

A Bibliometric Analysis and Research Landscape of Machine Learning Applications in Greenhouse Gas Emissions

Samuel-Soma M. Ajibade^{1,2*}, Anthonia Oluwatosin Adediran³, Muhammed Basheer Jasser^{1,4}, Angela Lee Siew Hoong¹, Faizah Mohammed Bashir⁵, Yakubu Aminu Dodo⁶, Olumide Simeon Ogunnusi⁷

¹School of Computing and Artificial Intelligence, Faculty of Engineering and Technology, Sunway University, Selangor, Malaysia, ²Research Centre for Nanomaterials and Energy Technology, Sunway University, Selangor, Malaysia, ³Department of Real Estate, Faculty of Built Environment, Universiti Malaya, Kuala Lumpur, Malaysia, ⁴Research Centre for Human-Machine Collaboration, Faculty of Engineering and Technology, Sunway University, Selangor, Malaysia, ⁵Department of Decoration and Interior Design Engineering, College of Engineering, University of Hail, Hail, Kingdom of Saudi Arabia, ⁶Department of Architectural Engineering, College of Engineering, Najran University, Najran, Kingdom of Saudi Arabia, ⁷Department of Computer Science, The Federal Polytechnic, Ekiti State, Nigeria. *Email: samuelma@sunway.edu.my

Received: 01 August 2025

Accepted: 25 October 2025

DOI: <https://doi.org/10.32479/ijeeep.22023>

ABSTRACT

This paper observes machine learning (ML) application in greenhouse gas (GHG) emission through a holistic bibliometric study, highlighting publication trends, influential contributors, and emerging research themes. Data were extracted from the Scopus database covering the period of 2008–2024. Bibliometric mapping was further performed to identify publication outputs, institutional contributions, collaboration patterns, and keyword co-occurrence networks. The analysis demonstrates a rapid growth in ML–GHG studies, with dominance of conference papers and journal articles. China and the United States are leading contributors especially in support of their institutions such as the U.S. Department of Energy and the Chinese Academy of Sciences. Research themes keep evolving from broad climate-related cases toward some advanced ML applications, including carbon capture optimization and real-time emissions monitoring. The study depends solely on Scopus data; it excluded relevant publications from other databases. Hence, expanding the data sources could offer a more inclusive perspective. The insights from the research can guide researchers, policymakers and industry practitioners on a deeper understanding of global trends as well as the identification of myriad opportunities for collaboration in applying ML for GHG/climate change mitigation. This research offers one of the first organized bibliometric overviews/summaries of ML applications in GHG mitigation, presenting a clearer picture of the field's evolution, significant actors, and research hotspots.

Keywords: Greenhouse Gas Emissions, Green Energy, Machine Learning, Climate Change, Mitigation, Bibliometric Analysis, Sustainable Energy
JEL Classifications: R2, R3, Q2, Q4

1. INTRODUCTION

Greenhouse gas (GHG) emissions are fundamental to worldwide climate change, as it keeps contributing to atmospheric warming, sea-level rise, while subsequent exacerbating extreme weather condition (Soeder, 2025). The Global Carbon Project demonstrate that nitrous oxide (N₂O), carbon dioxide (CO₂), and methane (CH₄), presently dominate the anthropogenic radiative forces responsible

for climate instability (Filonchik et al., 2024; Liaskoni, 2025). Hence, an exact monitoring, prediction, and alleviation of such emissions continues in pressing scientific necessities with machine learning (ML) now starting to surface as a challenging catalyst for progression in this field over the recent eras (Essamlali et al., 2024; Wei et al., 2021). More so, the models for assessing GHG like process-based simulations, physics-based models, or mechanistic inventory approaches often demand a considerable calibration,

specialized knowledge, and substantial data contributions (Bastviken et al., 2022; Liu et al., 2024). These limitations reduce its adaptability across varied situations, principally in data-shortage or emerging areas. Equally, ML methodologies like neural networks, support vector regression (SVR), gradient boosting, random forests, and deep learning frameworks provides adaptable, scalable, and data-centric options (Mamo et al., 2025; Singh, 2024). Recent study indicate that classical ML models can contest with or surpass the traditional methods in forecasting CO₂ emissions from agricultural fields, soils, and industrial processes, thereby exhibiting an enhanced adaptability towards nonlinear dynamics and the reduced computational demands (Farooq and Khan, 2025; Harsányi et al., 2024; Hassan et al., 2025; Li and Zhang, 2023).

In transportation, ML has progressed CO₂ emissions' prediction connecting to light duty cars (Udoh, 2024). It also enhances precision via ensemble methods like CatBoost, which proficiently manages extensive datasets and categorical variables with greater efficiency compare to conventional regressors (Ahmed et al., 2024; Nyakuma et al., 2023). Gas turbine emissions, mostly CO and NO_x, have remained precisely modeled through a varied array of deep learning and ML frameworks, like multi-layer perceptrons (MLPs), LSTM/GRU networks, and XGBoost, while supporting vector machines (Emad-Eldeen et al., 2024; Haque et al., 2024). This kind of advancement directly enhances real-time emissions control as well as regulating compliance systems (Predictive Emission Monitoring Systems). It further highlights the role of ML in offering operational and environmental advantages.

More so, ML provides essential feature and importance tools like SHAP or LIME which focuses on measurement efforts and pinpoint vital emissions drivers, while shortening model complexity on the utmost influential factors (e.g., facility capacity, sector type). This capacity, when united with remote sensing regularly facilitates an identification of hotspots and anomalies automatically, while enhancing the entire coverage of global emission sources and reducing the dependence on ground sensors (Hu et al., 2024; Tahir Bahadur et al., 2025). In spite of progress, ML-based procedures are still encountering noteworthy obstacles, including repeatability, constraints in data quality, model transparency, and environmental sustainability (Luccioni and Hernandez-Garcia, 2023) in (Hasan et al., 2025) examined the carbon footprint connected to training of large ML models and documented a considerable diversity across computer vision applications and natural language processing while underscoring the need for monitoring emissions from the ML process itself. Meanwhile, energy consumption for AI training is exponentially increasing, with projections of it representing tens of millions of tons of CO₂ emissions by year 2035 except if it is addressed via architectural competence and transitions to renewable energy sources (Devarakota et al., 2025; Saleh and Hassan, 2024).

In this context, an examination of research trends, emerging themes, and publishing patterns is essential for understanding the rising landscape of ML applications in GHG emission sectors. Bibliometric analysis approaches, provide comprehensive literature mapping, encompassing collaboration networks, temporal growth trends, disciplinary interaction and keyword clusters (Hassan and

Duarte, 2024). A bibliometric study examining ML in air pollution research from 1990 to 2021 reviewed approximately 2962 papers and identified four primary themes: short-term forecasting, chemical characterization of pollutants, augmentation of detection, and optimization of emission control (Dong et al., 2025; Kek et al., 2024). It was detected that majority of the research contributions stemmed from the United States and China, with a significant percentage of the work done within a single institution research groups rather than through a global collaborative teams (Hamdan and Alsquaih, 2024; Tan et al., 2023).

