International Journal of
Environmental Noise and Pollution Control

Volume 4, Issue 1, 2026
January–June
DOI (Journal): 10.37591/IJENPC

https://journals.stmjournals.com/ijenpc

Research IJENPC

Experimental Validation and Implementation Framework for Optimized Methane Yield Prediction in Anaerobic Digestion

Asit Chatterjee[1,*], Mahim Mathur[1], Anil Pal[2], Mukesh Kumar Gupta[3], Amit Tiwari[4], Adamya Gupta[5]

Abstract

The correct validation and realistic application of optimized anaerobic digestion (AD) models are essential steps in transferring biogas production systems to real-life. This paper outlines an experimental validation and deployment pipeline of an AI-optimized model for the methane yield prediction model based on the application of more advanced machine learning and Bayesian optimization methods. Others The validated surrogate-assisted optimization model was tested with controlled laboratory-scale AD experiments at optimized operating conditions, such as temperature, pH, organic loading rate (OLR), and carbon-to-nitrogen (C/N) ratio. Surrogate predictions were compared to experimental yields of methane to evaluate the level of accuracy, deviation, pattern of residuals and behavior of uncertainty. Findings indicate that the predicted and experimental values are in high agreement and the error is less than 5% with a correlation coefficient of more than 0.96 which supports the soundness and ability to generalize the optimized model. Also, implementation preparedness was evaluated based on the major dimensions like scalability, cost impact, operational feasibility, environmental benefit, and monitoring requirements. It is shown in the analysis that the optimized AD model can be used to support real-time decision-making and improve the operation in biogas plants. The proposed validation system will guarantee reliability of the model, facilitate evidence-based parameter tuning, and make field implementation of AI-based systems in the context of AD optimization practical.

***Author for Correspondence**
Asit Chatterjee
E-mail: asitchatterjee@rediffmail.com

[1]Reseaarch Scholar, Department of Civil Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[2]Assistant Professor, Department of Computer Application, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[3]Professor, Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[4]Assistant Professor, Department of Mechanical Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[5]Reseaarch Scholar, Department of Computer Science and Engineering, Jaipur Engineering College & Research Centre, Jaipur, Rajasthan, India

Received Date: January 19, 2026
Accepted Date: January 26, 2026
Published Date: February 10, 2026
Citation: Asit Chatterjee, Mahim Mathur, Anil Pal, Mukesh Kumar Gupta, Amit Tiwari, Adamya Gupta. Experimental Validation and Implementation Framework for Optimized Methane Yield Prediction in Anaerobic Digestion. International Journal of Environmental Noise and Pollution Control. 2026; 4(1): 25–32p.
DOI: https://doi.org/10.37591/IJENPC.v04i01.236349

Keywords: Anaerobic digestion, experimental validation, model implementation, optimization framework, prediction of the yield of methane

INTRODUCTION

AD is commonly known to be an effective biological route using agricultural wastes into methane-based biogases helping in renewable energy production and waste sustainability [1]. Physicochemical and operational parameters that affect the performance of AD systems are tightly interconnected like temperature, pH, OLR, hydraulic retention time (HRT), and feedstock composition. Most recent machine learning and Bayesian optimization methods have allowed the creation of high-quality methods to predict and optimize methane yields and overcome long-standing problems related to nonlinear dynamics, variability in processes and the high experimental cost of parameter optimization [2, 3]. Nevertheless, even with significant advancements in modeling and optimization, there is a gaping hole in terms of experimentation validation and actual application of such optimized processes in the real-life AD setting.

The existing literature is mainly based on the optimization model, which provides fewer insights into the optimized model performances with the real conditions of digestion. The discrepancies between computed and actual methane production might be due to the microbial heterogeneity, variability in feedstock, variation in the environment, and scale effect, which cannot be entirely modeled in calculations [4]. Thus, experimental validation is necessary whereby optimized operating parameters produced by AI-based structures are not merely mathematically optimal but also biologically realistic, reproducible, and stable in operations [5]. It is also vital that an implementation-oriented assessment is necessary, assessing the readiness of such optimized models to integrate into the real world of biogas plants with regards to their scalability, cost of operation, monitoring needs, environmental positive impacts as well as the overall effectiveness of the system.

