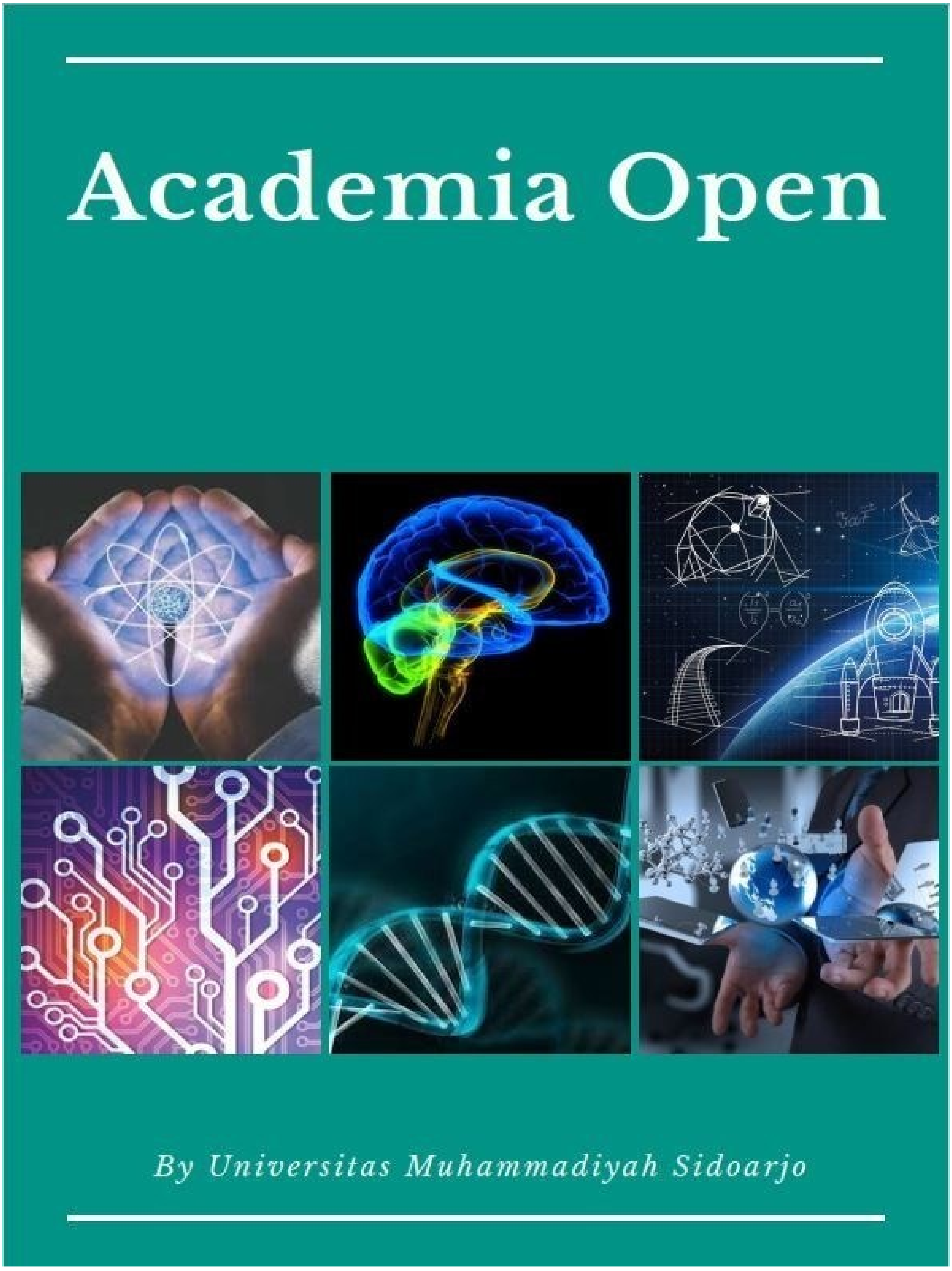
Academia Open

Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056



ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Academia Open

Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056

Table Of Contents

Academia Open
Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Academia Open

Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056

EDITORIAL TEAM

Complete list of editorial team (link)

