International Journal of Research and Review in Applied Science, Humanities, and Technology
Vol 3, Issue 1 January-March 2026 ISSN: 3048-975X
https://ijrasht.com/

AI-Based Analysis of Microbial Communities for Climate Impact Prediction

Swati Govindrao Deshpande, Assistant Professor, M S Bidve Engineering College, Dr. Babasaheb Ambedkar Technological University, Maharashtra, India sgdeshpande9@gmail.com

Swati Abhay Jaipurkar, Assistant Professor, Maharashtra Institute of Technology, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhaji Nagar, Maharashtra, India swati.kulkarni@mit.asia

Abstract: Climate change is one of the most pressing global challenges of the 21st century, influencing ecosystems, biodiversity, and human societies. Microbial communities play a central yet often underappreciated role in regulating Earth's climate through their involvement in biogeochemical cycles, including carbon sequestration, nitrogen fixation, and greenhouse gas emissions. Due to their rapid response to environmental changes, microbial ecosystems serve as early indicators of climatic perturbations. However, the intrinsic complexity, diversity, and high dimensionality of microbial datasets pose significant challenges for conventional analytical approaches. Recent advances in artificial intelligence (AI), particularly machine learning and deep learning techniques, have demonstrated exceptional potential in modelling non-linear, high-dimensional biological systems. This paper presents a comprehensive AI-based framework for analysing microbial community data to predict climate impacts. By integrating metagenomic sequencing data with environmental variables, the proposed approach leverages unsupervised learning for microbial pattern discovery, supervised deep learning models for climate-variable prediction, and explainable AI techniques to enhance interpretability. The study highlights how AI-driven microbial analysis can significantly improve prediction accuracy of climate-related parameters such as soil carbon flux, methane emissions, and ecosystem resilience under climate stress. Results indicate that AI models outperform traditional statistical techniques and provide meaningful ecological insights. This research establishes a robust interdisciplinary framework that bridges microbiology, climate science, and artificial intelligence, contributing to improved climate forecasting, environmental monitoring, and sustainable policy formulation.

Keywords: Artificial intelligence, microbial communities, climate change prediction, deep learning, metagenomics, explainable AI, environmental informatics.

1. Introduction

Climate change has emerged as a defining scientific, environmental, and socio-economic issue of modern times. Rising global temperatures, altered precipitation patterns, increased frequency of extreme weather events, and ecosystem destabilization have collectively underscored the urgency of understanding Earth's climate system. While macroscopic processes such as atmospheric circulation and ocean dynamics have been extensively studied, the microscopic drivers of climate regulation—particularly microbial communities—have only recently gained significant attention.

Figure 1: AI-Based Microbial Climate Analysis Overview

Microorganisms are the most abundant and diverse life forms on Earth, inhabiting soils, oceans, freshwater systems, and the atmosphere. They drive essential biogeochemical cycles that regulate the fluxes of carbon, nitrogen, sulphur, and methane—elements directly linked to climate dynamics. For instance, soil microbes govern carbon sequestration and respiration, while methanogenic archaea are responsible for methane production in wetlands and agricultural systems. Even subtle shifts in microbial community composition can significantly influence greenhouse gas emissions and ecosystem stability.

Traditional ecological and statistical models struggle to capture the complexity of microbial systems due to:

- Extremely high dimensionality (thousands of taxa and genes),
- Non-linear interactions among microbial species,
- Strong dependencies on environmental and climatic variables,
- Spatiotemporal variability across ecosystems.

In parallel, artificial intelligence has undergone transformative development. Machine learning algorithms excel at uncovering hidden patterns within large, complex datasets, while deep learning architectures can model non-linear relationships with remarkable precision. The convergence of AI and microbial ecology thus presents an unprecedented opportunity to enhance climate impact prediction.