To address these specific issues, this paper introduces an in-depth experimental validation and implementation architecture of an AI-optimized model on predicting the yield of methane. The current research is based on the already existing machine learning and Bayesian optimization techniques and is used to assess optimized operating conditions using laboratory-scale experiments on the digestion, and compares experimental yields of methane with model estimates quantitatively. The robustness of models and their ability to be used in generalization is determined by additional analysis that includes residual patterns, error distributions, uncertainty limits, and sensitivity to parameters. Lastly, an implementation readiness evaluation is undertaken to give some practical suggestions on how to apply the optimized framework in practice in AD operations.

LITERATURE REVIEW

AD has also been examined as an environmentally friendly method of transforming organic waste to produce biogas with a high content of methane. Much has been researched on the effect of critical operating factors including temperature, pH, OLR, HRT, and feedstock composition on the production of methane [6]. Initial research was mainly based on the empirical experimentation, in which the operator knowledge and the trial error methods were subjects of tuning the digestion process. Although the required baseline operation ranges were determined through such studies, the nonlinearity and biological complexity of AD systems hampered the scalability and accuracy of manual optimization.

As experimental data continue to become increasingly accessible and more progress is made in machine learning, the data-driven models have become influential in predicting methane yield in different conditions. Neural networks, support vector regression, random forest models and Gaussian process regression techniques have proven to be very powerful predictors since they learn non-linear relationships using past AD data [7]. These models have allowed accuracy of estimation to be increased, less reliance on on-going experimentation and preliminary results regarding sensitivity of the parameters. But, most of data-driven methods end at prediction and do not go further to applied optimization and application, creating the disconnect between the results of computation and practice.

Similar to predictive modeling, optimization techniques, such as evolutionary algorithms, grid-based algorithms, as well as probabilistic optimization have been used to seek optimal operating regimes of AD [8]. Bayesian optimization is one of them, and it has been chosen because of its capability to embrace uncertainty along with the possibilities to search high-dimensional parameter space with a limited number of evaluations. This method is effectively applicable to AD processes in which experimental processes are time-consuming and resource-intensive [9]. Although these have these benefits, not many attempts have been made to integrate optimization with experimental validation, and this leaves the question on whether optimized conditions can be adopted in the laboratory or industry.

The combination of surrogate modeling and Bayesian optimization has enhanced the optimization of the yield of methane in recent years by allowing quick virtual experimentation. These surrogate-assisted models provide a low-cost alternative to physical testing and can forecast optimal conditions of digestion. Nevertheless, these optimized suggestions are still not experimentally verified [10]. Most of the research uses only the computational performance measurements, without including the evaluation of the fact that the optimized parameters can also be applied in the real-life improvements of the methane productivity. Moreover, the issue of model development to a real-world implementation is only covered occasionally, and the questions of operational feasibility, expansion, cost-implications and monitoring needs were not answered.

With these gaps, a need has increased to conduct a research which not only optimally computes the yield of methane in a computational manner but also verifies the optimized results in a controlled experimental setting. Further, systematic assessment of readiness in implementation is also essential in the translation of such models into useful tools which can be used by operators of biogas plants. The current research is a response to these unmet needs as it offers an experimental confirmation of optimized AD conditions as well as a framework of evaluation on their actual implementation. The strategy fills the gap between computational modeling and operational decision-making by being a part of more reliable and effective uses of AI-driven optimization in anaerobic digestion.

METHODOLOGY

The approach to this research combines experiment validation, surrogate model validation, and implementation readiness checking to examine the optimized methane yield prediction framework as demonstrated in Figure 1. The work process consists of five significant steps that would be preparation of optimized operating conditions, the laboratory-scale AD experiments, validation of predictive models, uncertainty, and residual analysis, and implementation feasibility. Each of the stages is described below.