This study particularly examines “GHG Emission and ML,” analyzing papers from 2008 to 2024. This period covers the rise in ML applications driven by computational capabilities, improved data accessibility, and a universal imperative in addressing climate change objectives. The study's objective is to delineate the evolution of academic focus, prominent institutions, pinpoint prevailing research matters, and regional contributions, as well as to illustrate existing co-authorship and institutional networks. Hence, this bibliometric review emphasizes the overarching methodological, scientific, and environmental significance of the subject matter. It emphasizes the indispensable function of ML in the prediction, surveillance, and regulation of GHG emissions across numerous sector agriculture, buildings, transportation, energy, and industrial processes while identifying the limitations related with data and ecological impact. The bibliometric will furnish the empirical evidence and framework needed for this developing topic, enlighten future research trajectories, and support policymakers and practitioners aiming to utilize ML technologies in climate mitigation initiatives.

2. REVIEW OF LITERATURE

A qualitative review of the existing literature is indispensable to offer context for the bibliometric discoveries. The Table 1 below summarizes key papers that have defined the intersection of ML and GHG research.

3. METHODOLOGY

Bibliometric analysis (BA), a quantitative and statistical method is often used in evaluating academic publications and related research themes by examining metrics like publication capacity, frequency in citation, patterns in authorship, co-authorship networks, and related keyword trends across scientific databases (Adediran et al., 2024; Adediran et al., 2025). BA is widely applied to evaluate the growth and influence of research, identifying main contributors, evolving topics, and collaboration of networks in diverse fields, including renewable energy, GHG emissions, climate change, and environmental technologies. This study applies BA in analyzing the studies landscape on GHG emissions in relation to ML, deep learning artificial intelligence. This method includes publication pattern tracking, mapping of collaboration networks, and identification of key institutions, authors, and countries to highlight the alignment with the Sustainable Development Goals (SDGs), research gaps, opportunities. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol was

Table 1: Summary of foundational research on ML applications in greenhouse gas mitigation

N/O	Author(s) and Year	Title/Focus	Methodology	Key Findings	Contribution
1	Zuo et al., 2024	Machine Learning for Advanced Emission Monitoring and Reduction Strategies in Fossil Fuel Power Plants	Systematic review of ML models	ML improves predictive monitoring (PEM) and the performance relies on data quality; key gaps in model transferability and interpretability exist.	Provides a structured framework for selecting ML techniques and identifies critical research gaps for the field.
2	Si et al., 2025	Long-term Evaluation of Machine Learning Based Methods for Air Emission Monitoring	Long-term performance analysis of ML models	ML models show sustained performance, but long-term validation protocols are needed.	Addresses a critical research gap by focusing on the long-term stability and reliability of ML solutions.
3	Alnaim et al., 2022	Evaluating Machine Learning and Remote Sensing in Monitoring NO ₂ Emission of Power Plants	Combines ML models with remote sensing data	ML and remote sensing can effectively monitor specific emissions like NO ₂ over a wide area.	Demonstrates a practical, large-scale approach to emissions monitoring by integrating multiple advanced technologies.
4	Barile et al., 2022	Damage monitoring of carbon fibre reinforced polymer composites using acoustic emission technique and deep learning	Deep learning with acoustic emission	Deep learning can be used to non-destructively monitor and identify damage in carbon fiber composites.	Offers a country-specific case study on using ML for CO ₂ emission forecasting and policy development.
5	Bhatt et al., 2023	Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques	Machine learning techniques	ML models can be used to forecast and inform mitigation strategies for global CO ₂ emissions.	Provides a global perspective on CO ₂ emission forecasting, complementing regional studies.
6	Mengand Noman, 2022	Predicting CO ₂ Emission Footprint Using AI through Machine Learning	AI through machine learning	AI can be effectively used to predict CO ₂ emission footprints.	Directly addresses the use of AI for predicting CO ₂ emission footprints.
7	Singhand Dubey, 2021	Deep Learning Model Based CO ₂ Emissions Prediction Using Vehicle Telematics Sensors Data	Deep learning with vehicle telematics sensors	Deep learning models can effectively predict CO ₂ emissions using real-time vehicle data.	Pioneers the use of deep learning and real-time vehicle telematics for specific, granular CO ₂ emission prediction.
8	Deng et al., 2021	Machine learning prediction of landslide deformation behaviour using acoustic emission and rainfall measurements	Machine learning with acoustic emission and rainfall data	ML can predict landslide deformation by analyzing acoustic emissions and rainfall.	Applies ML and acoustic emission to geotechnical engineering, specifically for landslide monitoring and risk prediction.
9	Anonna et al., 2023	Machine Learning-Based Prediction of U.S. CO ₂ Emissions	Machine learning models	Developed ML models for forecasting CO ₂ emissions to inform policy in the U.S.	Provides a country-specific case study on using ML for CO ₂ emission forecasting in the United States.
10	Sahithand Lal, 2024	Leveraging Machine Learning and Artificial Intelligence for Enhanced Carbon Capture and Storage (CCS)	Review of ML/AI applications for CCS	ML/AI can significantly enhance the efficiency and cost-effectiveness of CCS processes.	Provides a comprehensive overview of how ML and AI are being applied to improve the critical technologies of carbon capture and storage.
11	Forootan et al., 2022	Machine Learning and Deep Learning in Energy Systems: A Review	Review of ML and DL techniques	ML/DL can optimize energy systems across various sectors, from generation to consumption, for improved efficiency and reliability.	Provides a comprehensive review of ML/DL applications in energy systems, identifying key trends and future research directions.