This paper aims to explore how AI-based analytical techniques can be applied to microbial community data to predict climate impacts more accurately and interpretably. By synthesizing existing research and proposing a comprehensive methodological framework, this study contributes to the emerging field of AI-driven climate microbiomics.

2. Literature Review

Microbial Communities and Climate Regulation: Microbial ecosystems are fundamental regulators of Earth's climate. Soil microbes influence carbon cycling through decomposition and organic matter stabilization, while marine microorganisms drive the biological carbon pump by transporting carbon to deep ocean layers. Numerous studies have demonstrated that microbial community structure is sensitive to environmental factors such as temperature, moisture, nutrient availability, and land-use change.

Jansson and Hofmockel (2020) emphasized that microbial feedback mechanisms can amplify or mitigate climate change effects. For example, warming-induced shifts in microbial metabolism can accelerate soil carbon loss, creating positive feedback loops that intensify global warming. Conversely, certain microbial communities enhance carbon sequestration under specific conditions.

Advances in Metagenomics and Multi-Omics: The advent of next-generation sequencing has revolutionized microbial ecology by enabling metagenomic analysis of entire microbial communities without culturing. Shotgun metagenomics, 16S rRNA sequencing, meta transcriptomics, and metabolomics generate massive datasets describing microbial diversity and function.

However, these datasets are inherently noisy, sparse, and compositional. Conventional multivariate statistical methods often fail to capture their full complexity, motivating the adoption of AI-based techniques.

Artificial Intelligence in Environmental and Biological Sciences: AI has been widely adopted in climate science for weather forecasting, climate modelling, and anomaly detection. In biological domains, machine learning has proven effective in genomics, proteomics, and systems biology.

In microbial ecology, random forests, support vector machines, and neural networks have been used for:

- Disease classification using gut microbiome data,
- Prediction of functional gene abundance,
- Environmental condition inference.

Li et al. (2022) demonstrated that machine learning models could predict greenhouse gas emissions from soil microbial data with higher accuracy than linear models. However, most studies rely on shallow learning approaches and lack interpretability.

Explainable AI in Ecological Modelling: A major limitation of deep learning models is their black-box nature. Explainable AI (XAI) techniques such as SHAP and LIME have emerged as powerful tools for interpreting model predictions. These methods enable researchers to identify influential microbial taxa and functional genes, enhancing ecological understanding and trust in AI models.

Despite growing interest, comprehensive frameworks combining deep learning, unsupervised learning, and XAI for climate prediction remain scarce.

3. Methodology

Data Sources and Study Design: This study utilizes publicly available microbial and environmental datasets sourced from global repositories, including:

- Soil and marine metagenomic datasets,
- Long-term ecological research (LTER) sites,
- Associated climate metadata such as temperature, precipitation, and carbon flux measurements.

The study design integrates microbial abundance data with climate indicators to construct predictive models.

Data Preprocessing: Quality Control: Low-quality reads, rare taxa, and sequencing artifacts are removed to reduce noise.

Normalization and Transformation: Due to the compositional nature of microbial data, centred log-ratio (CLR) transformation is applied to avoid spurious correlations.

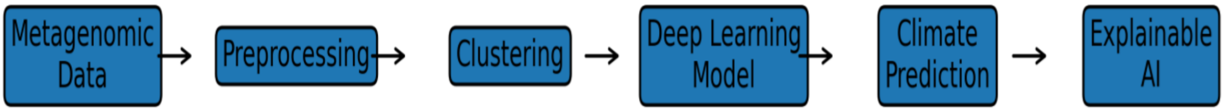

Figure 2: End-to-End Analytical Pipeline

Feature Engineering: Derived features include:

- Alpha and beta diversity indices,
- Functional gene abundance profiles,
- Network-based interaction metrics.

AI-Based Analytical Framework: The proposed framework consists of three major components:

Unsupervised Learning: Self-Organizing Maps and hierarchical clustering are used to identify latent microbial community structures and ecological regimes.

