

Performance Analysis and Science Mapping on High Performance Computing in the Era of Artificial Intelligence

Haruna Chiroma(✉)

College of Computer Science and Engineering, University of Hafr Batin, Hafr Batin, Saudi Arabia
charuna@uhb.edu.sa

Abstract. The demand for high-performance computing (HPC) systems is increasing due to the growing need for large-scale computation and the rapid evolution of artificial intelligence. Previous bibliometric analyses on HPC are outdated, leaving the research community without an up-to-date investigation of the field. To bridge this gap and support the development of HPC while informing policymakers, this paper conducts a performance analysis and science mapping of HPC research. The study examines different aspects, including publication characteristics, research disciplines, trends, keyword co-occurrence, prolific authors, contributing countries, funding agencies, and document types. Findings indicate that researchers from the United States of America lead in HPC research. The Oak Ridge National Laboratory contributed the highest number of articles in the field of HPC. It is found that HPC research results are predominantly published in conferences**.** The top two leading funders of HPC research are from the United States of America. The United States of America play a pivotal role in linking different regions, facilitating cross-border research and innovation in HPC research. Mapping analysis shows that High-Performance Computing, Cloud Computing, Scientific Computing and Sparse Matrix-Vector Multiplication as frequent keywords highlights the core focus areas in HPC research. The article serves as a valuable resource for the research community, policymakers, funding agencies, industry professionals, and students seeking insights into the landscape of HPC research in this era of AI.

Keyword: Bibliometric Analysis · Co-Authorship · Collaborating Countries · High Performance Computing

1 Introduction

The world is currently experiencing an artificial intelligence (AI) boom, where AI technologies are penetrating almost every aspect of the society. Businesses and organizations are increasingly adopting AI to keep pace with current trends and remain competitive in the market.

AI technology requires high-cost computational power to achieve its target efficiency [1, 2]. Because of the AI highly ambitious nature, it relies on the support of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15887, pp. 27–39, 2026.
https://doi.org/10.1007/978-3-031-97589-9_3

high-performance computing (HPC) paradigms [3]. The HPC capabilities go beyond conventional computers, which typically have a single or a few processors (fewer than 10) [4, 5]. HPC systems consist of hundreds of thousands or even millions of cores (e.g., CPUs and accelerators such as GPUs) designed for complex modeling, scientific computations, and climate simulations. More recently, HPC has been increasingly deployed for pre-training and inference of advanced AI models [6], such as ChatGPT. The demand for the HPC infrastructure is increasing as a result of fast evolution in AI technology. This development has led to the intersection of AI with HPC [7–9].

Bibliometric analysis, also known as Scientometric analysis, is conducted to gain insights into the landscape and structure of a research field [10, 11]. It has been applied in different areas of computing, such as modeling the built environment on HPC [12], Apache Hadoop research trends [13], HPC scalability and deep learning architecture [14], AI in the Internet of Medical Things [15], and metaheuristic algorithms [16]. These studies help track developments, identify key contributors, and reveal emerging trends in computing research.

A bibliometric analysis was conducted on the research progress in HPC, incorporating both performance analysis and science mapping to provide insights into the field [17]. However, this bibliometric analysis is now outdated, leaving the research community without an up-to-date investigation of HPC for over a decade. Recently, a bibliometric analysis on the intersection of quantum computing and HPC was conducted to provide an overview of current developments in this area and the correlation between the HPC and the quantum computing [18]. However, that study specifically focuses on the intersection between quantum computing and HPC, whereas the current study primarily focuses on the bibliometric analysis of HPC as a whole to provide an updated landscape of bibliometric analysis of HPC research to close the over 10 years gap.

The paper compares the publication trends of the top five sources of HPC research documents over a period of more than 10 years to understand the evolution of publication patterns in the era of AI. Secondly, it examines the publication behavior and collaboration patterns among countries and authors, as well as word co-occurrence to identify the core focus areas in HPC research. Thirdly, it investigates funding agencies, subject areas, and types of publications to understand HPC community priorities. Achieving these objectives provides valuable insights into publication trends and the overall landscape of HPC research.

The article is structured as follows: Sect. 2 presents the methodology, followed by the results and discussion in Sect. 3. Finally, Sect. 4 provides the concluding remarks.