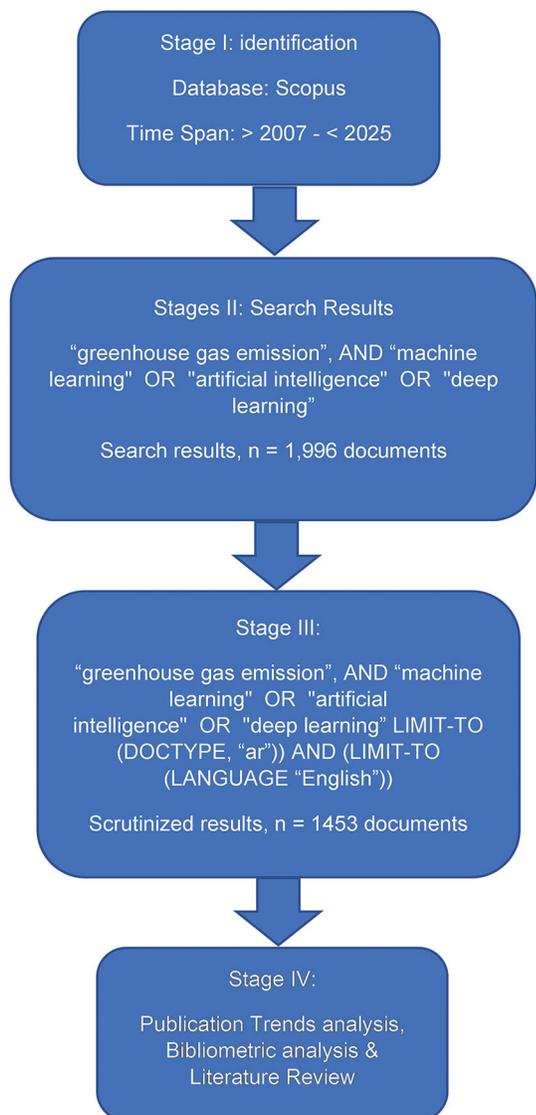
implemented for identifying and screening appropriate literature (Ayodele et al., 2024). Figure 1 below indicates the flowchart for the searching and screening process.

The preliminary data retrieval was initiated in Elsevier Scopus, one of the most all-inclusive databases for peer-reviewed literature. Boolean operators and search keywords were used as follows: (TITLE-ABS-KEY (“greenhouse gas emission”) AND TITLE-ABS-KEY (“machine learning” OR “artificial intelligence” OR “deep learning”)) AND PUBYEAR > 2007 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “cp”) OR LIMIT-TO (DOCTYPE, “re”) OR LIMIT-TO (DOCTYPE, “ch”)) AND (LIMIT-TO (LANGUAGE,

“English”)) AND (LIMIT-TO (SRCTYPE, “j”) OR LIMIT-TO (SRCTYPE, “p”) OR LIMIT-TO (SRCTYPE, “b”) OR LIMIT-TO (SRCTYPE, “k”)).

The initial search returned 1,996 documents prior to further refinement. Non-English publications, low-relevance records and unrelated topics, were all removed at the screening stage. The screening successfully eliminated 543 unrelated documents, which resulted in 1,453 documents, which were subsequently subjected to publication trends, and bibliometrics data analyses Also, author productivity, trends in publication output, institutional partnership, and funding sources were all analyzed. The BA software VOSviewer (v1.6.18) was used for keyword co-occurrence

Figure 1: Flow chart of study methodology on ML–GHG landscape



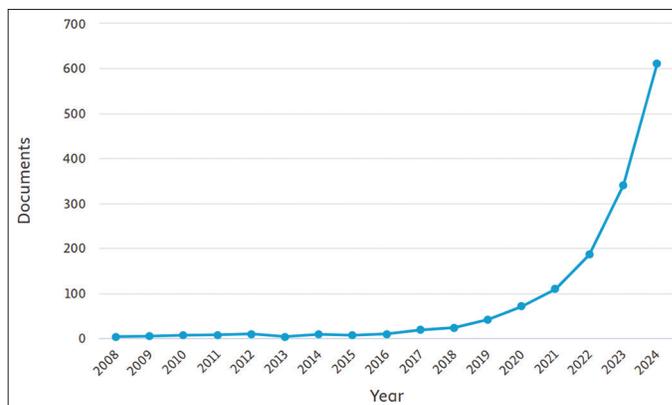
analysis, co-authorship mapping, and citation analysis. Data were exported to Microsoft Excel (2016) for publication trend analysis and subsequently to VOSviewer for visual mapping. This united PT (Publication Trend) and BA approach offer good insights into the development of research at the intersection of GHG emissions and AI-driven analytical methods, therefore revealing the priority areas and directions for future study.

4. RESULT AND DISCUSSION

4.1. General Publication Trends

The temporal distribution of published documents shown in Figure 2 from 2008 to 2024, demonstrate an exponential growth in research publications connecting (GHG) emissions with ML, artificial intelligence (AI), and deep learning (DL), emphasizing the growing role of data driven intelligence in the mitigation of climate change. In the early stage between 2008 and 2016, publication production was minimal, often in single digits, owing to the nascent adoption of ML/AI in environmental studies, the absence of large scale, high-quality emissions datasets, as well

Figure 2: Temporal distribution of published document growth on GHG and ML



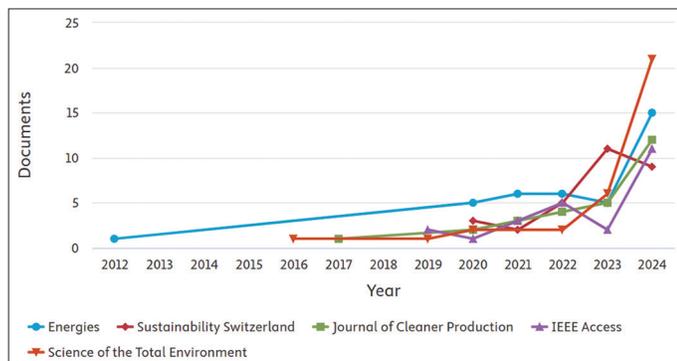
as the dominance of traditional statistical and simulation models. A slow uptick began around 2017, which coincided with policy milestones like the Sustainable Development Goals (SDGs) and Paris Agreement, as well as advancements in IoT-enabled sensor networks and satellite remote sensing, and the first wave of proof-of-concept research applying ML for energy efficiency optimization, carbon flux estimation, and air quality prediction. Further, the acceleration phase between 2020 and 2022 was driven by strong and heightened awareness that existed during COVID-19 with the link between human activity and emissions, the release of open access climate datasets by ESA, NASA, and NOAA, the mainstream acceptance of deep learning architectures, and increased cross disciplinary collaborations between climate scientists and AI experts. By 2021, yearly publications exceeded 100 and nearly doubled in the next year, signaling the entrance into a high-productivity phase. Notable, the most dramatic flow occurred in 2023 with over 340 publications and 2024 with about 610, prompted by urgency to meet climate targets by 2030, the integration of generative AI and hybrid models which combines ML with economic and physical climate systems, corporate investments and large-scale government, and the commercial readiness in real-time AI-powered emissions monitoring services. The keyword growth illustrates a shift from generic terms such as “climate change” and “machine learning” and toward specialized topics like “deep reinforcement learning for energy grids,” “carbon capture optimization,” and “GHG inventory automation.” From a BA perspective, the field has reached a mature, high-output stage with varied methodologies and thematic coverage spanning agriculture, renewable energy, industry, and transportation, with projections suggesting annual publications might exceed 1,000 by 2026. Further research however is anticipated to focus on the socio-technical integration merging AI with expanding datasets that will include the Global South, behavioral models governance, prioritizing the interpretable of AI for policy trust, and developing of standardized evaluation benchmarks, marking of a paradigm shift toward the blending of computational innovation with an urgent climate action in the narrow time window that is available to avert severe climate outcomes.