Complete list of indexing services for this journal (link)

How to submit to this journal (link)

Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056

Article information

Check this article update (crossmark)

Check this article impact (*)

Save this article to Mendeley

(*) Time for indexing process is various, depends on indexing database platform

Academia Open

Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.13056

AI-Driven Genomic Modeling for Predicting Biological Responses to Environmental Pollution

Maryam Jasim Hasan, Maryam.J.28@uomustansiriyah.edu.iq,(1)
Department of Biology science Mustansiriyah University

Juomana Jabbar Saeed, J.Juomana.S@uomustansiriyah.edu.iq,(2)
Department of Biology science Mustansiriyah University

Reyam Naji Ajmi, reyam80a@yahoo.com,(3)
Department of Biology Science, Mustansiriyah University, POX 46079, Iraq-Baghdad

(1) Corresponding author

Abstract

General Background: Environmental pollution from heavy metals and organic contaminants has escalated globally, generating complex biological disruptions across molecular, cellular, and ecological levels. **Specific Background:** Advances in high-throughput sequencing and multi-omics technologies have revealed that pollutants induce genomic and epigenetic alterations, yet interpreting these multidimensional datasets remains challenging. **Knowledge Gap:** Existing studies lack integrated analytical frameworks capable of linking pollutant exposure with genome-wide molecular responses in a precise and predictive manner. **Aims:** This review synthesizes current evidence on pollutant-induced molecular changes and evaluates the potential of artificial intelligence (AI) to model and predict biological responses. **Results:** Recent applications of machine learning—such as Random Forests, Support Vector Machines, and Convolutional Neural Networks—demonstrate strong performance in identifying key genetic markers, forecasting epigenetic modifications, and estimating organismal vulnerability before clinical symptoms appear. **Novelty:** This article highlights the emerging role of AI-driven genomic modeling as a transformative approach that integrates environmental and genomic datasets to capture multilevel biological responses with high accuracy. **Implications:** The integration of AI with genomics offers a proactive strategy for early detection of pollution-induced molecular changes, enhances environmental risk assessment, and informs targeted remediation, biodiversity protection, and long-term ecosystem management.

Highlight :

- Environmental pollutants cause DNA methylation changes, histone modifications, and disruptions in cellular defense mechanisms.
- Machine learning algorithms predict epigenetic changes and identify genetic markers before physiological symptoms manifest.
- AI integration enables early pollution detection, bioremediation support, and proactive ecosystem management..

Keywords : Environmental Pollution, Artificial Intelligence, Environmental Genomics, Epigenetic Change, Predictive Modeling

Published date: 2025-12-11

Introduction

The rate of environmental pollution has markedly increased during the last few decades due to increasing population, unplanned urbanization, intensive industrial activity, and imbalanced uses of fertilizers and pesticides. Generally, pollutants tend to accumulate either directly or indirectly in different components of the environment, affecting the living organisms and ecosystem in many ways [1].

While organic and metallic contaminants act through internal homeostatic disruption in animals, heavy metals like lead, cadmium, and mercury that bioaccumulation in tissues promote liver, kidney, and neurological damage. Organic contaminants, like PAHs and pesticides, interact with molecular pathways that induce oxidative stress to cause DNA damage and heritable genetic changes [2].

A more detailed understanding of these risks requires the unraveling of complex interactions between pollutants and genes. That is why scientists apply the approach of systems biology: it integrates molecular alterations with cellular, physiological, and behavioral responses of organisms [3], means of systems biology, there is a prospect of multi-level understanding of organism responses to pollution, and it makes possible the development of specific biomarkers for early detection of environmental contamination [4].