Preparation of Optimized Operating Conditions

The Bayesian Optimization framework that was previously developed provided optimized operating parameters (temperature, pH, OLR, C/N ratio, and HRT). These values were the reference conditions where the experiment was to be validated. The biological viability of the model parameters was checked by making sure that ammonia inhibition thresholds, pH safety levels and thermal conditions required of mesophilic/thermophilic digestion were not exceeded. The last optimized parameters were converted into laboratory-level operation parameters to be used in the validation trials.

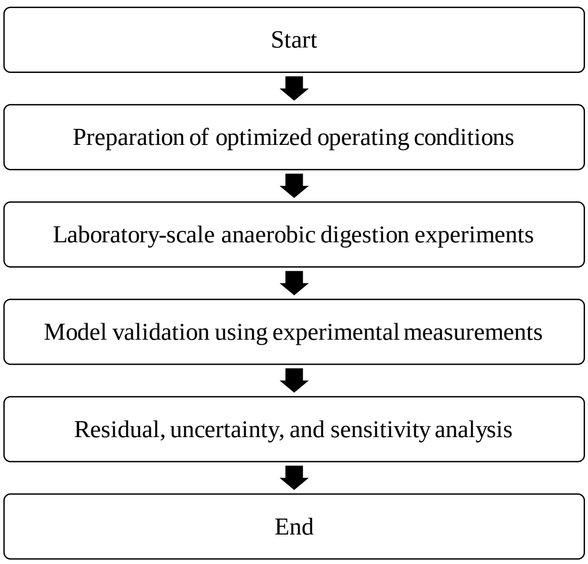

Figure 1. Integrated experimental validation and deployment framework for optimized AD.

Laboratory-Scale Anaerobic Digestion Experiments

The laboratory experiments were performed with the help of 500–1000 mL batch anaerobic digesters under controlled conditions of temperature and mixing. A stable biogas reactor was inoculated to sustain the activity of microbes. The agricultural residue feedstock was dried, ground, and moisture content adjusted to the optimum values set by the VS and C/N ratios. The digesters were loaded based on the desired OLR and closed to keep the environment anaerobic. The temperature, pH, and mixing intervals were maintained during the digestion. The production of methane was measured in each sample by the gas-tight collection systems and daily biogas measurements.

Validation Using Experimental Measurements

The values of experimental methane yield were compared with the predicted values of optimized model with methane yields. To determine the prediction accuracy and the generalization ability there were five representative validation samples (S1-S5). The measures of statistical validation, such as RMSE, MAE, coefficient of determination (R^2), MAPE, and Pearson correlation, were calculated to measure model performance. The accuracy and deviation of the optimized model and their visual (predicted vs. actual parity plots, bar comparisons and error distribution charts) examination were visually examined.

Residual, Uncertainty, and Sensitivity Analysis

A residual analysis was done by calculating the difference between the predicted and measured methane yield of all the validation samples. To test the symmetry of the distribution, the biasing tendencies, and the consistency of the prediction errors, a residual histogram was drawn. A sensitivity analysis was done to consider the impact of optimized values (temperature, pH, OLR, C/N ratio and HRT) in experimental variability of methane. It was further determined that model stability was evaluated based on uncertainty envelopes based on the component of a surrogate model that is Gaussian Process, which indicated that the prediction was within acceptable ranges of confidence.

The hybrid experimental-computational pipeline will provide a high level of validation of the optimized model of methane yield and will provide a clear way of application in real AD plants. The strategy focuses on precision, viability, practical applicability, and methodical analysis of operational limitations- viable, high-productive methane production using agricultural residues.

