International Journal of Research in Engineering, Science and Management
Volume 9, Issue 2, February 2026
www.ijresm.com | E-ISSN: 2581-5792 | RESAIM Publishing | www.resaim.com

Machine Learning Based Prediction of Energy Generation in Plant–Microbial Fuel Cells Using Environmental and Soil Data

Praveen Kumar[1*], Rajendra Singh[1]
[1]*School of Engineering and Technology, Raffles University, Neemrana, India*

***Abstract*: Plant–Microbial Fuel Cells (P-MFCs) are emerging as sustainable, bio electrochemical systems that harness energy from the interaction between plant root exudates and microbial communities in soil. Predicting energy generation from P-MFCs is challenging due to the influence of multiple environmental factors (temperature, moisture, pH, soil nutrients, etc.). This study proposes a machine learning (ML) framework that predicts electrical output in P-MFCs using environmental and soil sensor data. We evaluated multiple ML models including Random Forest (RF), Support Vector Regression (SVR), and Gradient Boosting Machines (GBM) using data collected from field trials and laboratory experiments. Results show that ensemble methods (RF and GBM) achieve strong predictive accuracy ($R^2 > 0.92$) across varied environmental scenarios, outperforming SVR. Feature importance analysis reveals soil moisture and pH as key predictors of power output. The proposed model enables real time prediction and optimization of P-MFC energy generation, offering a path toward smarter bioenergy harvesting from agricultural ecosystems.**

***Keywords*: Plant–Microbial Fuel Cells (P-MFCs), Bio electrochemical Systems, Machine Learning, Energy Prediction, Soil Sensor Data, Environmental Factors, Random Forest, Support Vector Regression (SVR), Gradient Boosting Machines (GBM), Ensemble Learning, Soil Moisture, Soil pH, Feature Importance Analysis, Real-Time Monitoring, Bioenergy Optimization, Sustainable Energy, Agricultural Ecosystems.**

1. Introduction

Plant–Microbial Fuel Cells (P-MFCs) are a class of bio electrochemical systems that convert chemical energy stored in soil and plant metabolites into electrical energy. When placed in the rhizosphere, P-MFCs harness electrons released by root exudates and microbial metabolism, enabling sustainable micro-scale power generation for sensors, irrigation control systems, and remote monitoring devices in agricultural landscapes.

Despite their promise, P-MFCs display highly variable electrical performance due to changes in environmental conditions such as soil moisture, temperature, pH levels, and nutrient availability. Traditional empirical methods for modeling P-MFC output are limited in their capacity to generalize to diverse field conditions.

Recent advances in Machine Learning (ML) provide powerful tools for modeling complex nonlinear relationships in environmental systems. In this study, we apply supervised ML algorithms to predict electrical energy generation in P-MFCs using environmental and soil features recorded via sensor arrays. By integrating IoT data streams with predictive models, this approach supports real-time optimization of bioenergy harvesting from agroecosystems.

2. Literature Review

A. Plant–Microbial Fuel Cells

P-MFCs have attracted research interest due to their renewable energy potential. Several studies have shown that power output depends on:

- Soil moisture levels, which influence ion mobility and microbial activity.
- Temperature, which affects enzyme activity and microbial metabolism.
- Soil pH, which influences redox potential and electron transfer rates.
- Soil nutrient content, especially organic carbon and nitrogen levels.

However, research on predictive modeling of P-MFC output remains limited.

B. Machine Learning in Bioenergy Prediction

Machine Learning has been applied extensively to renewable energy prediction:

- Wind and solar power forecasting using RF and GBM models.
- Soil carbon modeling and soil health classification.
- Biofuel yield prediction using support vector machines.

Notably, ML has not been widely applied to P-MFCs, where the environmental influence is complex and nonlinear.

3. Methodology

A. Data Collection

Environmental and soil data were collected using a 12-month field trial involving rice paddies with installed P-MFC units.

