, p. 100437, 2022, doi: 10.1016/j.clet.2022.100437.

[2]   R. S. de F. Dam, W. L. Salgado, R. Schirru, and C. M. Salgado, "Application of radioactive particle tracking and an artificial neural network to calculating the flow rate in a two-phase (oil–water) stratified flow regime," Appl. Radiat. Isot., vol. 180, no. August 2021, 2022, doi: 10.1016/j.apradiso.2021.110061.

[3]   C. Suosheng, F. Lide, and J. Qingyong, "Slip property of oil-water two phase flow in horizontal pipe," 2010 2nd Int. Conf. Comput. Autom. Eng. ICCAE 2010, vol. 4, pp. 200–203, 2010, doi: 10.1109/ICCAE.2010.5451724.

[4]   M. Du, H. Yin, X. Chen, and X. Wang, "Oil-in-water two-phase flow pattern identification from experimental snapshots using convolutional neural network," IEEE Access, vol. 7, pp. 6219–6225, 2018, doi: 10.1109/ACCESS.2018.2888733.

[5]   M. S. G. Nandagopal, E. Abraham, and N. Selvaraju, "Advanced neural network prediction and system identification of liquid-liquid flow patterns in circular microchannels with varying angle of confluence," Chem. Eng. J., vol. 309, pp. 850–865, 2017, doi: 10.1016/j.cej.2016.10.106.

[6]   J. dos Santos Ambrosio, A. E. Lazzaretti, D. R. Pipa, and M. J. da Silva, "Two-phase flow pattern classification based on void fraction time series and machine learning," Flow Meas. Instrum., vol. 83, no. December 2020, p. 102084, 2022, doi: 10.1016/j.flowmeasinst.2021.102084.

[7]   L. Zhao, T. Dai, Z. Qiao, P. Sun, J. Hao, and Y. Yang, "Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse," Process Saf. Environ. Prot., vol. 133, no. 92, pp. 169–182, 2020, doi: 10.1016/j.psep.2019.11.014.

[8]   M. Al-Naser, M. Elshafei, and A. Al-Sarkhi, "Artificial neural network application for multiphase flow patterns detection: A new approach," J. Pet. Sci. Eng., vol. 145, pp. 548–564, 2016, doi: 10.1016/j.petrol.2016.06.029.

[9]   C. M. Salgado, R. S. F. Dam, E. J. A. Puertas, and W. L. Salgado, "Calculation of volume

fractions regardless scale deposition in the oil industry pipelines using feed-forward multilayer perceptron artificial neural network and MCNP6 code," Appl. Radiat. Isot., vol. 185, no. March 2021, 2022, doi: 10.1016/j.apradiso.2022.110215.

[10]   J. E. Juliá, Y. Liu, S. Paranjape, and M. Ishii, "Upward vertical two-phase flow local flow regime identification using neural network techniques," Nucl. Eng. Des., vol. 238, no. 1, pp. 156–169, 2008, doi: 10.1016/j.nucengdes.2007.05.005.

[11]   M. Roshani et al., "Evaluation of flow pattern recognition and void fraction measurement in two phase flow independent of oil pipeline's scale layer thickness," Alexandria Eng. J., vol. 60, no. 1, pp. 1955–1966, 2021, doi: 10.1016/j.aej.2020.11.043.

[12]   N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," J. Bus. Res., vol. 133, no. May, pp. 285–296, 2021, doi: 10.1016/j.jbusres.2021.04.070.

[13]   J. Zhu and W. Liu, "A tale of two databases: the use of Web of Science and Scopus in academic papers," Scientometrics, vol. 123, no. 1, pp. 321–335, 2020, doi: 10.1007/s11192-020-03387-8.

[14]   A. R. Anugerah, P. S. Muttaqin, and W. Trinarningsih, "Social network analysis in business and management research: A bibliometric analysis of the research trend and performance from 2001 to 2020," Heliyon, vol. 8, no. 4, p. e09270, 2022, doi: 10.1016/j.heliyon.2022.e09270.

[15]   D. Yu, Z. Xu, Y. Kao, and C. T. Lin, "The structure and citation landscape of IEEE transactions on fuzzy systems (1994-2015)," IEEE Trans. Fuzzy Syst., vol. 26, no. 2, pp. 430–442, 2018, doi: 10.1109/TFUZZ.2017.2672732.