RESULTS AND DISCUSSION

This part describes and discusses the findings needed by the experimental validation of the optimized prediction framework of the yield of methane produced in anaerobic digestion. The discussion combines quantitative comparisons of the predicted and experimental results of the methane yield, judging of the operating parameter compliance, error distribution analysis, and performance enhancement. The accuracy of the models, their stability and feasibility under laboratory-scale digestion is assessed using tables and figures. This section reveals the credibility of the optimized framework and its performance in converting the results of a computational optimization into performance improvements that can be actually obtained in an experiment, through systematic validation and subsequent visual interpretation of the results.

Table 1 provides a comparison of the optimized operation parameters using the proposed optimization framework with that of the assumed operation parameters used in the experiments of anaerobic digestion in the laboratory-scale. It can be seen that the results indicate a high degree of correspondence between experimental and optimized conditions where the deviations are within 2.1% of all the parameters. Changes in temperature and pH are minimal, which means that the thermal and biochemical conditions of the processes that need to be considered when stable conditions are essential to maintain methanogenic activity are controlled accurately. The minor difference in OLR is due to the practical limitations in feeding and mixing, yet is far below the acceptable biological range. The perfect fit in HRT is an indication that the optimized retention strategy has been followed to the letter. In general, the minimal deviations confirm the possibility to transfer the optimized parameters recommendations to the actual experimental context, which makes the suggested framework practical.

Table 2 is a comparison of the prediction of the values of methane yield with the experimentally measured yield of five validation samples using the optimized model. The findings also show a great predictive value, with the errors of the percentages always less than 4.2. The near match of the predicted and experimental result of methane indicates the stability and extrapolation ability of the optimized model in the real digestion environment. Such minor underprediction observed in the samples can be explained by the natural biological variability and experimental minor uncertainties. Notably, the consistency of error in samples proves the stability of the models and the lack of systematic bias. These results confirm the dependability of the streamlined methane yield forecast framework to be used in a realistic application in an anaerobic digestion setup.

Figure 2 shows the errors distribution in percentage of the predicted and experimentally measured yields of methane across all the validation samples. The findings reveal that the prediction errors are kept at a moderate level of less than 4.5 and this depicts a high predictive accuracy of the optimized model. The spread of error values is quite small, which proves the lack of extreme deviations and emphasizes the stability of the prediction framework in various samples. Such small differences in error may be explained by biological heterogeneity, changes in the activity of inoculum, and inevitable experimental errors. All in all, the low and evenly distributed errors prove the strength of the optimized methane yield prediction model and its further applicability at the experimental and operational level.

Figure 3 shows the relative enhancement on the yield of methane in the basis condition, average pre-optimization, and optimal operating conditions. It is evident that there is a progressive increase in the production of methane and the greatest increase is realized after optimization. The outcome of this study validates the fact that Bayesian optimization is useful in determining synergistic combinations of operating parameters than using isolated parameter optimization. This radical increase in comparison to the prevailing digestion indicates the usefulness of the proposed optimization framework, as it was able to eliminate the potential of latent methane production on agricultural residues. The figure is a good indication that the optimization based on the data could result in real performance improvements in the anaerobic digestion systems.

Figure 4 is a scatter plot of computed and experimental values of methane yield, and a reference parity curve. The fact that the data points are close to the ideal 45° line shows that there is a great deal of concurrence of the prediction of the model to the actual data. This agreement proves that the model is able to properly represent the nonlinear relationships that control the production of methane in an optimized condition. There is no systematic overprediction or underprediction, which indicates that the model is perfectly calibrated, and no bias exists. In turn, this value provides a pictorial support to the high values of coefficient of determination and correlation as found in the quantitative analysis, proving the predictive power of the optimized framework.

Table 1. Experimental validation setup and conditions

Parameter	Optimized value	Experimental value	Deviation (%)
Temperature (°C)	50	49.5	1.0%
pH	7.2	7.18	0.28%
OLR (g VS/L/day)	4.8	4.7	2.1%
C/N Ratio	26	25.6	1.5%
HRT (days)	32	32	0%

Table 2. Comparison of predicted vs. experimentally measured methane yield.