The contributions to knowledge in summary are provided as follows:

- The study provides an updated bibliometric analysis of HPC research, addressing a 10-year gap in the field making it a valuable resource for researchers and policymakers.
- The study analyzes publication trends, keyword co-occurrence, and subject areas to identify core focus areas in HPC research.
- The study reveal that High-Performance Computing, Cloud Computing, Scientific Computing, and Sparse Matrix-Vector Multiplication are dominant themes, providing insights into the evolving landscape of HPC research.
- The study ranks the most prolific countries, institutions, and authors in HPC research.

- The study identifies four major co-authorship clusters, emphasizing international collaboration trends.
- The study finds that HPC research prioritizes rapid dissemination, favoring conferences for quick knowledge sharing and networking.

2 Methodology

This section present the bibliometric analysis procedure comprising of different stages including aim and scope, selection of technique, search term and data; lastly, data analysis and visualization (Fig. 1). The bibliometric procedure is adopted from Donthu et al. [19]. The detailed activities in each of the stage is discussed as follows:

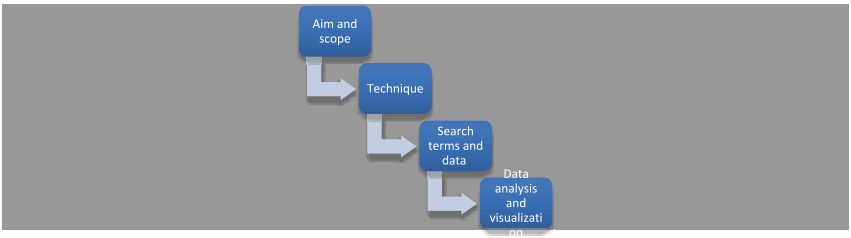

Fig. 1. Procedure for conducting bibliometric analysis

The aim of this study is to identify prolific authors, funders, institutions, countries, conferences, and journals in HPC research from the perspective of performance analysis**.** From a science mapping perspective, the study seeks to uncover the structural network of author-country collaborations, co-authorship relationships, and co-word occurrences, which contribute to the intellectual structure of HPC research. The scope of HPC research is extensive enough to justify both performance analysis and science mapping. The study aims to identify general themes, influential collaborating countries, and authorship collaborations within the period under study. To achieve this, co-word analysis, co-authorship analysis between countries, and co-authorship analysis between authors were adopted. Additionally, performance analysis was conducted by examining the total number of publications.

Keywords were formulated within the scope of the study defined at the first stage. Querying database requires keywords for effective retrieval of information about the HPC research publications. The curated search terms are as follows: *"SUPERCOMPUTERS" OR "HIGH AND PERFORMANCE AND COMPUTING" OR "PARALLEL AND COMPUTING" OR "CLUSTER AND SYSTEMS" OR "EXASCALE AND COMPUTING"*. The search terms were obtained from the literature. The academic database chosen for the data collection is Scopus database alone instead of multiple databases because of different data format and combining them can introduce error. Therefore, focusing on only Scopus eliminate the possibility of introducing error in the data. The data were collected from 2013 to 2024 limiting to English language. The results returned is 8,337 documents. The number of articles published within the study period exceeds

8,000, well above the 500-article benchmark considered sufficient for bibliometric analysis. The data was pre-processed by removing duplicates and unwanted field (e.g., DOI), reconciling discrepancies, ensure only valid affiliations were allowed and merged multiple entries. The data retrieved from the Scopus database was analysed using the analysis module of the Scopus. The analysed data based on the Scopus analysis module produce the performance analysis visualization for different components such as institutions, funders, subject area, authors, publication types, countries. On the other hand, the data was run to produce the intellectual structural networks.

3 Results and Discussion

This section presents the results obtained from the bibliometric analysis, including performance analysis and mapping. The analysis covers the sources of documents, subject areas, publication trends, prolific authors, prolific countries, prolific institutions, sponsors, and document types. On the other hand, the it includes co-authorship analysis between countries, co-authorship analysis between authors and word co-occurrence analysis between keywords.

3.1 Publications Trend

Figure 2 presents the publication trends in HPC from 2023 to 2024 for the top five sources with the highest document counts. From 2013 the publications trend start increasing for all the top five sources. While publication numbers have fluctuated over time, there was generally an increase followed by a decline, particularly starting in 2019 when all the top five sources experienced a sharp drop in output. The sharp drop in output from 2019 suggests that the field may be experiencing changes in research priorities, funding, or publication strategies.

