Civil Engineering Beyond Limits 1 (2024) 1908
Civil Engineering Beyond Limits
journal home: www.acapublishing.com/journals/1/cebel

Research Article

A Bibliometric Review of Earthquake and Machine Learning Research

Osman Nuri Akarsu*[1], Oğuzhan Akarsu[2], Abdulkadir Cüneyt Aydın[2]
[1]Kafkas University, Institute of Social Sciences, Kars, Turkey
[2]Atatürk University, Department of Civil Engineering, Erzurum, Turkey

Keywords

Earthquake research,
Bibliometric analysis,
Machine learning,
Research trends.

Abstract

This article presents a bibliometric review of earthquake research and its integration with machine learning techniques. Over the past two decades, there has been a growing interest in using machine learning to enhance earthquake prediction and research. The review collected 1172 scholarly articles from the Web of Science database, focusing on the keywords "earthquake" and "machine learning." Machine learning has shown promise in improving earthquake forecasting models and aiding decision-making in disaster management, infrastructure design, and emergency response. However, it is noted that the application of machine learning in earthquake engineering is still in its early stages and requires further exploration. Key findings of this review include the increasing importance of certain keywords in earthquake and machine learning research, such as "prediction," "neural network," "classification," "logistic regression," and "performance." These keywords highlight the central areas of research focus within this field. The review also identifies research trends and gaps, including the need for more exploration of large-scale, high-dimensional, nonlinear, non-stationary, and heterogeneous spatiotemporal data in earthquake engineering. It emphasizes the necessity for novel machine learning algorithms tailored specifically for earthquake prediction and analysis. Furthermore, it highlights the need for addressing uncertainty in earthquake research and improving forecasting models. The review underscores the growth in interest and collaboration in earthquake research and machine learning, evident in the increasing number of scholarly contributions over the years. In summary, this bibliometric review highlights the importance of accurate forecasting and the potential of machine learning techniques in advancing this field.

1. Introduction

Earthquakes are natural disasters that can cause significant damage and loss of life. Predicting earthquakes accurately has been a challenging task for scientists and researchers due to the complex nature of seismic activities. This study undertakes a comprehensive bibliometric analysis to shed light on the landscape of earthquake and machine learning research, with the aim of providing valuable insights for the advancement of both fields. The striking fundamental question for this study: How has the use of machine learning in earthquake research evolved over the past two decades, and what trends and gaps can be identified from a bibliometric perspective? The symbiotic relationship between machine-learning and earthquake science holds immense potential for improving prediction accuracy and enabling more effective disaster management. However, a systematic assessment of scientific contributions in this intersection is needed to identify research trends, knowledge gaps, and avenues for future exploration. Earthquakes are unpredictable natural disasters that pose a significant threat to communities around the world. The complexity of seismic activity adds to the challenge of accurate forecasting. Machine learning methods offer a promising solution by tackling the complexity of earthquake data, but the extent of their implementation and impact is still under investigation.

In recent years, there has been an increasing interest in utilizing machine learning methods to predict earthquakes [1]. Various studies have been conducted to explore the potential of machine learning algorithms in earthquake prediction [2]. Moreover, the development and implementation of these machine-learning methods in the field of earthquake engineering have been the subject of extensive research [3]. According to a study conducted by Xie et al., machine learning methods offer a promising avenue for advancing earthquake research and practice, particularly in dealing with large-scale, high-dimensional, nonlinear, non-stationary, and heterogeneous spatiotemporal data. Given the potential for rapid data growth in the field of seismic activities, machine learning has emerged as a valuable tool for advancing earthquake engineering research and improving earthquake prediction models [4]. However, despite the growing interest and potential of machine learning in earthquake engineering, the implementation of these methods is still in its early stage and requires further investigation.

In the present study, an assessment is conducted regarding the research endeavors pertaining to earthquakes and machine learning. This evaluation is predicated upon a bibliometric analysis encompassing 1172 scholarly contributions disseminated across various academic forums, including journals, within the confines of the Scientific Citation Index (SCI) and the Social Science Citation Index (SSCI) bibliographic repositories. The time frame under consideration spans from 2003 to 2023.

Numerous bibliometric investigations have previously been conducted on the earthquake studies and machine learning [5], [6], [7], [8], [9], [10]. For instance, Liu offers a bibliometric examination of earthquake investigation throughout the years 1900-2010, to evaluate worldwide research advancement and furnish insights into research patterns in this field [11]. Furthermore, Xie et al. conducted a bibliometric analysis of recent research on earthquakes and machine learning to identify the main research topics, influential authors, and key publications in the field [12]. Machine learning is also widely used in structural engineering [13], [14], [15], [16], [17], [18]. For example,

*Corresponding Author: osmannuriakarsu@gmail.com
Received 08 Apr 2024; Revised 21 Apr 2024; Accepted 21 Apr 2024
https://doi.org/10.36937/cebel.2024.1908

Ruggieri et al. present a machine learning framework called VULMA (VULnerability analysis using MAchine-learning) for assessing the seismic vulnerability of existing buildings by processing photographs, and highlights that yield favourable results for earthquake risk assessment [19]. Recently, machine learning methods have become increasingly important in the field of disaster management[20].

The use of machine learning in earthquake research offers numerous advantages and opportunities for advancement. Firstly, machine learning (ML) techniques can effectively handle the complex and dynamic nature of earthquake data. Using ML methods allows for the analysis of large-scale and high-dimensional spatiotemporal data in earthquake engineering, which traditional approaches may struggle to handle. Secondly, machine learning models can provide computational efficiency, allowing for faster analysis and processing of earthquake data. Thirdly, ML methods have the ability to propagate and treat uncertainties in earthquake engineering [12].

This is particularly important in earthquake research, where uncertainties are inherent due to the unpredictable nature of seismic events. By using machine learning techniques, researchers can improve their earthquake forecasting models and make more accurate predictions. Additionally, the use of machine learning in earthquake research can facilitate decision-making. Researchers can use ML models to analyze and interpret complex data sets, leading to better-informed decisions regarding disaster management, infrastructure design, and emergency response.

This study enriches the discourse on earthquake and machine learning integration by offering a comprehensive overview of research trajectories. By identifying trends and unexplored domains, this analysis aims to guide future research and applications in earthquake engineering and disaster management.