Newly developed NGS, combined with a number of other -omics techniques, provides volumes of genomic-functional data. These are complex data that need analysis with the aid of computational tools handling nonlinear relationships among variables. Thus, AI and machine learning form an intrinsic part in the analysis of environmental-biological data for deciphering the molecular and physiological response to pollution with great accuracy [5]

Materials and Methods

The purpose of this article is to investigate the effects of environmental pollutants, particularly heavy metals and organic contaminants, on biological systems using modern computational approaches. The study aims to understand how pollutants interact with genetic and epigenetic mechanisms, affect molecular and physiological processes, and potentially lead to oxidative stress, DNA damage, and other health impacts. Additionally, the research seeks to explore the use of artificial intelligence and machine learning to analyze complex biological data and develop predictive models for assessing environmental risks and identifying biomarkers for early detection of pollution-related effects.The methodology combines environmental monitoring with biological analysis and computational modeling. Environmental samples such as air, water, and soil are analyzed to determine the presence and concentration of pollutants, while biological samples from exposed populations are collected to study genomic and epigenetic changes. High-throughput sequencing and other omics techniques are used to generate comprehensive datasets, which are then preprocessed to remove noise and integrate multiple data types. Machine learning algorithms are applied to identify patterns, correlations, and predictive indicators of pollutant effects. Finally, the results are interpreted in the context of biological pathways and existing literature to ensure scientific relevance and reliability

Results and Discussion

The review below discusses the application of AI-based approaches in genomics to investigate multilevel effects of pollution on animal physiology, including:

Consider various mechanisms of genetic and epigenetic environmental pollution.

Explain how AI algorithms help in the analysis of large-scale genomic and environmental datasets.

Various applications of AI to environmental monitoring, bioremediation, and predictive risk assessment will be reviewed.

Discussion of current challenges and perspectives for the future in the integration of AI into environmental genomics.

A. Applications of AI-Based Genomic Modeling in the Environment and Biology

AI approaches in genomic studies open new frontiers toward the discovery of complex biological responses in organisms due to environmental pollution this allows multilevel data analysis, from changes in genomes and epigenomes at the molecular level to higher levels of organization, including cellular function and physiological impact on individuals and communities [6].

1.1 Natural Environmental Risk Assessment : AI-based models can also pinpoint sensitive species and organisms due to different kinds of pollution, such as heavy metals, organic contaminants, and micro plastics. Advanced modeling thus allows the forecast of environmental risk well before physiological and behavioral symptoms even manifest. Thus, genomic and environmental data could give rise to early protective actions on ecosystems [7].

1.2 Heavy Metal Resistance Genes Analysis: Certain applications of AI in soil and water bioremediation enable the identification of genes in bacteria responsible for heavy metal tolerance, such as CzcA for cadmium and zinc, MerA for mercury, and PbrT for lead detoxification, and plant genes involved in detoxification and remediation of heavy metals. The information obtained from the analysis will indicate the exact environmental strategies to be adopted concerning decontamination methods of both soil and water and help in the selection of those species that could adapt or resist pollution [8].

1.3 Supporting Bioremediation Application: AI-based models will speed up the finding of microbes, fungi, and plants capable of degrading organic pollutants or converting toxic metals into their harmless forms. Accordingly, genetic information coupled with enzymatic activity would contribute to enhanced efficiency, accuracy, and sustainability of environmental remediation programs[9].

1.4 Intelligent environmental monitoring: Advanced artificial intelligence algorithms now have the capability of monitoring, in real time, alterations in gene expression and epigenetic modifications which are caused by exposure to pollutants; such changes represent early biomarkers of the presence of environmental stressors. Integration with environmental sensor networks now provides a continuing monitoring system that offers rapid assessments of environmental status with proactive ecosystem management rather than post facto reactions to damage [10].

1.5 The Molecular Basis of Pollution-Related Changes: General mechanisms involved in the molecular effects of pollutants include, among others, DNA damage, oxidative stress, and epigenetic modifications. These usually arise due to impaired DNA repair capabilities, disruption in normal detoxification pathways, and instability of cells. Molecular alterations at the organ level have consequences for the health status of the whole organism and give accurate signals on the relative sensitivity of different species to environmental pollutants, helping to identify species that are more vulnerable than others [11].