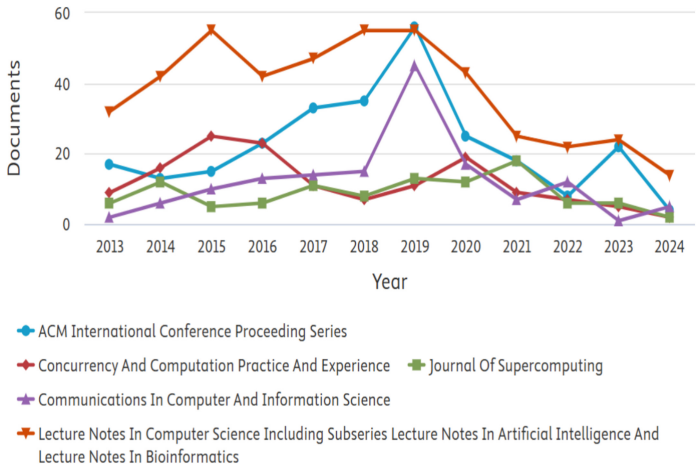

Fig. 2. Publications trend by source per year

This could reflect shifts in the HPC community or broader challenges in sustaining high publication volume. Lecture Notes in Computer Science (LNCS) and the ACM International Conference Proceedings (ACMICP) series occupied the first and second positions, respectively, in terms of publication output. The prominence of LNCS and ACMICP highlights their critical role as primary channels for disseminating HPC research.

3.2 Productivity by Country

Figure 3 depicts the ten most prolific countries publishing research on HPC. Researchers from the United States of America lead in HPC publications, contributing articles more than other countries. They are followed by China, which occupies a distant second position, emphasizing the significant role of the United States of America in HPC research.

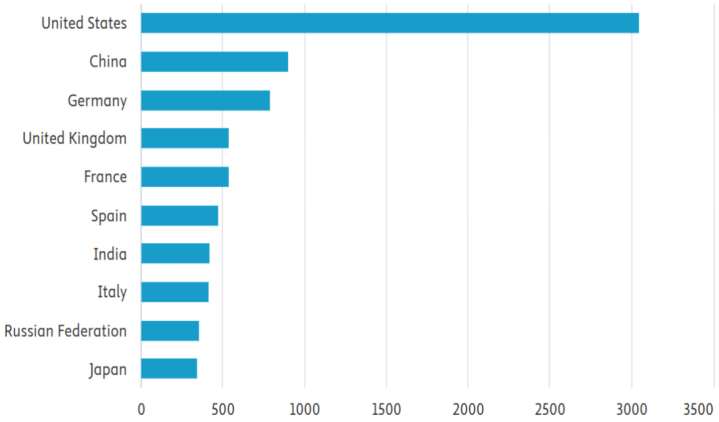

Fig. 3. High performance computing publication productivity by country

The United States of America leads HPC publications by a significant margin, indicating robust funding, research infrastructure, and strong international collaborations that position it as a global leader in HPC research. Although China ranks second, the large gap between its output and that of the United States of America suggests that while China is rapidly advancing, the United States of America maintains a substantial lead in HPC research activity. Figure 3 emphasizes that the United States of America plays a pivotal role in advancing HPC research worldwide, with China emerging as a key contributor, albeit with room to grow relative to the United States of America. A closer observation indicate that most of the countries on the chart are from Europe, emphasizing the strong presence of HPC researchers in that region.

3.3 Prolific Institutions

Figure 4 illustrate the 10 prolific institutions contributing research in the field of HPC. The illustration shows that the researchers from *Oak Ridge National Laboratory* contributed the highest number of articles followed by authors from *Argonne National Laboratory* and those from *Lawrence Berkeley National Laboratory* occupying the third position. Close observations indicate that most of the institutions are from United Sates of America.