2. Approach and Methodology

To review the significance of various machine learning techniques for earthquake engineering applications, this study will utilize a systematic approach. We collected publications on earthquakes and machine learning using the Web of Science bibliographic databases. We conducted a bibliographic search using the keywords "earthquake" and "machine learning" and identified publications that included these words in their titles, abstracts or keywords. We collected the author's name(s), author affiliation(s), subject category(ies), journal name(s), publication title(s), and publication year(s) for each publication. Using the search methodology outlined, a total of 1172 scholarly articles were identified within the Web of Science (WoS) databases, covering the period of 2003 to 2023. The bibliometric analysis of the study was conducted using the R-Studio software program.

2.1. Bibliometric and Scientific Mapping Methods

Bibliometrics serves to evaluate the scientific performance of authors, articles, journals, institutions, and countries in a given field of knowledge. Bibliometrics serves to evaluate the scientific performance of authors, articles, journals, institutions, and countries in a given field of knowledge. Bibliometrics serves to evaluate the scientific performance of authors, articles, journals, institutions, and countries in a given field of knowledge. It relies on the analysis of keywords and citations to map the state of the art. El Mohadab et al. have highlighted new aspects of scientific research related to this theme [21].

Bibliometric analysis has become a well-liked quantitative research technique to assess scientific productivity and detect trends in specific research fields [22]. This statistical method evaluates the quality and quantity of published scientific literature and scrutinizes trends in a particular area by conducting various citation analyses. Bibliometric analysis offers numerous advantages. The main benefit of bibliometric analysis lies in its capacity to perform quantitative analysis based on measurable, objective, readily-accessible data generated by coded information. Hence, this method is an effective and valuable tool to identify research trends across various disciplinary domains. The analysis is carried out by employing literature statistics on the concerned subject, pertinent keywords, or identifying the highest quality journals [23].

In recent years, there has been a gradual increase in the study of reference patterns in scientific literature, known as bibliometric analysis. This field refers to the application of statistical techniques to data from books, articles, and other publications. Bibliometrics is commonly used to explore the academic influence of specific authors, groups of authors, or publications on a given topic [24].

A significant array of software has been designed and put into practice for scientific mapping analysis. Certain software that has been developed for general purposes can also be utilized for scientific mapping. R, UCINET, VOSviewer, Gephi, Cytoscape, Science of Science Tool, CiteSpace II, CoPalRed, IN-SPIRE, Pajek, Bibexcel, and VantagePoint are among such software [25]. Fig. 1 shows a diagram of the bibliometric analysis process.

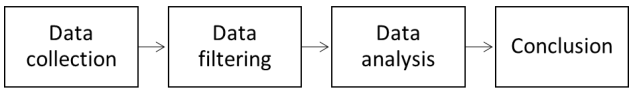

Figure 1. Bibliometric analysis procedure [26]

2.1.1. Data analysis

When conducting data analysis, the WoS analysis tool was utilized to preliminarily analyze the main authors, universities, journals, and other factors that characterize scientific production in financial accounting, including the impact evaluated by the h-index. To perform the evolutionary evaluations, the "R" program was utilized, which conducts content analysis of publications. In addition, this program is employed to monitor a scholarly domain, comprehend its intellectual, social, and cognitive structure by defining research areas, analyze its structural evolution, and generate maps of scientific output. The "R" software was also utilized to construct the network graph of publications and the density graph depicting the scientific publication performance of countries [27].

3. Results and discussions

3.1. Document types

The bibliometric analysis results demonstrate the wide range of academic productivity observed in the investigated field, which mirrors the variety of document types employed. In the context of this study, which encompasses a total of 1172 documents, the breakdown includes 943 articles, 29 article revisions, 5 book chapters and reviews, 146 conference papers, and 49 early access documents. Tab. 1 succinctly summarizes the showcased distribution of the diversity within document types in the analyzed field.

Table 1. Distribution of Document Types

Total Document Count for Analysis	1172
Articles	943
Article Revisions	29
Book Chapters and Reviews	5
Conference Papers	146
Early Access	49

The findings emphasize the significant academic production and the use of diverse communication channels.

3.2. Publication outputs

The research output descriptors demonstrate a significant progress in earthquake and machine-learning research, including improved scientific productivity and increased collaborative research efforts. The analysis of the data reveals discernible trends in publication counts over the years. From the initial years, spanning 2003 to 2008, it is evident that publication counts remained at modest, single-digit levels. However, commencing from the year 2009, a distinct upward

trajectory in publication counts becomes evident, with this trend intensifying notably during the 2010s. Post-2012, a more pronounced acceleration in publication counts is observable, with this momentum gaining further impetus from 2015 onward. Between 2015 and 2020, publication counts experience an almost threefold increment, signifying a noteworthy surge in scholarly output. Subsequently, a marked escalation becomes apparent from 2020 onwards. Particularly in the year 2022, a substantial leap in publication counts is observed, while the robust publication activity during the first seven months of 2023 also merits distinct attention. This analysis effectively elucidates the evolving academic interest and research inclinations within the studied domain over time. The heightened publication counts during the initial seven months of 2023 underscores the enduring relevance of the subject matter and portends a sustained vigor in research pursuits in the foreseeable future. The figure illustrates the number of publications over the years, as presented in Fig. 2. It is noteworthy that the data for the year 2023 encompasses only the initial seven months, thus potentially underrepresenting the annual aggregate. Considering this, it is plausible to anticipate that the cumulative reference count for the entirety of 2023 may well surpass the current recorded value, emblematic of the continuous expansion and diversification of the scholarly dialogue.

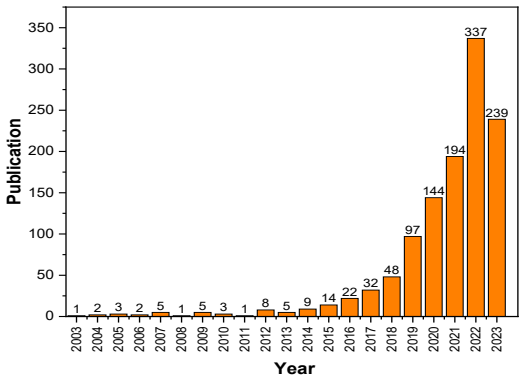

Figure 2. Publication Count Across Years.