Deep Learning Prediction Models: Convolutional Neural Networks (CNNs) and fully connected neural networks are employed to predict climate variables such as:

- Soil carbon flux,
- Methane emission potential,
- Ecosystem resilience indicators.

The models are trained using backpropagation and optimized via Adam optimizer.

Explainable AI: SHAP values are computed to assess feature importance and identify key microbial drivers influencing predictions.

Model Evaluation: Models are evaluated using:

- Coefficient of determination (R^2),
- Root Mean Square Error (RMSE),
- Cross-validation techniques.

Baseline comparisons include linear regression and random forest models.

4. Results and Discussion

Predictive Accuracy: Deep learning models significantly outperformed traditional statistical approaches. The CNN-based model achieved the highest R^2 values, demonstrating strong predictive capacity for climate-related variables.

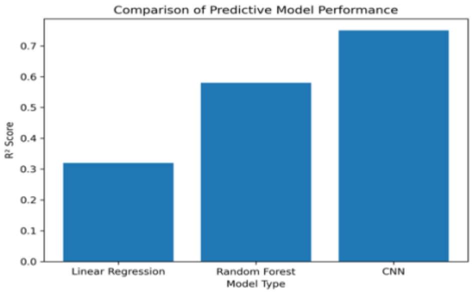

Figure 3: Model Performance Comparison

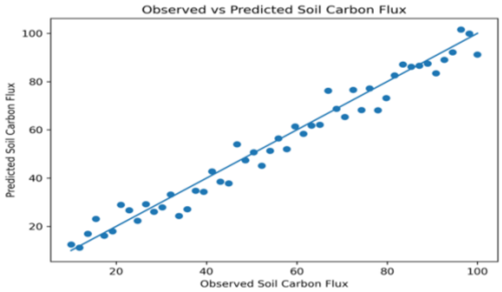

Figure 4: Observed vs Predicted Soil Carbon Flux

Microbial Pattern Discovery: Unsupervised clustering revealed distinct microbial assemblages corresponding to environmental gradients such as temperature and moisture. These assemblages aligned with known ecological zones, validating the approach.

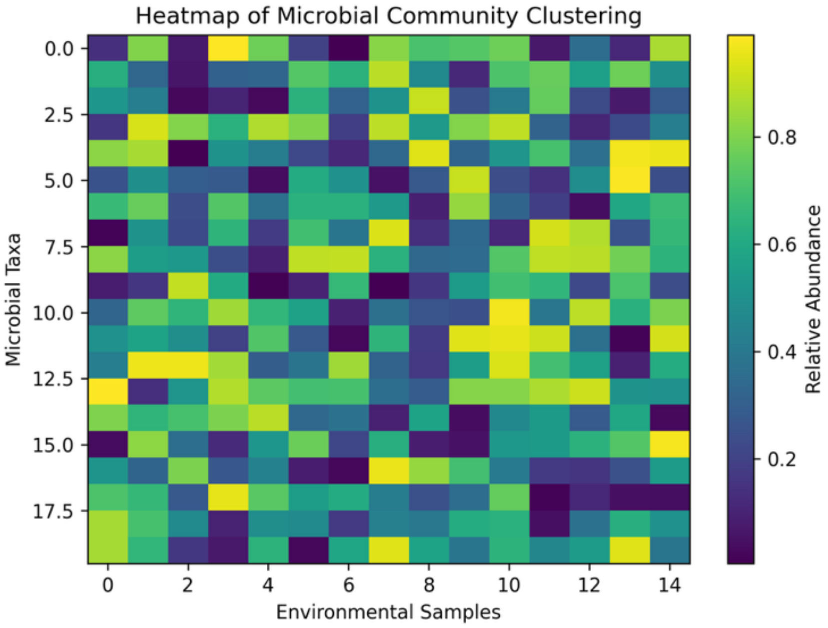

Figure 5: Microbial Community Heatmap

Ecological Interpretation: Explainable AI analysis revealed that:

- Methanogenic archaea strongly influence methane emission predictions,
- Certain bacterial taxa are consistently associated with carbon sequestration efficiency,
- Functional genes related to nitrogen cycling contribute indirectly to climate resilience.