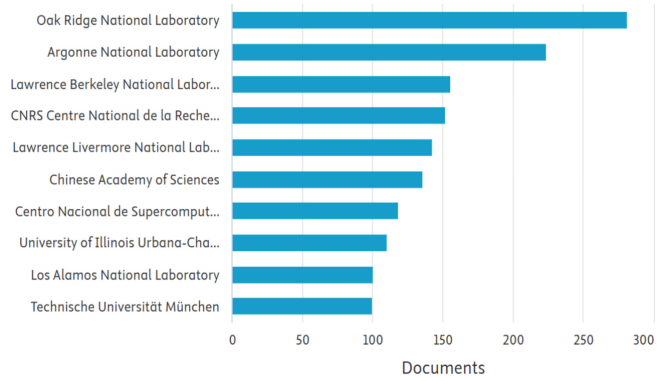

Fig. 4. Prolific institutions in high performance computing research

The trend illustrate the significant contributions of the authors from those institution in the field of HPC. This insight corroborate with the findings that indicate United States of America is the most prolific country in the area of HPC (Fig. 3). In the November 2024 TOP500 where the top-performing HPC system is El Capitan at *Lawrence Livermore National Laboratory*, United Sates of America. Other leading HPC systems include Frontier at *Oak Ridge National Laboratory*, United Sates of America; Aurora at *Argonne National Laboratory*, United Sates of America; and Jean Zay H100 at *CNRS Centre National de la Recherche Scientifique*, France.

3.4 Documents Type and Subject Area

Figure 5 presents HPC research output by document type. It highlights that HPC research results are predominantly published in conferences, followed by journals, with these two sources accounting for over 94% of all documents on HPC. This suggests that the HPC community prioritizes rapid communication and discussion of emerging findings, as conferences offer faster publication cycles and direct networking opportunities. It also implies that while journals remain important for in-depth research, other publication types play a minimal role in this fast-evolving field.

Figure 6 shows that HPC is a multidisciplinary research area that integrates different subject fields. However, a large portion of the research is dominated by Computer Science (41.3%) and Mathematics (16.7%), which occupy the first and second positions, respectively, followed by a third position in Engineering (16.3%). This distribution clearly confirms the multidisciplinary nature of HPC. HPC research spans different

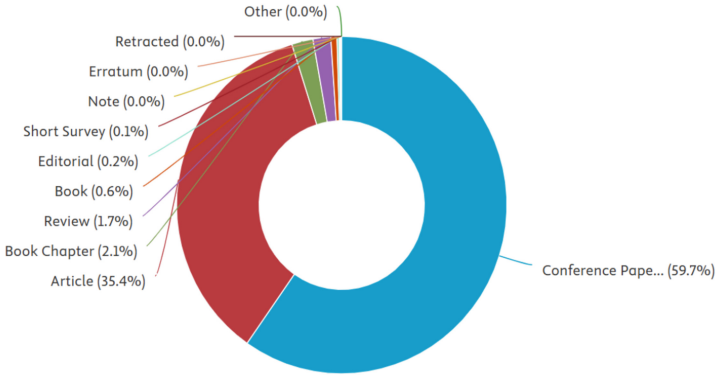

Fig. 5. High performance computing publications of documents by type

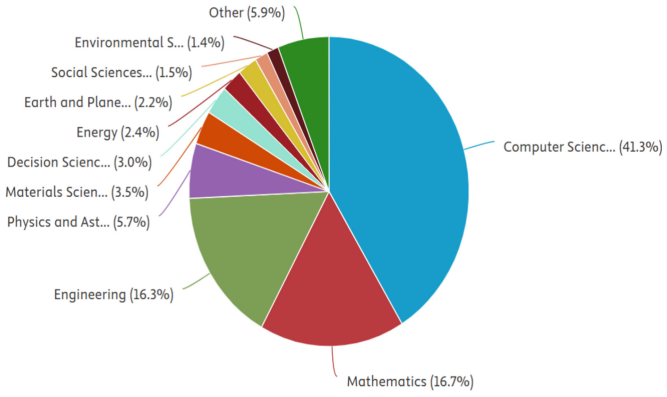

Fig. 6. High performance computing documents by subject area

fields, confirming that its development depends on the integration of multiple disciplines. With 41.3% of the research focused on computer science, it's clear that software development, algorithms, and computational methods are the primary drivers behind HPC systems. This emphasises the fact that regardless of hardware sophistication, effective HPC solutions fundamentally rely on robust computational frameworks. Mathematics, accounting for 16.7% of the research, highlights the essential role of quantitative analysis and theoretical foundations in advancing HPC capabilities. Engineering, making up 16.3%, plays a critical role in the design, implementation, and practical application of HPC systems, bridging the gap between theoretical research and real-world use.