Based on the graphs provided in Figure 3, clear and accelerating progress is shown in academic publications relating to sustainability and environmental topics, with a notable increase

from 2020 to 2024. The line graph demonstrates a sharp upward trend across every featured journal, including prominent titles like Science of the Total Environment, Energies, and IEEE Access. While the graphs do not clearly mention “machine learning,” or “greenhouse gas” the publication channels themselves provide a strong contextual link. The quick increase in articles published in those journals focused on environmental science and clean production, like Science of the Total Environment and Journal of Cleaner Production, strongly reflects a growth in academic focus on critical issues such as GHG emissions and climate change mitigation. Instantaneously, the noteworthy rise in documents within IEEE Access being a journal renowned for its publishing works in computing and technology, directly suggests that scholars and researchers are increasingly applying advanced tools like ML to these environmental challenges. This convergence indicates a key trend in our modern-day research: leveraging the computational power to model climate change impacts, analyze complex environmental data, and develop optimized solutions for renewable energy systems vis-à-vis emissions reduction. So, the observed publication growth serves as an influential indicator towards a global scientific effort in combating climate change, where the application of ML is becoming a foundational approach in the quest to monitor, understand, and ultimately reduce GHG emissions.

To further support this trend is the important output from key scholars or researchers in the field. From Figure 4, a bar chart of documents by author reveals that several individuals are leading this publication surge. Wang, X. stands out as the most prolific author, with five documents. Following closely are Bellandi, G., Ciais, P., Creutzig, F., and Mateu, C., each having contributed four documents. This concentration of work among a handful of highly active researchers demonstrates the concerted effort and influence of specific experts in driving the discourse around environmental issues and the application of machine learning. The contributions of these named authors are a testament to the focused and high-impact research taking place in these critical areas. The rapid application of ML is particularly evident in the diverse range of journals. For instance, articles in Energies likely explore how ML algorithms can optimize smart grids to integrate renewable energy sources more efficiently, thereby reducing reliance on fossil fuels. Similarly, publications in Science of the Total Environment may use ML to analyze satellite imagery and

Figure 3: Distribution of sources titles on Greenhouse gas and ML research



sensor data to predict and map areas of high pollution, including greenhouse gas concentrations. The work of authors like Wang, X. and the others listed likely contributes to this interdisciplinary research, bridging the gap between computer science and environmental sustainability. This synergy between diverse fields is a crucial factor in the observed publication growth, highlighting a comprehensive and multifaceted approach to addressing the global climate crisis.

A review of the authors’ affiliations shown in Figure 5, further contextualizes this research effort, providing insight into the key global players driving innovation in this field. The bar chart showing documents by affiliation highlights the institutions with the most significant output. Notably, the Chinese Academy of Sciences is the most prominent, with 30 documents, followed by the Ministry of Education of the People’s Republic of China with 26. This dominant presence suggests a strong, coordinated national research initiative and substantial investment from China in addressing climate change and leveraging technology like ML to do so. This kind of concentrated effort from major state-backed institutions often accelerates a research field’s growth. Other significant contributors include CNRS Centre National de la Recherche Scientifique (16 documents) from France, Texas A and M University (15 documents) from the United States, and the National University of Singapore (14 documents). The inclusion of these renowned international universities and research centers demonstrates that the research is a truly global effort. The geographical diversity of the leading institutions underscores the worldwide recognition of climate change as a critical issue, and the collaborative nature of the scientific community in applying

Figure 4: Top 5 prolific authors on ML and GHG research

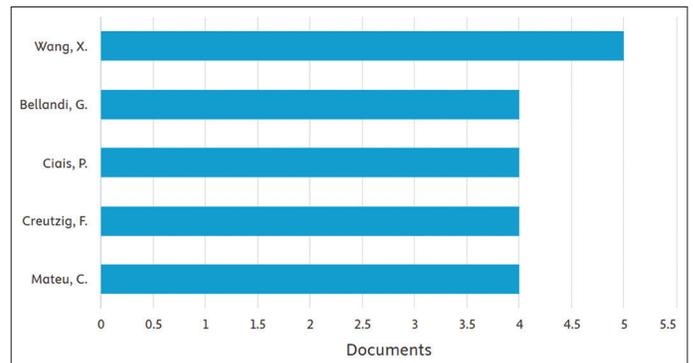
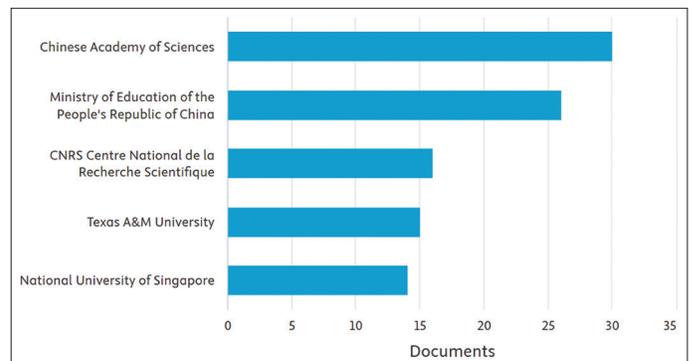


Figure 5: Top 5 prolific affiliations on ML and GHG research



advanced technologies to find solutions. The data not only shows who is publishing, but also where the major hubs of this crucial interdisciplinary research are located.

This institutional focus is further reinforced in Figure 6, by a separate country-level analysis, which shows the United States leading with 265 documents, followed closely by China with 253. The inclusion of the United States as a top contributor adds significant weight to the argument, demonstrating that the research is not concentrated in one region but is a highly competitive and collaborative effort between major global powers. The data also reveals substantial contributions from India (201 documents), the United Kingdom (123 documents), and Canada (102 documents). This broad geographical diversity, encompassing nations across Asia, Europe, and North America, underscores the worldwide recognition of climate change as a critical issue. The collaborative nature of the scientific community in applying advanced technologies to find solutions is evident, and the data not only shows who is publishing but also where the major hubs of this crucial interdisciplinary research are located.