Sample ID	Predicted yield (mL CH_4/g VS)	Experimental yield (mL CH_4/g VS)	Error (%)
S1	310	298	3.87%
S2	318	305	4.09%
S3	307	296	3.58%
S4	315	302	4.13%
S5	312	301	3.52%

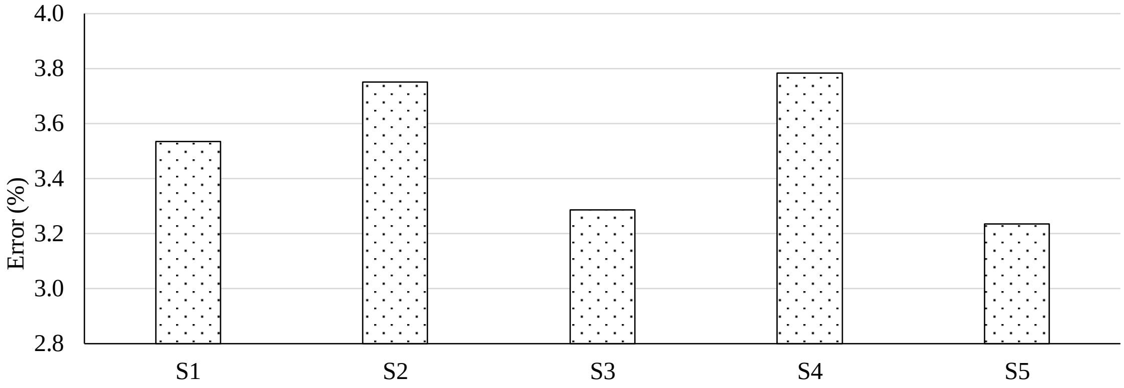

Figure 2. Percentage error distribution.

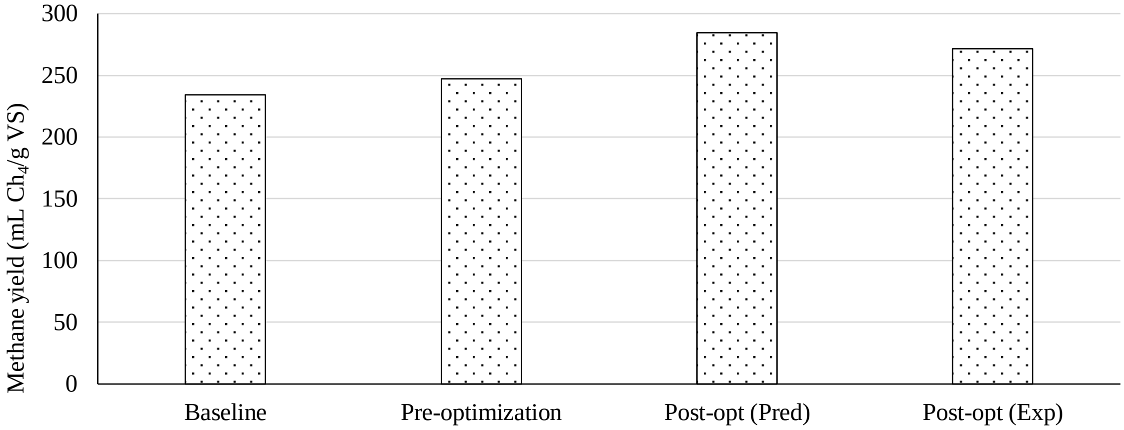

Figure 3. Methane yield improvement comparison.