3.5 Prolific Authors

Figure 7 highlights the top 10 prolific authors contributing the highest number of publications in the area of HPC. Chen, Y. is the researcher with the most publications, as indicated by the longest bar, followed by Dongarra, J. and Quintana-Orti, E.S. occupying the second and third positions respectively.

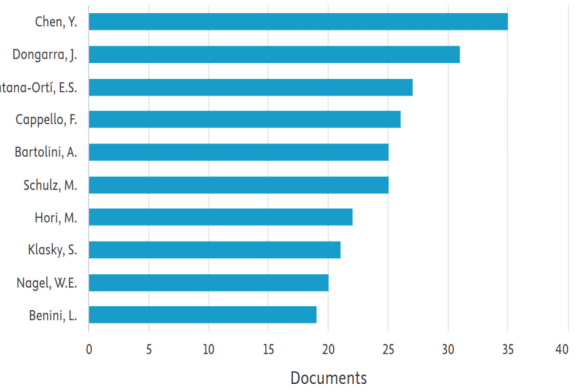

Fig. 7. High performance computing prolific authors by productivity

This suggests that these individuals are not only highly active in HPC research but may also be key influencers driving innovation and setting research trends in the field. Their substantial output indicates a deep engagement with HPC topics, which likely translates into significant impact on advancing HPC technologies and methodologies.

3.6 Funders

Figure 8 presents the top 10 funders of research in the field of HPC, ranked by the number of published documents. Notably, the National Science Foundation and the U.S. Department of Energy occupy the first and second positions, respectively. The leading funders are from the United States, Europe, and China, with the top two agencies located in the United States of America.

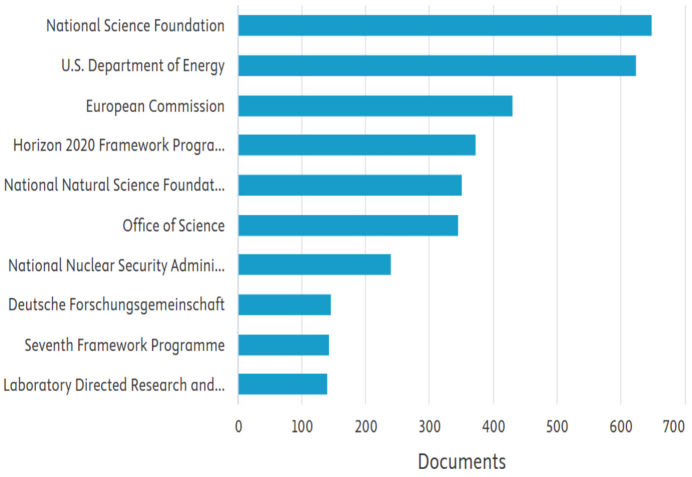

Fig. 8. Sponsors of high performance computing research

This indicates that the United States of America is a major driver of HPC research, benefiting from substantial governmental support that likely contributes to its leadership in the field. Additionally, the presence of top funders from Europe and China reflects a global commitment to HPC, although the U.S. funding agencies have a clear edge. This concentrated funding landscape highlights the strategic importance placed on HPC research, as robust financial support is essential for advancing HPC technologies and maintaining competitive research outputs on a global scale.

3.7 Country Co-authorship Analysis

Figure 9 presents the co-authorship network of HPC publications among authors from different countries. Each node represents a country, while the edges between nodes indicate co-authorship relationships. Thicker edges signify stronger collaboration between countries, and larger nodes represent a higher volume of authorship from that country. The network formed four major clusters: Cluster 1 is led by the United States of America; Cluster 2, by the United Kingdom; Cluster 3, by China; and Cluster 4, by Germany. There is strong collaboration between authors from the United States and China, between the United States and the United Kingdom, and between Germany and the United Kingdom, as indicated by the thick edges connecting these countries. Four major clusters have formed, led by the United States of America, UK, China, and Germany, indicating that while there are strong regional research bases, these countries serve as central hubs driving HPC research. The thick edges between the USA and China, as well as between the USA and the UK, suggest robust collaborative relationships.