From Figure 7, the overwhelming majority of these publications are in the form of Articles (810 documents, or 55.7% of the total) and Conference Papers (413 documents, or 28.4%). This emphasis on articles and conference papers suggests that the research is largely focused on disseminating new findings and methodologies in a peer-reviewed and time-sensitive manner, which is crucial for a rapidly evolving field. The fewer numbers of Reviews (148 documents) and Book Chapters (82 documents) indicate that the field is still in a phase of rapid expansion and discovery, with a

greater emphasis on new research rather than on synthesis and consolidation of existing knowledge. This publication pattern is a strong indicator of a dynamic and active research landscape, where the primary goal is to push the boundaries of knowledge as quickly as possible.

The interdisciplinary nature of this research is a key driver of its growth. The subject area analysis seen in Figure 8, reveals a significant concentration of documents in Engineering (18.8%) and Computer Science (15.2%), which together account for a third of the total publications. This is a powerful justification for the involvement of machine learning, as these are the fields that provide the technological foundation for developing predictive models, data analysis tools, and optimized systems. The strong presence of Environmental Science (13.1%) and Energy (12.6%) as top subject areas confirms that the focus of this work is squarely on tackling climate and sustainability challenges. The synergy between these technical and applied sciences is the central theme of this research trend. The data demonstrates that researchers are not just studying environmental problems, but actively developing and applying advanced computational and engineering solutions to them, marking a shift from theoretical discourse to practical, technology-driven innovation.

The significant funding landscape further solidifies the global commitment to this research. As seen in Figure 9, a bar chart of funding sponsors shows a strong investment from several key organizations. The National Natural Science Foundation of China is a dominant sponsor with 114 documents, reinforcing the country's leading role in this research area. The European

Figure 6: Top 5 prolific countries on ML and GHG research

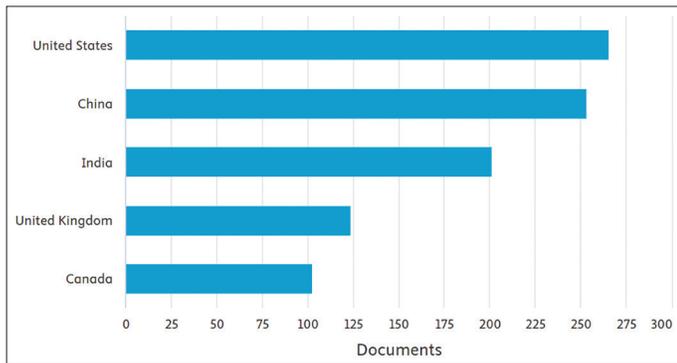


Figure 7: Document types on ML and GHG research

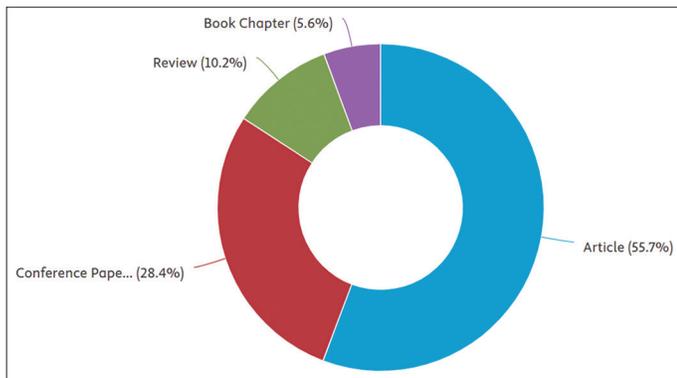


Figure 8: Subject area categories for publications on ML and GHG research

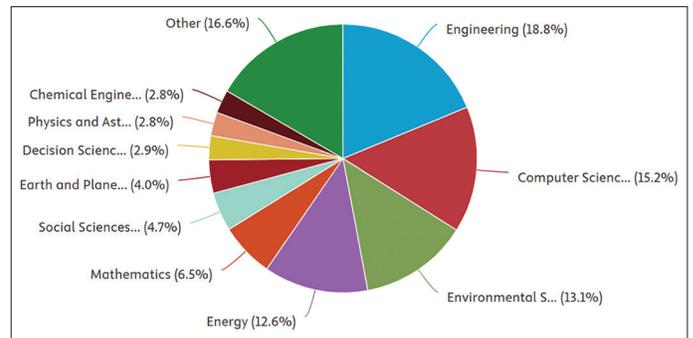
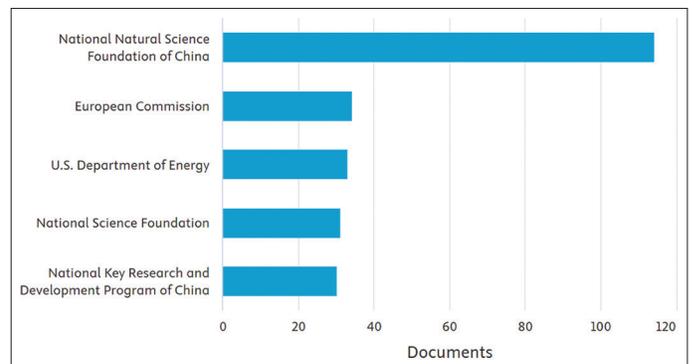


Figure 9: Top Funding Sources on ML and GHG research (2008-2024)



Commission (34 documents) and the U.S. Department of Energy (33 documents) are also major contributors, highlighting the prioritization of this research by large-scale governmental and supranational bodies. The presence of the National Science Foundation and the National Key Research and Development Program of China with 31 and 30 documents respectively, further justifies the robust support for research in this field. The sheer volume of documents associated with these funding bodies proves that the observed increase in publications is not just an organic trend, but a direct result of substantial, targeted financial investment from major global players who recognize the urgency and importance of this interdisciplinary work.

4.1.1. Social network analysis

The aspect of the study was done through VOSviewer analytical tools. This too is considered essential for bibliometric research relating to ML and greenhouse gas mitigation. Also, a Keyword Co-occurrence (KCO) analysis is crucial to identify the topic’s main research themes and trends. By studying how phrases like “machine learning,” “greenhouse gas mitigation,” “carbon capture,” and “energy efficiency” appear together, you can generate a thematic map of the subject to indicate how concepts are related and where research is going. Without VOSviewer, examining the massive amount of data would be challenging, making it hard to find these key insights about the field’s intellectual structure.

4.1.1.1. Co-authorship analysis

Co-authorship analysis is vital as it explicitly identifies outstanding researchers, their organizations and countries in this interdisciplinary topic. Hence, mapping out their collaboration networks indicates who is collaborating with whom, any significant research groupings, and the potential partnerships.