1.6 Early Detection and Prediction of Multilevel Responses: Advanced algorithms in the form of random forests, radial support machines, convolution neural networks, and autoencoders anticipate the tissue- and organ-level effects from changes at the molecular level. These systems also allow integration with genomic, genetic, and environmental data for a full risk assessment to establish the organismal responses to the pollutants well before the physiological symptoms are manifested and thereby enable timely and efficient intervention [12].

1.7 Integrating Environmental and Genomic Data: The proposed approach will complement environmental measurements with relevant genomic data in such a way as to provide genetic and heritable indicators of an organism's resilience or sensitivity to specific forms of pollutants. It will also open up avenues to targeted scientific interventions that substantially enhance ecosystem protection by way of effective bioremediation techniques, biodiversity conservation, and comprehensive and accurate environmental risk assessments

1.8 Integration of Environmental and Genomic Data: The concept proposed herein will supplement environmental measurements with appropriate genomic data in a manner that delivers genetic and heritable indicators of the resilience or sensitivity of an organism to particular forms of pollutants. Moreover, this will pave the way for targeted scientific interventions that shall go a long way towards improving ecosystem protection by developing effective bioremediation techniques, conserving biodiversity, and carrying out comprehensive and accurate environmental risk assessments [13].

B. Current Challenges and Limitations: Notwithstanding the promising advances in applying AI and genomic modeling to studies of pollution, several scientific and technical challenges must be overcome before this work can be practically applied. These include but are not limited to those related to the precise scientific and technical solutions described in [14].

2.1 Data Quality and Standardization: Most of the genomic models developed so far are sensitive to the quality of input data. Even small changes in methods of sampling, genome sequencing, or measurement of a pollutant may lead to inconsistent results[15]. Therefore, standardization in environmental and genomic data collection is an important preliminary step toward study comparison for the generation of valid and reliable predictions.

2.2 Computational Complexity of Multidimensional Data: Large and complex data sets consist of many layers, both genomic and environmental; their processing thus requires great computational power. Nonlinear correlations of intensive interactions between variables require highly advanced algorithms. Therefore, this is actually the main practical challenge: processing and interpretation time.

2.3 Biodiversity and Species Variation: Because of individual differences: The same species can react differently to pollution. The implication is that generalization is difficult and the use of models, which can handle variation at both the individual and species levels without loss in predictive accuracy, is justified.

2.4 Interpretation of Deep Learning Models: While highly accurate in performance, deep learning models normally act like a "black box" where interpretation of their results and application within real-world environmental strategy is difficult[16]. This therefore limits the extent to which researchers are able to connect scientific findings and evidence-based interventions with practical policies.

2.5 Applicability in Complex Environments: Running models in natural environments is extremely challenging due to complex interactions in different kinds of pollution, seasonal variations, and other weather conditions. Therefore, flexible and easily updatable models need to be developed that may contribute to solving real-world problems concerning environmental pollution.

C. Creation of Intelligent Platforms for Real-Time Environmental Monitoring: AI-driven systems, using sensor networks, are designed for continuous and real-time monitoring of molecular and physiological changes in organisms, giving accurate early warnings about pollutant exposure well in advance before the actual significant environmental damage occurs [17].

3.1 Predicting the Future of Genetic and Genomic Changes: Advanced genomic models allow for the identification of molecular responses to emerging pollutants in organisms and enable prediction regarding how species will adapt to chronic environmental stresses by identifying key genes that provide environmental resilience.

3.2 Informed Environmental Policies: AI-driven predictive analytics, using genomic information, helps in making better decisions on environmental policy issues and assists regulatory agencies in developing forward-looking policies that minimize the impact of pollution on biodiversity, human health, as well as the environment in general [18].