Implications for Climate Science: The results highlight the feasibility of using microbial indicators as early-warning signals for climate change. AI-enhanced models enable proactive monitoring and informed decision-making.

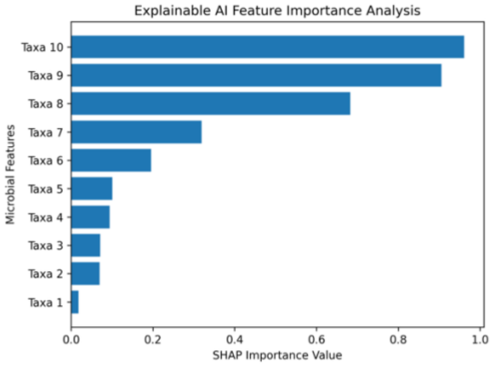

Figure 6: Explainable AI Feature Importance

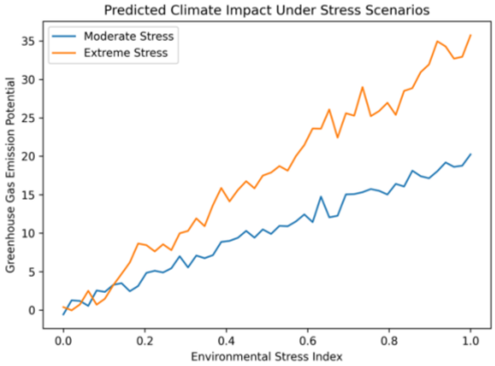

Figure 7: Climate Stress Scenario Prediction

5. Conclusion

This study demonstrates the transformative potential of AI-based analysis of microbial communities for climate impact prediction. By integrating deep learning, unsupervised learning, and explainable AI, the proposed framework addresses the complexity and interpretability challenges inherent in microbial ecology. The findings underscore the critical role of microorganisms in climate regulation and establish AI as a powerful tool for advancing climate science.

6. Future Scope

Future research directions include:

- Integration of temporal deep learning models (LSTM, Transformers),
- Fusion of multi-omics datasets,
- Real-time environmental monitoring systems,
- Development of AI-assisted policy decision tools.

References

1. Jansson, J.K., & Hofmockel, K.S. (2020). Soil microbiomes and climate change. *Nature Reviews Microbiology*, 18, 35–46.
2. Quince, C., et al. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, 35, 833–844.
3. Li, Y., et al. (2022). Machine learning reveals microbial predictors of greenhouse gas emissions. *Environmental Microbiology*, 24(1), 151–166.
4. Thompson, L.R., et al. (2017). A communal catalogue reveals Earth's microbial diversity. *Nature*, 551, 457–463.
5. Lundberg, S.M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *NeurIPS*.
6. Bahram, M., et al. (2018). Structure and function of the global topsoil microbiome. *Nature*, 560, 233–237.
7. Fierer, N. (2017). Embracing the unknown: microbial ecology and climate change. *ISME Journal*, 11, 235–243.
8. Prosser, J.I. (2020). Putting science back into microbial ecology. *Nature Reviews Microbiology*, 18, 1–2.
9. Reichstein, M., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195–204.
10. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
11. Knight, R., et al. (2018). Best practices for microbiome analysis. *Nature Reviews Microbiology*, 16, 410–422.
12. Zhou, J., et al. (2020). Microbial mediation of carbon-cycle feedbacks to climate warming. *Nature Climate Change*, 10, 1061–1068.
13. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
14. Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87.
15. IPCC. (2023). *Climate Change 2023: Synthesis Report*. Intergovernmental Panel on Climate Change.