The examination of the data pertaining to the reference citations, analyzed longitudinally, offers insights into the dynamic evolution of academic publications in terms of their referencing behavior. In Fig. 3, the reference citation analysis is presented.

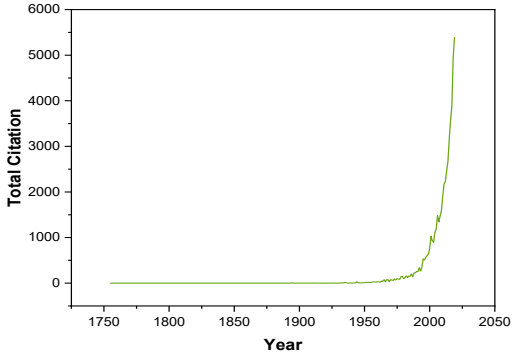

Figure 3. Reference Citation Analysis

Noteworthy observations are discerned throughout the trajectory. In the latter half of the 1950s, scholarly works exhibited relatively modest referencing practices, a trend that underwent transformative changes in the subsequent decades. The 1960s marked a pivotal point where the referencing behavior exhibited a discernible increase, presaging a more pronounced paradigm shift. However, it is the mid-1970s that constitute a notable juncture, characterized by a significant surge in the referencing intensity. This crescendo in referencing proclivity persistently prevailed, extending from the 1970s well into the first years of the new millennium. The commencement of the 21st century witnessed an accelerated escalation in this referencing trend, with the 2010s witnessing a remarkable proliferation in the number of references employed in scholarly discourse. The evident surge in the reference count unveils the deepening interconnections and multifaceted engagements within the scholarly landscape. The progressive augmentation in referencing patterns further reflects the burgeoning complexity of academic inquiries and the expanding horizons of research inquiries within the academic domain.

3.3. Subject categories and major journals

Earthquake and machine-learning research spanned over 92 subject categories in database. The five most common categories were Geosciences Multidisciplinary, Geochemistry, Geophysics, Engineering Civil, Engineering Geological, Remote Sensing. The supplied data set presents the quantities of articles across different academic fields, offering an interesting opportunity for examination and academic interpretation.

The illustrated data set in Fig. 4 depicts the quantities of articles within various academic disciplines. This dataset can offer insight into which subjects have received more scholarly attention and which fields have experienced heightened academic exploration. On examination of the top 10 prevalent subjects, it is clear that "Geosciences Multidisciplinary" is the foremost subject with 386 articles. This indicates the broad research within the realms of geography and environmental sciences. The "Geochemistry Geophysics" category contains 224 articles, demonstrating a significant interest in geochemistry and geophysics. The prominence of engineering disciplines is also notable, with "Engineering Civil" (197 articles) and "Engineering Geological" (168 articles) highlighting a substantial emphasis on infrastructure and geological studies. Furthermore, "Engineering Electrical Electronic" (115 articles) reflects devoted research in the fields of electrical and electronic engineering. Computer science also has a noteworthy presence. "Computer Science Artificial Intelligence" (94 articles) and "Computer Science Information Systems" (82 articles) emphasize the significance of artificial intelligence and information system research. The field of Imaging Science and Photographic Technology (112 articles) demonstrates the importance of visual media and image processing. Meanwhile, "Environmental Sciences" (137 articles) and "Remote Sensing" (139 articles) demonstrate a significant interest in environmental sciences and remote sensing technologies.

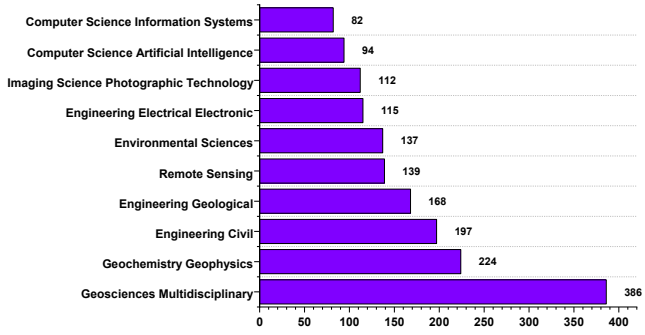

Figure 4. Top 10 academic disciplines

This analysis offers valuable insights into how academic research focuses on specific subjects and indicates which disciplines are given priority. For researchers, this dataset can serve as a guide to plan future studies and investigate their areas of interest more deeply. It provides guidance for directing research efforts in line with current academic trends.

In Table 2, we provide a list of the 13 most prominent journals publishing research on earthquakes and machine learning. The table includes the number of articles, citations, and impact factors for each journal. Its columns correspond to the journal name, total number of citations (TC), number of published papers (NP), the ratio of total citations to total papers (TC/TP), impact factor (IF) for 2022, and the start year of publication (PY_start).

Table 2. The 13 most prominent journals in earthquake and machine-learning research.

Journal	TC	NP	TC/TP	IF (2022)	PY_start
Remote Sensing	1328	57	23,30	5	2014
Geophysical Research Letters	1064	33	32,24	5,2	2017
Journal of Geophysical Research-Solid Earth	631	42	15,02	3,9	2018
Soil Dynamics and Earthquake Engineering	628	37	16,97	4	2014
Seismological Research Letters	569	38	14,97	3,3	2019
Catena	528	7	75,43	6,2	2018
Earthquake Spectra	381	16	23,81	5	2014
Engineering Structures	347	20	17,35	5,5	2019
Geophysical Journal International	322	35	9,20	2,8	2016
Construction And Building Materials	268	10	26,80	7,4	2019
Earthquake Engineering & Structural Dynamics	267	24	11,13	4,5	2014
IEEE Transactions on Geoscience and Remote Sensing	254	19	13,37	8,2	2018
Natural Hazards	254	19	13,37	3,7	2015

By analyzing this information, one can gain insights into the influence and impact of these journals within the academic community. Journals that have higher citation counts, impact factors, and ratios of citations to papers are likely to be more influential and widely referenced in the field. For example, "Remote Sensing" is notable for its high total citation count of 1328, a substantial number of published papers (57), and a high TC/TP ratio of 23.30. Moreover, it has an impact factor of 5 for the year 2022. "Geophysical Research Letters" also has a notable total citation count of 1064 and an impressive impact factor of 5.2 in 2022. The "Journal of Geophysical Research-Solid Earth" demonstrates substantial influence in the field with a considerable total citation count of 631. It is noteworthy that the "Catena" journal has published only 7 papers but has an exceptionally high TC/TP ratio of 75.43. This may indicate that the limited publications in this journal have received an unusually high level of citations. Additionally, certain journals, such as "Construction and Building Materials" and "IEEE Transactions on Geoscience and Remote Sensing," have distinguished impact factors (7.4 and 8.2, correspondingly) for 2018, demonstrating their significant impact and standing within their corresponding fields. Overall, this data emphasises the significance of citation metrics and impact factors in evaluating the importance and influence of academic journals. It also highlights the variety of journals in terms of citation counts, impact factors, and publication histories, demonstrating the fluid nature of research publication within the earthquake and machine-learning fields.

3.4. Author productivity

The dataset contains data on 4171 authors in total, with 31 single-authored documents. On average, each document has 5.08 authors. Additionally, there is a significant annual growth rate of 31.5%. Lotka's Law is an important concept in academia used to understand author productivity and publication distribution. This law shows that a minority of authors contribute the majority of scholarly output while the majority of authors contribute comparatively fewer works.

In our dataset, we have categorized the number of documents written by authors and the corresponding number of authors in each category. This categorisation enables an investigation into author productivity distribution. There are 3211 authors who have written just one document, accounting for approximately 77% of the total authors. This aligns with Lotka's Law, which indicates a larger number of authors who contribute fewer documents. Around 15.9% of authors, or 665 individuals, have written two documents. This remains in compliance with Lotka's Law as it reveals that a lesser proportion of authors create a slightly greater number of documents compared to those who produced only one. As we transition to authors who composed 3 to 10 documents, the quantity of authors in each group considerably decreases, as expected. For instance, authors who have produced 3 documents constitute roughly 3% of the overall authors. The pattern persists with fewer authors creating a greater number of documents. For example, only 0.8% of all authors have written five documents. The distribution found here follows Lotka's Law, which suggests that a minority of authors are responsible for the majority of documents. This common phenomenon is present across various fields and highlights how author productivity tends to be uneven. This underscores the importance of abundant authors' contributions to the academic literature. The data analysis demonstrates this pattern effectively, reaffirming the applicability of Lotka's Law to the distribution of author productivity in a specific context. The graphical representation relating to the research on the Lotka Law can be observed in Fig. 5.

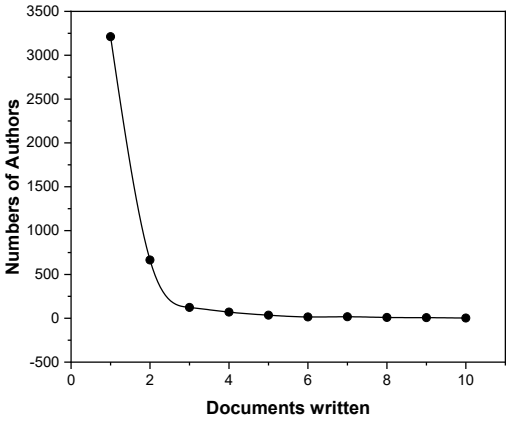

Figure 5. Author productivity through Lotka's Law

Table-3 shows researchers who have made significant contributions to the fields of earthquake and machine learning. Specifically, Li Y has the highest number of articles with 26, denoting their extensive involvement. Chen Y and Wang Y follow closely with 19 and 18 articles respectively, indicating their noteworthy research efforts. Other authors, including Li Z, Zhang H, Chen J, Huang Y, Zhang M, Li H and Wang W, have also made valuable contributions, as evidenced by their respective article counts. The metric of fractionalized articles provides valuable insight into the distribution of contributions. Fundamentally, the data underscores a set of authors who have made a substantial impact on their relevant fields via their extensive scholarly output, indicative of their commitment to the advancement of knowledge.

Table 3. The 10 most productive authors

Authors	Articles	Articles Fractionalized
Li Y	26	4.84
Chen Y	19	5.85
Wang Y	18	4.38
Li Z	15	4.88
Zhang H	14	2.61
Chen J	13	2.97
Huang Y	13	3.43

Zhang M	13	2.53
Li H	12	2.43
Wang W	12	3.18

In Table 4, the authors Devries Pmr, Viegas F, and Wattenberg Bj have each received 27 local citations, indicating the considerable impact of their work. Equally, Huang Y has amassed 25 citations, and Zhao L has been cited 24 times, further demonstrating their scholarly influence. Authors Harirchian E and Lahmer T, as well as Derendyaev Ab and Gitis Vg, have received 14 and 11 local citations respectively, indicating their noteworthy contributions. Additionally, Rasulzade S has received eight citations, indicating their presence in the academic discourse.

Table 4. The 10 most cited authors

Author	Citations
Devries Pmr	27
Viegas F	27
Wattenberg Bj	27
Huang Y	25
Zhao L	24
Harirchian E	14
Lahmer T	14
Derendyaev Ab	11
Gitis Vg	11
Rasulzade S	8

In summary, the data emphasises authors whose research has gained significant attention, demonstrating their crucial role in advancing scholarly conversations within their respective fields.

We examine the collaboration pattern for the 50 most productive authors, and the collaboration map is shown in Fig. 6.

The size of the circles is proportional to the number of publications, while the distance between circles shows the extent of collaboration among authors; shorter distances reveal higher levels of collaboration.

Finally, Table-5 presents the most frequently cited articles alongside their DOIs and the respective numbers of citations they have received.

The examination of the top ten most cited papers provides useful perspectives into earthquake and machine learning research. These papers, spanning different years and journals, showcase the considerable influence of particular studies within academic discourse. It is noteworthy that the abundant citations these papers have received demonstrate the importance of their conclusions and approaches, highlighting their contribution to the progression of both areas. The broad scope of topics encompassed, spanning seismic activity prediction to computational methodologies, exemplifies the multifaceted nature of research initiatives within this field. By garnering increasing interest and shaping academic dialogues, these papers add to the communal knowledge pool and facilitate future explorations at the converging junction of earthquake investigations and machine learning.

3.5. Temporal Dynamics of Keyword Occurrences

We conducted a keyword analysis to gain insight into current trends and the forefront of earthquake and machine-learning research. While conducting the word analysis, we categorised the reviews into three time periods: 2003-2010, 2011-2018 and 2019-2023. We have provided Table-6 containing the keywords that were most frequently used from 2003 to 2023.

In Fig. 7, a tree map displays the percentage breakdown of the most commonly used words. Within this examination, the term "earthquake" exhibited the most pronounced prevalence, accounting for 9% of the occurrences. Subsequently, the terms "model," "prediction," and "classification" followed with 7%, 6%, and 6% of the total occurrences, respectively.

The research involved the application of Multiple Correspondence Analysis (MCA) to the keywords included in our dataset.

The abstract configuration of the keywords associated with the articles studied is visually illustrated in Fig. 8. This analytical technique condenses the extensive dataset, replete with multiple variables, into a reduced-dimensional space, thereby creating an intuitive two-dimensional (or possibly three-dimensional) visualization. The planar distances within this graph substantiate the similarities between keywords. In particular, the keywords that converge towards the central locus denote the increased attention they have received in the contemporary period.

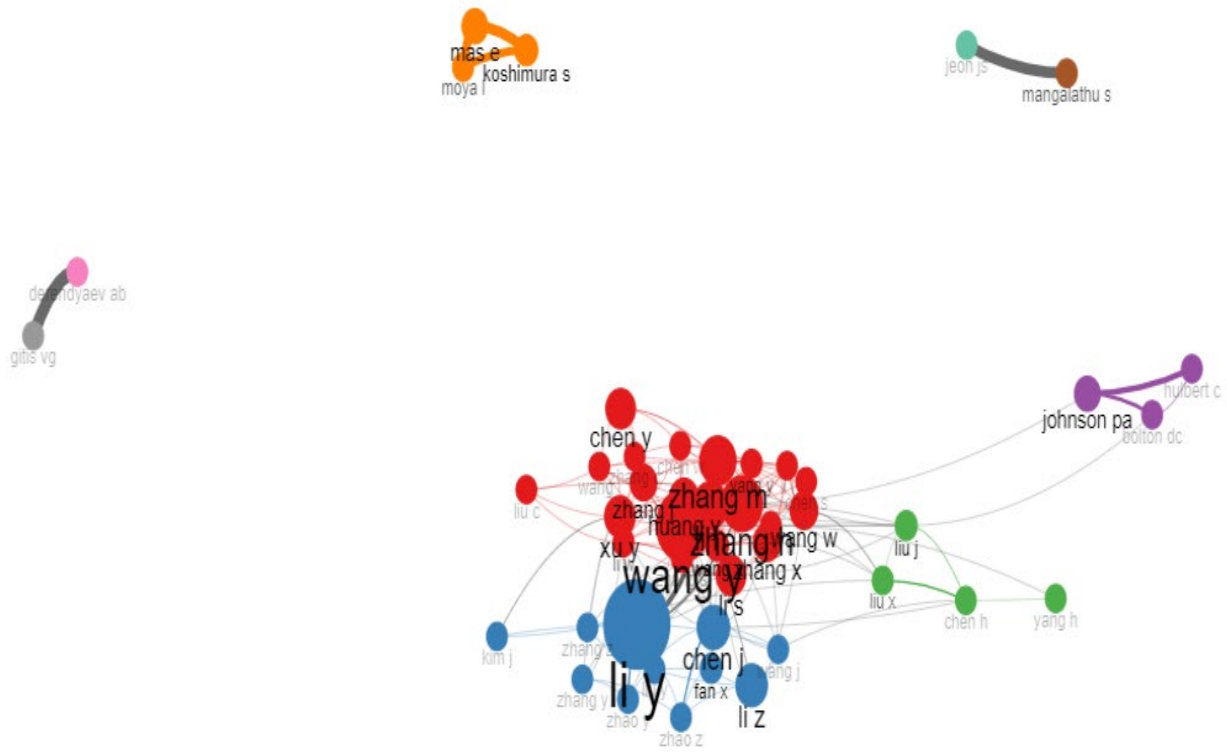

Figure 6. Integrated Analysis and Clustering of the Top 50 Prolific Authors in Earthquake and Machine Learning Research (2003-2023)

Table 5. Top 10 Highly Cited Papers in Earthquake and Machine-Learning Research

Paper	DOI	Total Citations	TC per Year
Ghorbanzadeh O, 2019, Remote Sensing	10.3390/rs11020196	363	72,60
Huang Y, 2018, Catena	10.1016/j.catena.2018.03.003	281	46,83
Ross Ze, 2018, J. Geophys. Res.-Solid Earth	10.1029/2017JB015251	266	44,33
Imran M, 2014, Www'14 Companion: Proceedings 23rd International Conf. World Wide Web	10.1145/2567948.2577034	241	24,10
Kong Q, 2019, Seismological Res. Lett.	10.1785/0220180259	226	45,20
Rouet-Leduc B, 2017, Geophys. Res. Lett.	10.1002/2017GL074677	191	27,29
Wang Q, 2020, IEEE Trans. Emerging Topics Comput.	10.1109/TETC.2017.2699169	160	40,00
Devries Pmr, 2018, Nature	10.1038/s41586-018-0438-y	158	26,33
Li Z, 2018, Geophys. Res. Lett.	10.1029/2018GL077870	151	25,17
Xie Y, 2020, Earthq. Spectra	10.1177/8755293020919419	145	36,25

Table 6. The Evolution of Prominent Keywords Across Time.

2003-2010		2011-2018		2019-2023	
Words	Occurrences	Words	Occurrences	Words	Occurrences
prediction	4	earthquake	20	earthquake	192
neural-network	2	model	19	model	191
assessment	1	classification	18	prediction	98
atmospheric-pressure fluctuations	1	neural-network	15	classification	88
classification	1	algorithm	7	neural-network	88
criterion	1	selection	7	logistic-regression	61
demeter satellite	1	prediction	6	performance	53
design	1	discrimination	5	algorithm	52
deterministic	1	magnitude	5	network	42
driven	1	area	4	artificial neural-network	35
earthquake	1	failure	4	behavior	34
emanation	1	pattern-recognition	4	hazard	33
geochemistry	1	picking	4	identification	32
entry	1	random forest	4	area	31
exposure	1	recognition	4	damage	31
extraction	1	satellite	4	frequency ratio	30
field	1	building damage	3	gis	30
geochemistry	1	areas	3	fault	28
linear-systems	1	artificial neural-network	3	system	27
liquefaction	1	behavior	3	magnitude	26

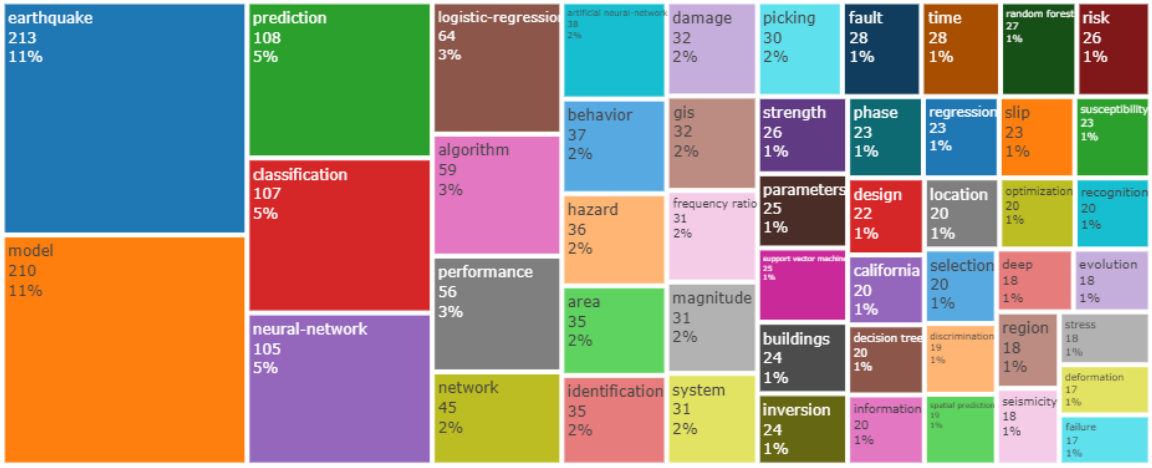

Figure 7. Tree map of Keywords.

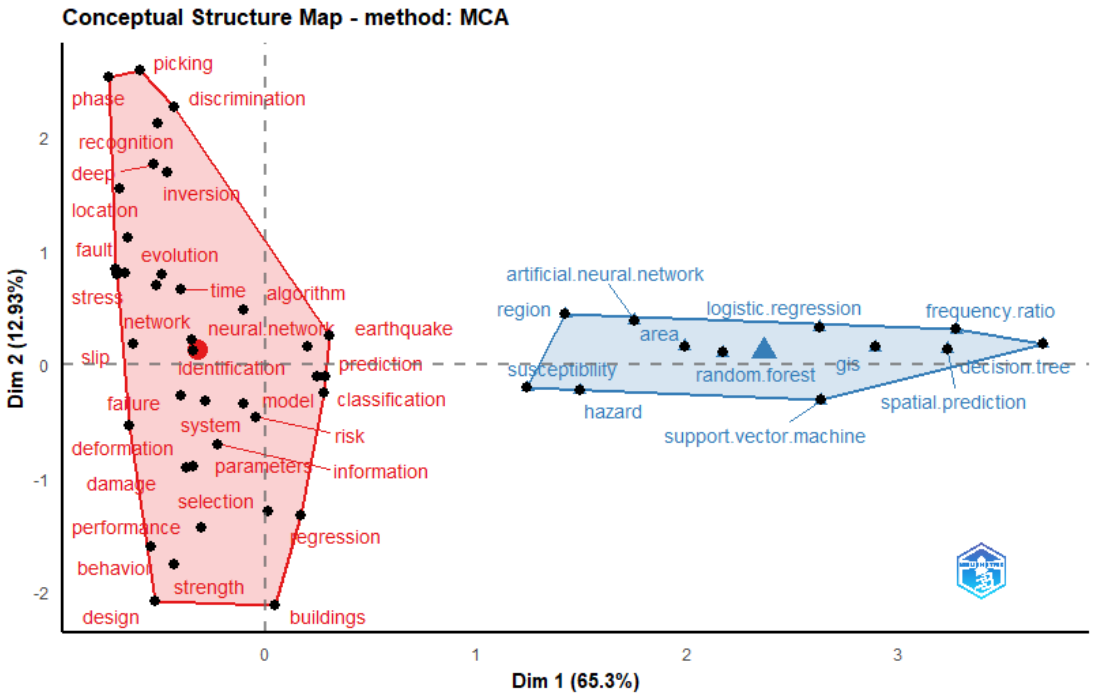

Figure 8. Factorial analysis of conceptual structure map-method: MCA of high-frequency keywords

A dendrogram is a graphical representation commonly used in statistics and data analysis. It is useful in hierarchical clustering analyses and classification problems. Consisting of a tree-like structure, the dendrogram visually illustrates how data points are grouped together based on measures of similarity or dissimilarity. The dendrogram shown in Fig. 9 illustrates the hierarchical arrangement and interrelationships between keywords resulting from hierarchical clustering. The segmentation shown in the figure, together with the vertical boundaries, helps to explore and understand the different clusters. It is important to note that the purpose of Fig. 9 is not to determine the final level of association between clusters, but rather to approximate the number of clusters present, thus fostering an environment conducive to subsequent discourse.