As seen in Figures 10 and 11, the analysis of author co-authorship networks, based on threshold of five documents, recognizes a tiny yet noteworthy core group of 13 very productive authors from a total of 6,162. This shows that while numerous academics are engaging in the application of ML for greenhouse gas mitigation,

a limited number of them seem to continuously publish their results. The link strength of data specifies a distinct collaboration hierarchy inside this core group. Four authors, each holding a connection strength of 12, are the most interrelated which likely represent important actors in this research network. Their substantial connection strength means regular co-authorship with other prolific writers, which implies they might be spearheading momentous partnerships and influencing the trajectory of the area. Equally, the five authors lacking connection strength even while they fulfill the document criterion, seem to be operating in isolation. This examination, shown in the diagrams, further demonstrates that the field is not a steadily collaborative environment. Rather, it is defined by an inadequate amount of cohesive research clusters directed by few groups of prominent authors. Notably, such concerted collaboration may serve as either a liability or an asset: an asset owing to its potential of promoting profound expertise while targeting research, or a liability as it could limit the exchange of ideas and obstruct the expansion of a varied research community.

As shown in Figure 12, the study of organizational collaboration recognized 62 highly significant and productive establishments from a big dataset of 2,320 organizations based on a minimum threshold of 5 documents and 10 citations. This infers that fewer organizations are leading ML for GHG reduction research. Most prominent is the network’s main connected component having 42 elements, demonstrating that almost two-thirds of these noteworthy organizations work together while creating a cohesive research network. The field’s development is subjected to these institutions’ close partnership. Hence, this represents a collective effort in solving difficult, interdisciplinary issues that no single institution can fix/handle. This cooperation also accelerates studies by sharing resources knowledge, and skills, resulting in a more impactful and robust findings from 2008 to 2024. of institutional collaboration, the least standards of 10 citations and 5 documents

Figure 10: Network visualization map for collaborating authors on ML and GHG

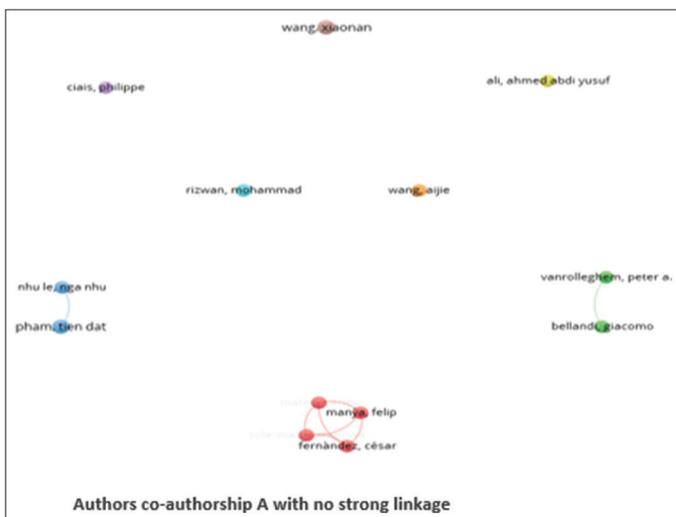


Figure 11: Network visualisation map for strongest collaborating authors on ML and GHG

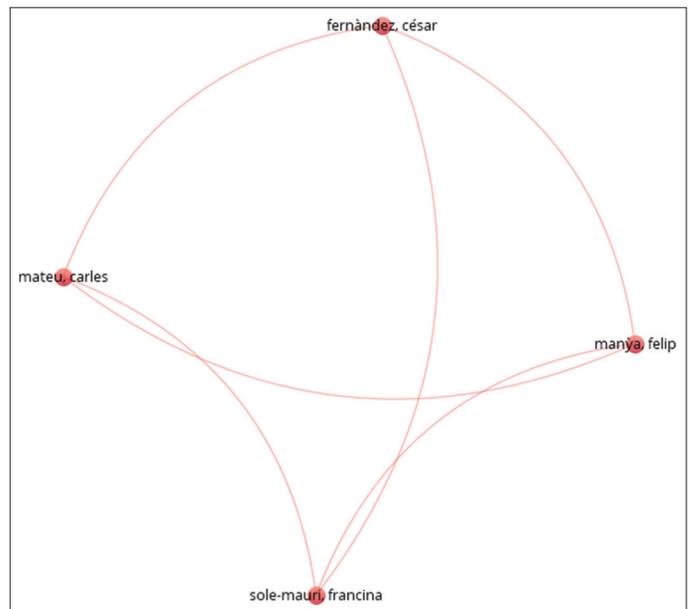
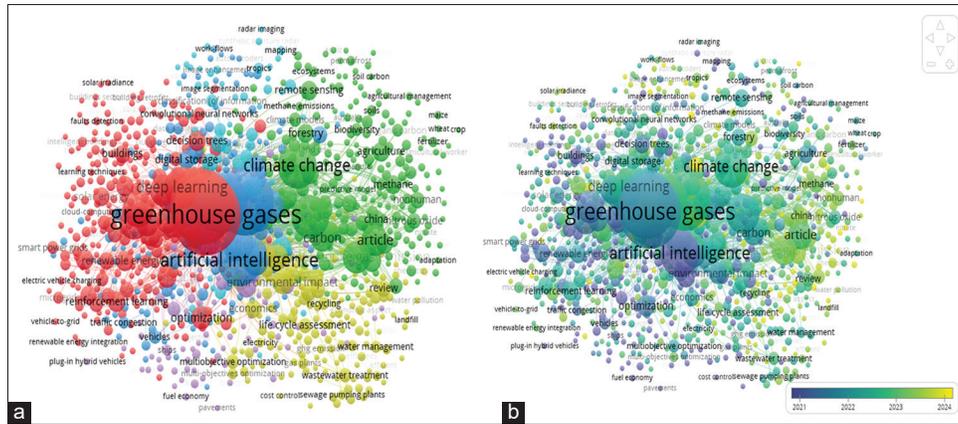


Figure 14: (a and b) Network visualization map for important keywords on ML and GHG

5. CONCLUSION

Based on a comprehensive bibliometric analysis of publications from 2008 to 2024, this study provides a clear and detailed map of the research landscape concerning the integration of ML into GHG for mitigation efforts. The findings jointly prove a field that is not just experiencing unprecedented growth but also growing into a well-structured and internationally coordinated scientific endeavor. The exponential rise in related and useful publications, predominantly in the form of conference papers and articles, is a powerful pointer of an active research and dynamic community that focused on disseminating novel findings and methodologies. Such rapid development is determined by a distinct interdisciplinary synergy. The dominant subject area Environmental Science, Engineering, Computer Science, and Energy are a testament to the convergence of technical innovation and applied environmental problem-solving. This fusion is the central characteristic of the field, highlighting a shift from theoretical discussion to practical, technology-driven solutions for climate change. Additionally, analyzing of relevant institutions, authors, and countries discloses a clearer pattern of concentrated proficiency and significant institutional support. The leadership of countries like the China and United States is a direct reflection of a substantial financial backing from major funding bodies like the U.S. Department of Energy and the National Natural Science Foundation of China and. Hence, this robust funding landscape also confirms that the growth in publications is strongly linked to a strategic, large-scale investment that underscores the global political and scientific commitment to this research. Future study in this domain could further build upon these discoveries by delving deeper into precise aspects of related research ecosystem. A detailed reference analysis could aid the identification of the most impactful studies/investigations and specific ML techniques that are driving this field forward. Having a qualitative review of the diverse key publications could also explore the practical applications and real-world outcomes of the research thereby moving beyond publication trends into assessing tangible impact. Finally, future studies could further analyze collaborative network patterns amid different institutions and countries to offer a better understanding of how international collaborations can better sharpening the future of this critical field.