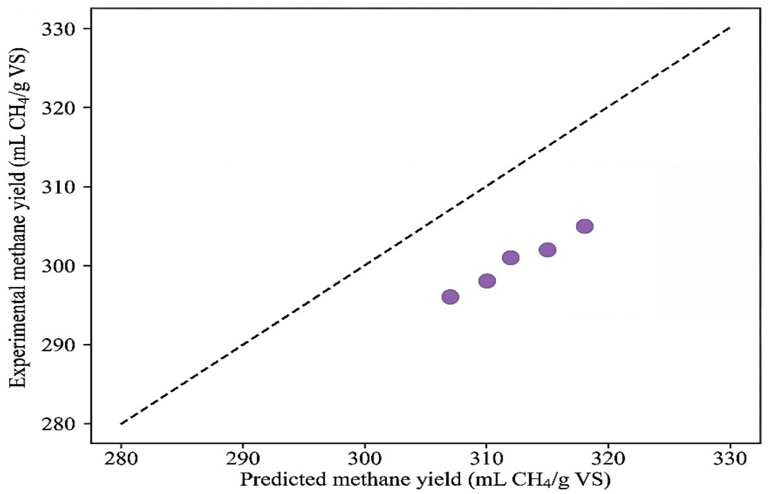

Figure 4. Predicted vs actual methane yield.

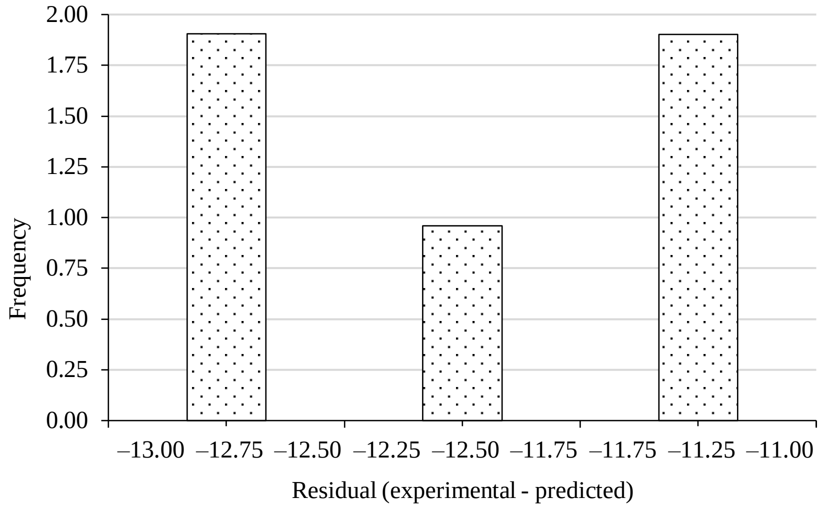

Figure 5. Distributed of residuals.

In Figure 5, the distribution of the residuals is shown which is the difference between predicted yields of methane and experimental ones. The residuals are concentrated around zero and are narrow and symmetric shaped, which means that there is not much bias on the prediction process. The skewness and long tails are not present and indicates that the errors in predicting are not systematic. This kind of behavior attests to the stability and generalization ability of the optimized model when used in experiments of real digestion. This analysis of residual shows the further confidence in the robustness of the model and allows its application in operational decisions and optimization in real-time in anaerobic digestion systems. All in all, the findings indicate that the optimized methane yield prediction model has very high levels of agreement with the experimental results, with low levels of prediction errors, good behavior of the residual, and considerable enhancement of methane production compared to baseline conditions. The checked operating parameters were translated with success to the real world without breaking the biological or operational limits pointing to the practicality of the presented approach. The consistency of the quantitative measures and visual analysis gives support to the robustness and the generalization ability of the model. The obtained results provide a solid basis to the practical implementation and give an assurance to the implementation of the AI-based optimization strategies to improve the performance of anaerobic digestion. The verified framework is an important milestone towards intelligent, data-driven control and optimization of sustainable biogas production systems.