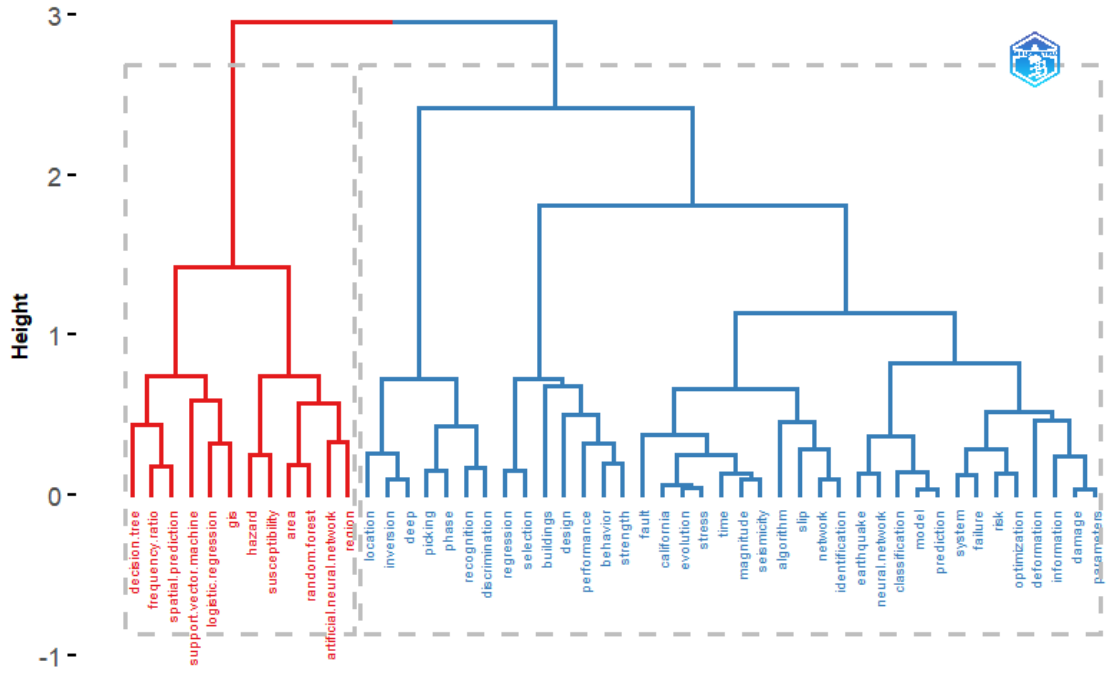

Figure 9. Topic dendrogram.

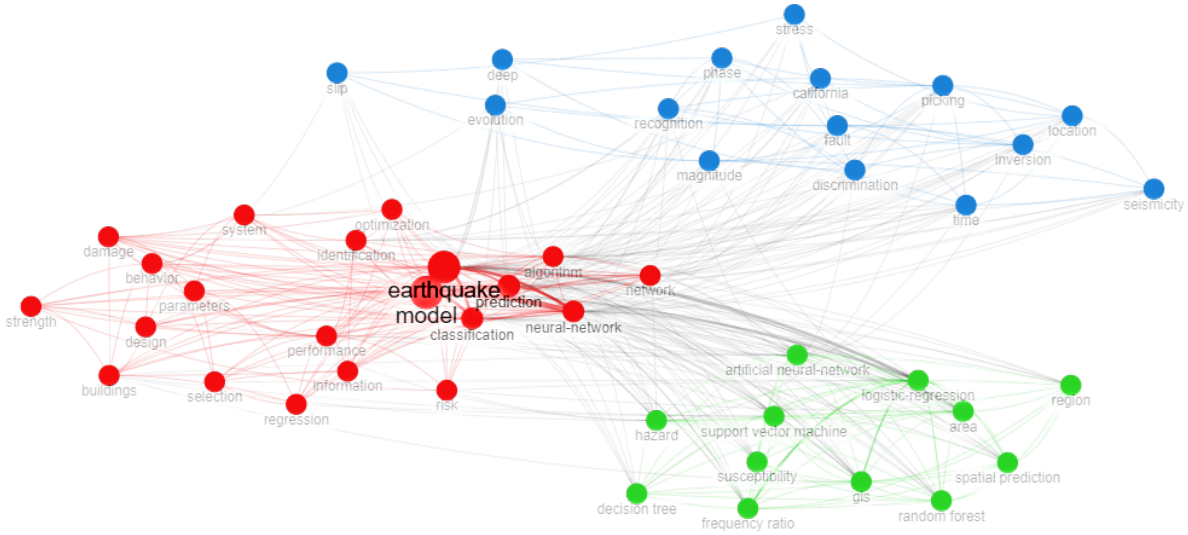

Figure 10. The network maps of keywords.

Lastly, Fig. 10 presents the network map of the keywords, illustrating their interconnected relationships.

In the first years of the analysis, the term "Earthquake" showed a significant presence with a constant frequency from 2003 to 2010, while the terms "Model", "Prediction" and "Classification" started to increase slowly. The frequency of the terms "Logistic Regression", "Neural Network" and "Performance" remained low during this period. In the following years, the term "Earthquake" showed a significant increase until 2023, while the frequency of other keywords also increased. Especially since the 2010s, the terms "Model", "Prediction", "Classification" and "Neural Network" show a remarkable increase. The terms "Logistic Regression" and "Performance" have also increased, but with a lower frequency compared to the other terms. Towards the beginning of the 2020s, the frequency of all keywords increased significantly. During this period, the terms "Earthquake" and "Model" had particularly high frequencies. The terms "Prediction", "Classification", "Neural Network" and "Performance" also increased rapidly. In conclusion, the results of the analyses show that the frequency of certain keywords tends to increase over the years. This trend reflects the significant growth and interest in studies related to earthquakes and machine learning.

4. Conclusions

In conclusion, the comprehensive bibliometric analysis conducted in the field of earthquake and machine learning research has shed light on several key aspects that define the landscape of scholarly endeavor in the field. The diversity and range of document types used, as shown in the dataset of 1172 documents, underlines the multifaceted nature of academic production. This is also reflected in the robust publication output over the years, indicating significant growth, collaborative efforts and evolving scholarly interest in the field.

The distribution of subject categories and major journals has provided valuable insights into the focus of research and the prominent avenues of scholarly discourse. This analysis not only provides a snapshot of the topics that have received considerable attention, but also points researchers towards areas of focus for future study. The spotlight on prolific authors has revealed the uneven distribution of author productivity, confirming Lotka's law and highlighting the importance of a select group of contributors who have had a significant impact on the field.

The temporal dynamics of keyword occurrences have illuminated the evolving trends and themes in earthquake and machine learning research. The prevalence of certain keywords over different time periods underscores the evolving nature of the research focus and technological advances in the field.

By examining the most cited papers, we have identified the influential works that have shaped the academic discourse in earthquake and machine learning research. These papers represent not only the broad range of topics within the field, but also the lasting impact that certain studies have had on the academic community.

In essence, the results of this study have highlighted the vibrant and evolving nature of earthquake and machine learning research. The collaborative efforts, evolving research interests, and impactful contributions of authors and papers collectively underscore the dynamic and multidisciplinary nature of the field. The insights gained from this bibliometric analysis serve as a compass for researchers, guiding them to areas of importance and providing valuable perspectives for shaping future research trajectories.

Declaration of Conflict of Interests

The authors declare that there is no conflict of interest. They have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1.] Gitis V G, Derendyaev A B, Petrov K N 2021 Earthquake prediction based on combined seismic and GPS monitoring data. In *International Conference on Computational Science and Its Applications* (pp. 601-612). Cham: Springer International Publishing.

[2.] Ridzwan N S M and Yusoff S H M 2023 Machine learning for earthquake prediction: a review (2017–2021). *Earth Science Informatics*, *16*(2), 1133-1149.

[3.] Thai H T 2022 Machine learning for structural engineering: A state-of-the-art review. In *Structures* (Vol. 38, pp. 448-491). Elsevier.

[4.] Fauvel K, Balouek-Thomert D, et.al 2020. A distributed multi-sensor machine learning approach to earthquake early warning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 01, pp. 403-411).

[5.] Huang Y, Xu C, Zhang X, Li L 2022 Bibliometric analysis of landslide research based on the WOS database. *Natural Hazards Research*, *2*(2), 49-61.

[6.] Liu L L, Zhang J, Li J Z, Huang F, Wang L C 2022 A bibliometric analysis of the landslide susceptibility research (1999–2021). *Geocarto International*, *37*(26), 14309-14334.

[7.] Huang Y, Xu C, Zhang X, Li L, Xu X 2023 Research in the Field of Natural Hazards Based on Bibliometric Analysis. *Natural Hazards Review*, *24*(2), 04023012.

[8.] He X 2022 Exploring the knowledge domain of earthquake prediction based on bibliometric analysis and text mining. *Journal of Intelligent & Fuzzy Systems*, *42*(6), 5885-5901.

[9.] Garcia J, et.al 2022 Machine learning techniques applied to construction: A hybrid bibliometric analysis of advances and future directions. *Automation in Construction*, *142*, 104532.

[10.] Li Y, Xu Z, Wang X, Wang X 2020. A bibliometric analysis on deep learning during 2007–2019. *International Journal of Machine Learning and Cybernetics*, *11*, 2807-2826.

[11.] Liu X, Zhan F B, Hong S, Niu B, Liu Y 2012 A bibliometric study of earthquake research: 1900–2010. *Scientometrics*, *92*(3), 747-765.

[12.] Xie Y, Ebad Sichani M, Padgett J E, DesRoches R 2020 The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra*, *36*(4), 1769-1801.

[13.] Kazemi F, Asgarkhani N, Jankowski R 2023 Predicting seismic response of SMRFs founded on different soil types using machine learning techniques. *Engineering Structures*, *274*, 114953.

[14.] Bayrak B, Akarsu O, Kaplan G, Aydın A C 2023 The plastic hinge length prediction of RC members by using ANN. *Sādhanā*, *48*(3), 114.

[15.] Leggieri V, Mastrodonato G, Uva G 2022 GIS multisource data for the seismic vulnerability assessment of buildings at the urban scale. *Buildings*, *12*(5), 523.

[16.] Pizarro P N and Massone L M 2021 Structural design of reinforced concrete buildings based on deep neural networks. *Engineering Structures*, *241*, 112377.

[17.] Xu J G, Hong W, Zhang J, Hou S T, Wu G 2022 Seismic performance assessment of corroded RC columns based on data-driven machine-learning approach. *Engineering Structures*, *255*, 113936.

[18.] Nguyen H D, Truong G T, Shi M 2021 Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. *Engineering Structures*, *235*, 112067.

[19.] Ruggieri S, Cardellicchio A, Leggieri V, Uva G 2021 Machine-learning based vulnerability analysis of existing buildings. *Automation in Construction*, *132*, 103936.

[20.] Linardos V, Drakaki M, Tzionas P, Karnavas Y L 2022. Machine learning in disaster management: recent developments in methods and applications. *Machine Learning and Knowledge Extraction*, *4*(2).

[21.] El Mohadab M, Bouikhalene B, Safi S 2020 Bibliometric method for mapping the state of the art of scientific production in Covid-19. *Chaos, Solitons & Fractals*, *139*, 110052.

[22.] Koo M, Lin S C 2023 An analysis of reporting practices in the top 100 cited health and medicine-related bibliometric studies from 2019 to 2021 based on some proposed guidelines. *Heliyon*.

[23.] Buber M and Koseoglu B 2022 The bibliometric analysis and visualization mapping of net environmental benefit analysis (NEBA). *Marine Pollution Bulletin*, *181*, 113931.

[24.] Nielsen S B, Lemire S, Bourgeois I, Fierro L A 2023 Mapping the evaluation capacity building landscape: A bibliometric analysis of scholarly communities and themes. *Evaluation and Program Planning*, 102318.

[25.] Demir H and Erigüç G 2018 Bibliyometrik bir analiz ile yönetim düşünce sisteminin incelenmesi. *İş ve İnsan Dergisi*, *5*(2), 91-114.

[26.] Mulyawati I B, Ramadhan D F 2021 Bibliometric and visualized analysis of scientific publications on geotechnics fields. *ASEAN Journal of Science and Engineering Education*, *1*(1), 37-46.

[27.] Derviş H 2019 Bibliometric analysis using bibliometrix an R package. *Journal of Scientometric Research*, *8*(3), 156-160.

How to Cite This Article

Akarsu, O.N., Akarsu, O., and Aydin, A.C., A Bibliometric Review of Earthquake and Machine Learning Research, Civil Engineering Beyond Limits, 1(2024), 1908.
https://doi.org/10.36937/cebel.2024. 1908