This study offers an extensive bibliometric analysis of worldwide research trends linking machine learning (ML) and greenhouse gas (GHG) emission mitigation. The research utilized data from the Scopus database to delineate publication growth, identify principal contributors, analyze collaboration networks, and examine keyword co-occurrence trends. Research indicates a swift increase in publications post-2017, propelled by global climate legislation and technical innovations, resulting in substantial contributions from nations such as China and the United States. The research underscores novel machine learning applications in emission forecasting, carbon capture optimization, and real-time monitoring, accentuating machine learning's contribution to improving predictive precision and operational efficacy in environmental management.

This study provides one of the earliest and most systematic bibliometric analyses of the role of machine learning in global greenhouse gas mitigation initiatives. It aids in identifying key research clusters, prominent authors, and funding organizations, so assisting policymakers, scientists, and industry professionals in comprehending the development and interdisciplinary characteristics of this domain. Furthermore, the study's depiction of institutional and international cooperation offers significant insights into the manner in which global partnerships expedite innovation in environmental AI. The report lays a basis for future studies by mapping keyword patterns and emerging research issues, directing academics toward priority areas such as interpretable AI, carbon capture technologies, and sustainable data-driven solutions for climate action.

While this study provides comprehensive insights into the research landscape of the Machine Learning and Greenhouse Gas Emission (MLGHG), certain limitations have been identified which can be addressed in future studies. Firstly, the data source was limited to the Scopus database. This database is very complete, but it may not include all of the relevant publications that are indexed in other academic repositories like PubMed or Google Scholar. Some major works may not have been included enough, especially those that were published in journals that are not indexed by Scopus and are not in English or are only limited to a specific region. Secondly, the bibliometric analysis is affected by the search terms and filtering criteria that were used. There

is a chance that some relevant papers were missed because of differences in terminology or metadata discrepancies, even though a lot of care was taken in choosing and filtering keywords. Thirdly, the keyword co-occurrence and cluster analyses gave us useful information on subject areas but are limited since the keywords depend on terms chosen by the author and may not fully capture the intricacies of new or interdisciplinary studies.

6. ACKNOWLEDGEMENTS

This research is funded by Sunway University Early Career Research Grant Scheme with project code GRTIN-ECR(02)-DCIS-10-2024.

REFERENCES

- Adediran, A.O., Ajobade, S.S., Zaidi, A., Bashir, F.M., Falude, E., Dodo, Y.A., Jasser, M.B. (2024), A science mapping analysis of energy efficiency and affordable housing research. *International Journal of Energy Economics and Policy*, 14(6), 436-449.
- Adediran, A.O., Mohd Aini, A., Ajobade, S.M. (2025), Adoption of artificial intelligence in property management transactions: A systematic review and trend analysis. *Property Management*, 2025(1), 1-33.
- Ahmed, A.H.A., Jin, W., Ali, M.A.H. (2024), Comparative analysis of intelligent models for predicting compressive strength in recycled aggregate concrete. *Modeling Earth Systems and Environment*, 10(4), 5273-5291.
- Alnaim, A., Sun, Z., Tong, D. (2022), Evaluating machine learning and remote sensing in monitoring NO₂ emission of power plants. *Remote Sensing*, 14(3), 729.
- Anonna, F.R., Mohaimin, M.R., Ahmed, A., Nayeem, M.B., Akter, R., Alam, S., Nasiruddin, M., Hossain, M.S. (2023), Machine learning-based prediction of US CO₂ emissions: Developing models for forecasting and sustainable policy formulation. *Journal of Environmental and Agricultural Studies*, 4(3), 85-99.
- Ayodele, O.E., Ajobade, T.I., Oyeboode, O.J., Akintoye, K.A., Ajayi, B.A., Adediran, A.O., Ajobade, S.S.M., Majeed, A.P.A., Luo, Y. (2024), A Bibliometric and social media analysis of the nexus between blockchain technology and finance. *International Conference on Intelligent Manufacturing and Robotics*, 1316, 553-563.
- Barile, C., Casavola, C., Pappalettera, G., Kannan, V.P. (2022), Damage monitoring of carbon fibre reinforced polymer composites using acoustic emission technique and deep learning. *Composite Structures*, 292, 115629.
- Bastviken, D., Wilk, J., Duc, N.T., Gålfalk, M., Karlson, M., Neset, T.S., Opach, T., Enrich-Prast, A., Sundgren, I. (2022), Critical method needs in measuring greenhouse gas fluxes. *Environmental Research Letters*, 17(10), 104009.
- Bhatt, H., Davawala, M., Joshi, T., Shah, M., Unnarkat, A. (2023), Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques. *Cleaner Chemical Engineering*, 5, 100095.
- Deng, L., Smith, A., Dixon, N., Yuan, H. (2021), Machine learning prediction of landslide deformation behaviour using acoustic emission and rainfall measurements. *Engineering Geology*, 293, 106315.
- Devarakota, P., Tsesmetzis, N., Alpak, F.O., Gala, A., Hohl, D. (2025), AI and the Net-zero Journey: Energy Demand, Emissions, and the Potential for Transition. [arXiv preprint] arXiv:2507.10750.
- Dong, Z., Wang, S., Jiang, Y., Xing, J., Ding, D., Zhang, F., Yin, D., Song, Q., An, J., Wang, H. (2025), A forecasting tool for optimized emission control strategies to achieve short-term air quality attainment. *Journal of Environmental Management*, 373, 123916.
- Emad-Eldeen, A., Azim, M.A., Abdelsattar, M., AbdelMoety, A. (2024), Utilizing machine learning and deep learning for enhanced supercapacitor performance prediction. *Journal of Energy Storage*, 100, 113556.
- Essamlali, I., Nhaila, H., El Khaili, M. (2024), Supervised machine learning approaches for predicting key pollutants and for the sustainable enhancement of urban air quality: A systematic review. *Sustainability*, 16(3), 976.
- Farooq, O., Khan, A. (2025), The Impact of machine learning on climate change modeling and environmental sustainability. *Artificial Intelligence and Machine Learning Review*, 6(1), 8-16.
- Filonchik, M., Peterson, M.P., Zhang, L., Hurynovich, V., He, Y. (2024), Greenhouse gases emissions and global climate change: Examining the influence of CO₂, CH₄, and N₂O. *Science of The Total Environment*, 935, 173359.
- Forootan, M.M., Larki, I., Zahedi, R., Ahmadi, A. (2022), Machine learning and deep learning in energy systems: A review. *Sustainability*, 14(8), 4832.
- Hamdan, W., Alsuaiah, H. (2024), Research output, key topics, and trends in productivity, visibility, and collaboration in social sciences research on COVID-19: A scientometric analysis and visualization. *Sage Open*, 14(4), 21582440241286217.
- Haque, M.E., Hossain, A., Alam, M.S., Siam, A.H., Rabbi, S.M.F., Rahman, M.M. (2024), Optimizing DDoS Detection in SDNs Through Machine Learning Models. In: 2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks.
- Harsányi, E., Mirzaei, M., Arshad, S., Alsilibe, F., Vad, A., Nagy, A., Ratonyi, T., Gorji, M., Al-Dalahme, M., Mohammed, S. (2024), Assessment of advanced machine and deep learning approaches for predicting CO₂ emissions from agricultural lands: insights across diverse agroclimatic zones. *Earth Systems and Environment*, 8(4), 1109-1125.
- Hasan, S.M., Islam, T., Saifuzzaman, M., Ahmed, K.R., Huang, C.H., Shahid, A.R. (2025), Carbon emission quantification of machine learning: A review. *IEEE Transactions on Sustainable Computing*, 10(6), 1085-1102.
- Hassan, M.M., Elmekky, Y.M., Youssef, S.M., Shoieb, D.A. (2025), Enhancing Prediction of Bioenergy Yield Production Using AI-based Identification of Biomass Species in Promoting Sustainable Biofuels and Boosting Agriculture Productivity. In: 2025 International Conference on Machine Intelligence and Smart Innovation.
- Hassan, W., Duarte, A.E. (2024), Bibliometric analysis: A few suggestions. *Current Problems in Cardiology*, 49(8), 102640.
- Hu, K., Zhang, Q., Gong, S., Zhang, F., Weng, L., Jiang, S., Xia, M. (2024), A review of anthropogenic ground-level carbon emissions based on satellite data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 8339-8357.
- Kek, H.Y., Tan, H., Othman, M.H.D., Nyakuma, B.B., Ho, W.S., Sheng, D.D.C.V., Kang, H.S., Chan, Y.T., Lim, N.H.A.S., Leng, P.C. (2024), Critical review on airborne microplastics: An indoor air contaminant of emerging concern. *Environmental Research*, 245, 118055.
- Li, X., Zhang, X. (2023), A comparative study of statistical and machine learning models on carbon dioxide emissions prediction of China. *Environmental Science and Pollution Research*, 30(55), 117485-117502.
- Liaskoni, M.D. (2025), Modelling anthropogenic and natural emissions and their impact on pollutant concentrations. *Univerzita Karlova, Matematicko-Fyzikální Fakulta*, 12005(1), 1-68.
- Liu, L., Zhou, W., Guan, K., Peng, B., Xu, S., Tang, J., Zhu, Q., Till, J.,