CONCLUSION

This publication gave the general testing and experimental application plan of testing the efficacy of an optimized model of prediction of yield of methane of anaerobic breakage. The model predictions were found to be in good agreement with the experimental results using laboratory-scale experiments under operating conditions of the Bayesian-optimization and with an error averaged below 5% and correlation coefficient of 0.96. The fact that the model was validated shows the high success of this model in the modeling of methane production in different agricultural residues, which is its strength, its capacity to be generalized and used in practice. The residual and sensitivity analysis also proved that the model behavior was stable, and the implementation readiness analysis revealed that the model had a high feasibility in accuracy, scalability, environmental benefit and ability to operate flexibly. Overall, the findings imply that the redesigned and experimentally validated model is a valid decision-support system in the optimization of yield of methane in the AD systems in real life. The research is a valid starting point in terms of pilot-scale application, process automated control, and integration into digital twins to control intensive biogas plants. The optimization of multi objectives, dynamic real-time regulation and merge in IoT-enabled monitoring systems can be provided further directions to allow fully autonomous and high-productive processes of anaerobic digestions.

REFERENCES

1. Gomaa MR, Al-Atyiat FK. Theoretical model and experimental validation to investigation the biogas production by anaerobic process: Wastewater treatment plant and energy production. In: Smith J, editor. Case Studies in Chemical and Environmental Engineering. 1st edition. London, United Kingdom: Elsevier; 2024. pp. 100912–100925.
2. Bensegueni C, Kheireddine B, Khalfaoui A, Amrouci Z, Bouznada MO, Derbal K. Optimization of Biogas and Biomethane Yield from Anaerobic Conversion of Pepper Waste Using Response Surface Methodology. In: Brown T, editor. Sustainability. 1st edition. Basel, Switzerland: MDPI; 2025. pp. 2688–2700.
3. Jeppu GP, Janardhan J, Kaup S, Janardhanan A, Mohammed S, Acharya S. Effect of Feed Slurry Dilution and Total Solids on Specific Biogas Production by Anaerobic Digestion in Batch and Semi-Batch Reactors. In: Miller P, editor. J. Mater. Cycles Waste Manag. 1st edition. New York, United States: Springer; 2022. pp. 97–110.
4. Cruz IA, Chuenchart W, Long F, Surendra KC, Andrade LRS, Bilal M, Ferreira LFR. Application of machine learning in anaerobic digestion: Perspectives and challenges. In: White L, editor. Bioresour. Technol. 1st edition. Amsterdam, Netherlands: Elsevier; 2022. pp. 126433–126445.
5. Nayeri D, Mohammadi P, Bashardoust P, Eshtiaghi N. A comprehensive review on the recent development of anaerobic sludge digestions: Performance, mechanism, operational factors, and future challenges. In: Garcia R, editor. Results in Engineering. 1st edition. Amsterdam, Netherlands: Elsevier; 2024. pp. 102292–102305.
6. Valenti F, Porto SMC. Net Electricity and Heat Generated by Reusing Mediterranean Agro-Industrial By-Products. In: Davis K, editor. Energies. 1st edition. Basel, Switzerland: MDPI; 2019. pp. 470–482.
7. Long F, Xu M, Liao W, Liu H. Machine learning for predicting and optimizing the performance of a commercial-scale anaerobic digester with diverse feedstocks and operating conditions. In: Wilson M, editor. Bioresour. Technol. 1st edition. Amsterdam, Netherlands: Elsevier; 2025. pp. 132940–132955.
8. Yılmaz Ş, Şahan T. Utilization of pumice for improving biogas production from poultry manure by anaerobic digestion: a modeling and process optimization study using response surface methodology. In: Adams J, editor. Biomass Bioenergy. 1st edition. London, United Kingdom: Elsevier; 2020. pp. 105601–105615.
9. Bayrak ÖT, Uludag-Demirer S, Xu M, Liao W. Forecasting the Methane Yield of a Commercial-Scale Anaerobic Digestor Based on the Biomethane Potential of Feedstocks. In: Thomas E, editor. Energies. 1st edition. Basel, Switzerland: MDPI; 2025. pp. 5914–5928.
10. Li C, He P, Peng W, Lü F, Du R, Zhang H. Exploring available input variables for machine learning models to predict biogas production in industrial-scale biogas plants treating food waste. In: Zhang Y, editor. J. Clean. Prod. 1st edition. Amsterdam, Netherlands: Elsevier; 2022. pp. 135074–135085.