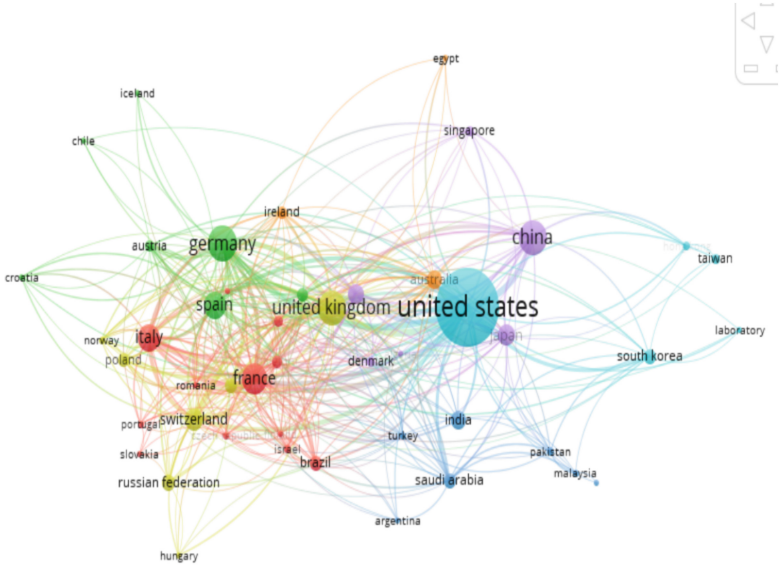

Fig. 9. Co-authorship analysis between countries

These connections highlight the pivotal role the USA plays in linking different regions, facilitating cross-border research and innovation in HPC research. The noticeable connection between Germany and the UK highlights active collaboration within Europe, contributing significantly to the HPC research landscape. Despite geopolitical differences, the strong collaborations especially between the USA and China indicate that scientific research in HPC transcends borders. This international cooperation is essential for pooling resources, sharing expertise, and advancing the field of HPC. In the November 2024 TOP500 list, out of the top 10 performing HPC systems globally, five are from the United States of America.

3.8 Author Co-authorship Analysis

Figure 10 presents the co-authorship collaborative network, illustrating the relationships among authors who publish research articles in the field of HPC. In this network, each node represents an author, and the connecting links indicate collaborations between authors. Thicker links denote stronger collaborative relationships, while larger nodes indicate a higher volume of publications by that author. The network forms five major clusters, each led by key authors: Cluster 1: Led by Dongarra, J.; Cluster 2: Led by Chen, Y., Cluster 3: Led by Wang, Y.; Cluster 4: Led by Yang, C.; Cluster 5: Led by Capello, F. and Liu, Y. These clusters highlight the emergence of central figures who play a pivotal role in fostering research collaboration in HPC.

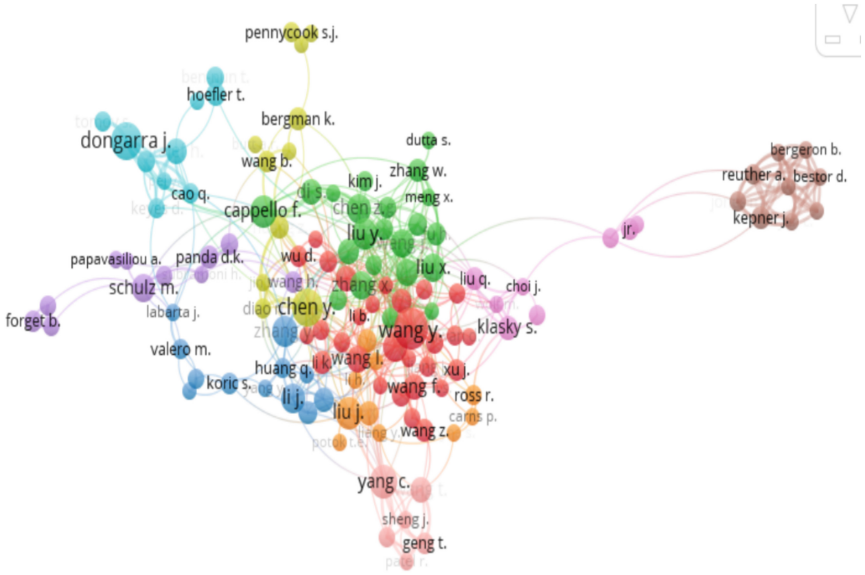

Fig. 10. Co-Authorship Analysis between authors

The thickness of the links indicates the intensity of collaboration among the authors. The network shows that these key authors not only publish frequently but also maintain

strong, influential partnerships with their peers. The larger nodes for these key authors suggest a high publication output, which coupled with their extensive collaborative ties, highlights their significant impact on advancing HPC research. Their centrality in the network indicates that they are likely to influence research directions and innovation within the field of HPC.

3.9 Keywords Co-occurrence

Figure 11 depicts the co-occurrence network of keywords frequently appearing in HPC research articles published by multiple authors. In this network: Node represent keywords, with larger nodes indicating a higher frequency of keyword appearances across multiple publications. Links between nodes represent the co-occurrence of keywords, with thicker links signifying that the connected keywords frequently appear together in research articles.

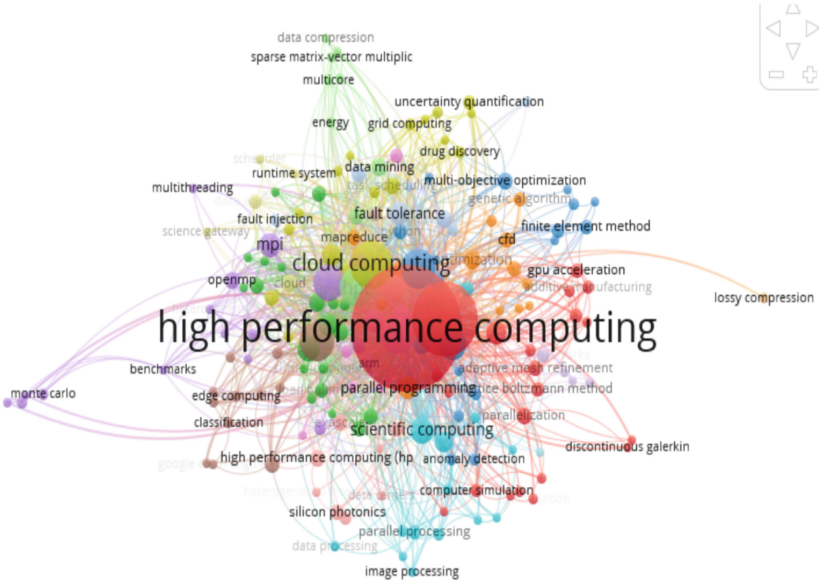

Fig. 11. Co-occurrence between authors keywords

The keywords identified from different clusters in the network include High-Performance Computing, Cloud Computing, Scientific Computing, and Sparse Matrix-Vector Multiplication. The presence of High-Performance Computing, Cloud Computing, and Scientific Computing as frequent keywords highlights the core focus areas in HPC research. The inclusion of Sparse Matrix-Vector Multiplication suggests a strong connection to numerical methods and computational efficiency, which are critical for optimizing HPC applications. The clustering of these keywords suggests interdisciplinary research, where HPC is frequently integrated with cloud computing and scientific computing methods. The larger nodes suggest that these keywords are dominant themes in HPC research, making them potential research hotspots.

4 Conclusions, Limitations and Future Research

The study present performance analysis and science mapping of HPC research for over 10 years period to uncover new pattern within the HPC research landscape. The study investigates funding agencies, subject areas, prolific countries, authors, institutions, publication behavior and collaborations. The study provides a comprehensive, data-driven assessment of the current state of HPC research in the era of AI. The area of HPC is dominated by USA. The study highlight the pivotal role of USA in linking different regions to facilitate cross-border research and innovation in HPC research. High-Performance Computing, Cloud Computing, Scientific Computing and Sparse Matrix-Vector Multiplication are found to be the core focus areas in HPC research. It is believed that this bibliometric study can offer valuable insights for researchers, funding agencies, and policymakers in HPC research decision-making process in this era of AI. These dimensions offer a holistic view of the field's dynamics, demonstrating how collaboration networks, funding mechanisms, and research agendas have evolved in response to the rapid evolving field of AI. However, the study is limited to data from the Scopus database, and the PRISMA guideline was not adopted, as the study primarily focuses on bibliometric analysis. Analyses such as h-index, m-index, g-index, i-index (e.g., i-10, i-100), single authorship analysis, proportion of cited publications, citations per cited publication, Bradford's law, academia–industry collaboration, and institutional collaborations were not included. It would be interesting for future studies to explore these comprehensive aspects in conjunction with in-depth state-of-the-art advances by adopting the PRISMA guideline using data from sources such as Web of Science, Scopus and the DBLP Computer Science Bibliography.