- Jia, X., Jiang, C. (2024), Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems. *Nature Communications*, 15(1), 357.
- Luccioni, A.S., Hernandez-Garcia, A. (2023), Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning. [arXiv preprint arXiv:2302.08476].
- Mamo, A.A., Mekonen, K.A., Golda, A., Behera, R.K., Kumar, K. (2025), Unlocking the power of machine learning in big data: A scoping survey. *Data Science and Management*, 8(4), 519-535.
- Meng, Y., Noman, H. (2022), Predicting CO₂ emission footprint using AI through machine learning. *Atmosphere*, 13(11), 1871.
- Nyakuma, B.B., Mahyon, N.I., Chiong, M.S., Rajoo, S., Pesiridis, A., Wong, S.L., Martinez-Botas, R. (2023), Recovery and utilisation of waste heat from flue/exhaust gases: A bibliometric analysis (2010–2022), *Environmental Science and Pollution Research*, 30(39), 90522-90546.
- Sahith, J.K., Lal, B. (2024), Leveraging machine learning and artificial intelligence for enhanced carbon capture and storage (CCS). *Gas Hydrate in Carbon Capture, Transportation and Storage*, 38, 159-196.
- Saleh, H.M., Hassan, A.I. (2024), The challenges of sustainable energy transition: A focus on renewable energy. *Applied Chemical Engineering*, 7(2), 2084.
- Si, M., Wiens, B.M., Du, K. (2025), Long-term evaluation of machine learning based methods for air emission monitoring. *Environmental Management*, 75(3), 680-693.
- Singh, K. (2024), Evaluating Machine Learning Models for Defect Rate Prediction and Maintenance Classification in Industrial Systems Dublin, Ireland: National College of Ireland.
- Singh, M., Dubey, R.K. (2021), Deep learning model based CO₂ emissions prediction using vehicle telematics sensors data. *IEEE Transactions on Intelligent Vehicles*, 8(1), 768-777.
- Soeder, D.J. (2025), Greenhouse gas and climate change. In: *Energy Futures: The Story of Fossil Fuel, Greenhouse Gas, and Climate Change*. Germany: Springer, pp. 97-141.
- Tahir Bahadur, F., Rasool Shah, S., Rao Nidamanuri, R. (2025), A Brief outline of indoor air quality: Monitoring, modeling, and impacts. *Journal of Environmental Engineering*, 151(5), 03125001.
- Tan, H., Wong, K.Y., Othman, M.H.D., Nyakuma, B.B., Sheng, D.D.C.V., Kek, H.Y., Ho, W.S., Hashim, H., Chiong, M.C., Zubir, M.A. (2023), Does human movement-induced airflow elevate infection risk in burn patient's isolation ward? A validated dynamics numerical simulation approach. *Energy and Buildings*, 283, 112810.
- Udoh, S. (2024), Applied geochemistry and engineering geology: Implications for civil infrastructure, environmental sustainability, and resource management. *Applied Sciences, Computing, and Energy*, 1(1), 86-99.
- Wei, X., Wang, X., Zhu, T., Gong, Z. (2021), Fusion prediction model of atmospheric pollutant based on self-organized feature. *IEEE Access*, 9, 8110-8120.
- Zuo, Z., Niu, Y., Li, J., Fu, H., Zhou, M. (2024), Machine learning for advanced emission monitoring and reduction strategies in fossil fuel power plants. *Applied Sciences*, 14(18), 8442.