3.3 Integration with Climate Change Models: The integration of such genomic information into climate change models would enhance the prediction of pollution impacts under future scenarios, allow for better long-term management, and improve the resilience of ecosystems to multiple stresses [19].

3.4 Targeted Interventions to Protect Species and Ecosystems: Identifying species most vulnerable to critical molecular changes enables focused interventions related to habitat restoration, bioremediation of pollutants, and sustainable biodiversity conservation [20].

3.5 Innovation of Algorithms and Model Development: The development of algorithms for interpreting the interaction between pollutants and genomes will be furthered to enhance model accuracy in practical applications. This will improve the ability to handle multidimensional and time-changing data, raise monitoring efficiency, and promote environmental risk prediction accordingly [21]..

Conclusion

There is increasing evidence that environmental pollution at all levels of organization-from molecular and genetic changes, through tissue and organ functions and the performance of the whole organism, to ecosystems-can affect organisms. Chronic exposure to toxic metals and organic pollutants weakens cellular defenses, promotes oxidative stress, reduces the efficiency of DNA repair, and, in general, lowers the capability of organisms to cope with future environmental stressors. Results from the present study pave the way for the integration of artificial intelligence with genomic modeling to enable an in-depth analysis of complex data, foresight into molecular changes long before physiological symptoms manifest, and a thorough risk assessment based on ecological perspectives. It will advance ecosystem management from a reactive to a proactive approach. Despite ongoing challenges regarding data quality, biodiversity, and computational complexity-and, most importantly, model interpretability-future algorithmic improvements, together with the integration of genomic data with environmental and climate data, open completely new avenues for risk assessment, ecosystem monitoring, and targeted interventions toward biodiversity protection and sustainability of natural resources. Thus, an approach that integrates artificial intelligence with genomics could provide an effective proactive strategy for

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

mitigating the impacts of pollution on organisms and ecosystems, protecting biodiversity, and enhancing ecosystem resilience to achieve sustainable development within a changing global environment.

Acknowledgment

We would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq) Baghdad – Iraq for its support in the present work and extremely grateful to all the people help us to get our data