*Corresponding author: praveen.kumar0906@gmail.com

Data were sampled at 1-hour intervals and include:

Table 1

Feature	Units	Sensor
Soil moisture	%	Capacitive soil moisture sensor
Soil temperature	°C	Thermistor
Soil pH	–	pH probe
Electrical conductivity (EC)	dS/m	EC sensor
Organic carbon	%	Laboratory assay
Nitrogen (N)	mg/kg	Soil nutrient probe
Phosphorus (P)	mg/kg	Soil nutrient probe
Potassium (K)	mg/kg	Soil nutrient probe
Ambient temperature	°C	Weather station
Relative humidity	%	Weather station
P-MFC voltage	V	DC multimeter
P-MFC current	A	Hall effect sensor

Output power was computed as:
Power (W) = Voltage (V) × Current (A)

B. Data Preprocessing

- Missing data imputed using k-nearest neighbours method.
- Features normalized using min–max scaling.
- Dataset split: 70% training, 30% testing.

C. Machine Learning Models

We evaluated the following supervised regression models:

1. Random Forest Regression (RF)
2. Gradient Boosting Machine (GBM)
3. Support Vector Regression (SVR)
4. Multilayer Perceptron (MLP) Neural Network

Models were evaluated using:

- R^2 (Coefficient of determination)
- RMSE (Root Mean Square Error)
- MAE (Mean Absolute Error)

4. Results and Evaluation

Table 2

Model	R^2 (Test)	RMSE	MAE
Random Forest	0.94	0.15	0.10
Gradient Boosting	0.92	0.18	0.12
SVR	0.88	0.22	0.16
MLP	0.90	0.21	0.15

A. Feature Importance

The top predictors identified by RF and GBM:

1. Soil moisture
2. Soil pH
3. Soil temperature
4. EC
5. Nitrogen

Soil moisture and pH were consistently the highest-ranked features, indicating their strong influence on P-MFC performance.

B. Model Comparison

- RF and GBM outperformed SVR and MLP.
- Ensemble learning provided robustness to noise and nonlinear interactions.
- Neural networks required larger datasets for stable generalization.

5. Discussion

The performance of ML models demonstrates the efficacy of predictive analytics in modeling P-MFC energy generation. High R^2 values suggest the models accurately capture the influence of environmental and soil factors.

Practical applications include:

- Real-time prediction for adaptive energy management.
- Sensor node autonomy by using predicted vs. actual energy supply.
- Decision support for irrigation and crop management based on soil conditions linked to energy output.

6. Conclusion

This study demonstrates that machine learning can successfully predict energy output from Plant–Microbial Fuel Cells using environmental and soil sensor data. Ensemble models (RF and GBM) delivered high predictive accuracy and identify key environmental drivers of power production. The integration of IoT data streams with predictive models provides a scalable framework to optimize bioelectric performance in agricultural systems.

Future work should explore:

- Deep learning models with temporal sequences (LSTM).
- Transfer learning across soil types and crop systems.

Integration with real-time control for automated resource management.

References

[1] B. E. Logan et al., "Microbial fuel cells: Methodology and technology," *Environ. Sci. Technol.*, vol. 40, no. 17, pp. 5181–5192, 2006.

[2] Z. He, H. Minteer, and J. Li, "Electricity generation from rhizodeposits in plant microbial fuel cells," *Environ. Sci. Technol.*, vol. 43, no. 9, pp. 3391–3397, 2009.

[3] K. Rabaey and W. Verstraete, "Microbial fuel cells: Novel biotechnology for energy generation," *Trends Biotechnol.*, vol. 23, no. 6, pp. 291–298, 2005.

[4] R. Alfred et al., "Towards paddy rice smart farming: A review on big data, machine learning, and rice production tasks," *IEEE Access*, vol. 9, pp. 50358–50380, 2021.

[5] M. K. Senapaty, A. Ray, and N. Padhy, "IoT-enabled soil nutrient analysis and crop recommendation model for precision agriculture," *Computers*, vol. 12, no. 3, Art. 61, 2023.

[6] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, nos. 1–2, pp. 5–32, 2001.

[7] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.jds-online+2

[8] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.