Disclosure of Interests. "The author declare that there is no any conflict of interest".

References

1. Swain, S.R., Parashar, A., Singh, A.K., Lee, C.N.: An intelligent virtual machine allocation optimization model for energy-efficient and reliable cloud environment. J. Supercomput. **81**(1), 1–26 (2025)
2. Rrapaj, E., Bhalachandra, S., Zhao, Z., Austin, B., Nam, H.A., Wright, N.J.: Power consumption trends in supercomputers: a study of NERSC's Cori and Perlmutter Machines. In: ISC High Performance 2024 Research Paper Proceedings (39th International Conference), pp. 1–10. Prometeus GmbH (2024)
3. Ettifouri, I., Zbakh, M., Tadonki, C.: The need for HPC in AI solutions. In: International Conference of Cloud Computing Technologies and Applications, pp. 137–159. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-78698-3_8
4. Shankar, S., Reuther, A.: Trends in energy estimates for computing in ai/machine learning accelerators, supercomputers, and compute-intensive applications. In: 2022 IEEE High Performance Extreme Computing Conference (HPEC), pp. 1–8. IEEE (2022)
5. Sterling, T., Brodowicz, M., Anderson, M.: High Performance Computing: Modern Systems and Practices. Morgan Kaufmann (2017)

6. Li, B., Basu Roy, R., Wang, D., Samsi, S., Gadepally, V., Tiwari, D.: Toward sustainable hpc: Carbon footprint estimation and environmental implications of hpc systems. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 1–15 (2023)
7. Sarma, R., Inanc, E., Aach, M., Lintermann, A.: Parallel and scalable AI in HPC systems for CFD applications and beyond. Front. High Perform. Comput. **2**, 1444337 (2024)
8. An, W., et al.: Fire-Flyer AI-HPC: a cost-effective software-hardware co-design for deep learning. In: SC24: International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 1–23. IEEE (2024)
9. Wulff, E., Girone, M., Pata, J.: Hyperparameter optimization of data-driven AI models on HPC systems. J. Phys. Conf. Ser. **2438**(1), 012092. IOP Publishing (2023)
10. Broadus, R.: Toward a definition of "bibliometrics." Scientometrics **12**(5–6), 373–379 (1987)
11. Pritchard, A.: Statistical bibliography or bibliometrics. J. Documen. **25**, 348 (1969)
12. Aghimien, E.I., Aghimien, L.M., Petinrin, O.O., Aghimien, D.O.: High-performance computing for computational modelling in built environment-related studies–a scientometric review. J. Eng. Des. Technol. **19**(5), 1138–1157 (2021)
13. Zhang, J., Lin, M.: A comprehensive bibliometric analysis of Apache Hadoop from 2008 to 2020. Int. J. Intell. Comput. Cybern. **16**(1), 99–120 (2023)
14. Gbedawo, W.V., Dzikunu, A., Nyamadi, M.: Scalability and efficiency of deep learning models on high-performance computing clusters: bibliometric analysis. African J. Appl. Res. **10**(2), 283–305 (2024)
15. Chiroma, H., Hashem, I.A.T., Maray, M.: Bibliometric analysis for artificial intelligence in the internet of medical things: mapping and performance analysis. Front. Artif. Intell. **7**, 1347815 (2024)
16. Ezugwu, A.E., et al.: Metaheuristics: a comprehensive overview and classification along with bibliometric analysis. Artif. Intell. Rev. **54**, 4237–4316 (2021)
17. Liu, Z., Liu, Y., Guo, Y., Wang, H.: Progress in global parallel computing research: a bibliometric approach. Scientometrics **95**, 967–983 (2013)
18. Garcia-Buendia, N., Muñoz-Montoro, A.J., Cortina, R., Maqueira-Marín, J.M., Moyano-Fuentes, J.: Mapping the landscape of quantum computing and high performance computing research over the last decade. IEEE Access (2024). https://doi.org/10.1109/ACCESS.2024.3411307
19. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M.: How to conduct a bibliometric analysis: an overview and guidelines. J. Bus. Res. **133**, 285–296 (2021)