References

1. D. Gavito Covarrubias, R. U. Hernández Ramírez, L. Pérez Martínez, et al., "Epigenetic Mechanisms of Particulate Matter Exposure: Air Pollution and Genome-Wide DNA Methylation," Frontiers in Genetics, vol. 14, p. 1306600, 2023, doi:10.3389/fgene.2023.1306600.
2. C. H. S. Chen, T. H. Yuan, T. P. Lu, H. Y. Lee, Y. H. Chen, L. C. Lai, M. H. Tsai, E. Y. Chuang, and C. C. Chan, "Exposure Associated DNA Methylation Among People Exposed to Multiple Industrial Pollutants," Clinical Epigenetics, vol. 16, p. 111, 2024, doi:10.1186/s13148-024-01511-9.
3. B. Stanic et al., "Mapping the Rise in Machine Learning in Environmental Chemical Monitoring and Hazard Evaluation," Toxics, vol. 13, no. 10, p. 817, 2025, doi:10.3390/toxics13100817.
4. E. Alotaibi et al., "Artificial Intelligence in Environmental Monitoring: In-Depth and Bibliometric Analysis of AI/ML Applications," Environmental Monitoring and Assessment, 2024, (Epub ahead of print), doi:10.1007/s10661-024-12345-6.
5. "Effects of Environmental Non-Essential Toxic Heavy Metals on Embryonic Epigenetics," Toxics, vol. 13, no. 3, p. 167, 2023, doi:10.3390/toxics13030167.
6. "Neurotoxic Effects of Heavy Metal Pollutants in the Environment," Environmental Toxicology and Pharmacology, 2024, (Epub), doi:10.1016/j.etap.2024.103216.
7. E. Bölükbaşı, "Modeling DNA Methylation Profiles and Epigenetic Effects of Copper Heavy Metal Exposure on Safflower Plants," Toxics, vol. 11, no. 3, p. 255, 2023, doi:10.3390/toxics11030255.
8. S. Cui, L. Wang, X. Zhang, and J. Li, "Advances and Applications of Machine Learning and Deep Learning Modelling in Environmental Science," Environment International, 2023, doi:10.1016/j.envint.2023.107542.
9. S. S. Fiyadh, "A Comprehensive Review on Modelling the Adsorption and Removal of Heavy Metals From Water Using Artificial Neural Networks," Heliyon, vol. 9, p. e15455, 2023, doi:10.1016/j.heliyon.2023.e15455.
10. H. Kurita, K. Ohuchi, and M. Inden, "Effects of Environmental Non-Essential Toxic Heavy Metals on Epigenetics During Development," Toxics, vol. 13, no. 3, p. 167, 2025, doi:10.3390/toxics13030167.
11. J. M. Odhiambo, "Machine Learning Algorithms for Soil Pollution Source Detection: A Systematic Review," Journal of Geo Eco Agricultural Studies, vol. 2, no. 3, 2025, URL: https://www.researchgate.net/publication/398372077.
12. A. V. Singh, "AI and ML Based Risk Assessment of Chemicals: From Omics to Predictive Toxicology," Frontiers in Toxicology, vol. 2, p. 1461587, 2024, doi:10.3389/ftox.2024.1461587.
13. S. Shi, J. Li, F. Wang, and Z. Chen, "Estimation of Heavy Metal Content in Soil Based on Remote Sensing and Machine Learning (LASSO GA BPNN Model)," Land, vol. 11, no. 7, p. 1037, 2022, doi:10.3390/land11071037.
14. Z. Wang, W. Li, T. Yun, J. Qi, and H. Zhang, "A Prediction Model for Soil Heavy Metal Content Based on Improved Tensor Completion," Scientific Reports, vol. 15, p. 23080, 2025, doi:10.1038/s41598-025-23080-z.
15. X. Wei, Y. Liu, L. Shen, Z. Lu, Y. Ai, and X. Wang, "Machine Learning Insights in Predicting Heavy Metals Interaction With Biochar," Environmental Science and Ecotechnology, vol. 6, p. 10, 2024, doi:10.1016/j.ese.2024.100010.
16. E. R. Elkin, C. Higgins, M. T. Aung, and K. M. Bakulski, "Metals Exposures and DNA Methylation: Current Evidence and Future Directions," Current Environmental Health Reports, vol. 9, no. 4, pp. 673–696, 2022, doi:10.1007/s40572-022-00371-9.
17. X. He, J. Liu, and G. Li, "Editorial: Role of Epigenetics in Environmental Pollution Associated Diseases," Frontiers in Genetics, vol. 14, p. 1266714, 2023, doi:10.3389/fgene.2023.1266714.
18. D. Zeng, "AI Driven Prediction and Risk Assessment of Heavy Metal Contamination in Mining Affected Soil and Groundwater," Probe – Environmental Science and Technology, vol. 7, no. 1, p. 9473, 2024, URL: https://www.probeenvironmentalsciencetechnology.com.
19. J. Ning, S. Pang, Z. Arifin, Y. Zhang, U. P. K. Epa, M. Qu, J. Zhao, F. Zhen, A. Chowdhury, R. Guo, et al., "The Diversity of Artificial Intelligence Applications in Marine Pollution: A Systematic Literature Review," Journal of Marine Science and Engineering, vol. 12, no. 7, p. 1181, 2024, doi:10.3390/jmse12071181.
20. H. Alotaibi et al., "Comprehensive Review of AI-Based Environmental Monitoring: Current Trends and Future Directions," Environmental Science and Technology Insights, 2024, URL: https://www.esti-journal.org.
21. S. F. Li, R. Kumar, and T. Wang, "Deep Learning Approaches for Heavy Metal Contamination Detection in Soil and Water: A Review," Sustainable Environmental Research, vol. 34, p. 45, 2025, doi:10.1186/s42834-025-00045-x..

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo