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 methane yield prediction 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 the 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]Research Scholar, Department of Civil Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[2]Professor, Department of Civil Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[3]Assistant Professor, Department of Computer Application, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[4] Professor, Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[5]Assistant Professor, Department of Mechanical Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[6]Research 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.

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

INTRODUCTION

Anaerobic digestion (AD) is commonly known to be an effective biological route for converting agricultural waste into methane-based biogases, which helps in renewable energy production and waste sustainability [1]. The physicochemical and operational parameters that affect the performance of AD systems are tightly interconnected, such as temperature, pH, OLR, hydraulic retention time (HRT), and feedstock composition. 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 non-linear 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 gap in experimental validation and the actual application of such optimized processes in real-life AD settings.

The existing literature is mainly based on optimization models, which provide fewer insights into the performance of optimized models under real digestion conditions. The discrepancies between the computed and actual methane production might be due to microbial heterogeneity, variability in feedstock, variation in the environment, and scale effects, 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 conducted to assess the readiness of such optimized models to be integrated into the real world of biogas plants with regard to their scalability, cost of operation, monitoring needs, positive environmental impacts, and overall effectiveness of the system.

To address these specific issues, this study introduces an in-depth experimental validation and implementation architecture of an AI-optimized model for predicting the yield of methane. The current research is based on existing machine learning and Bayesian optimization techniques and is used to assess optimized operating conditions using laboratory-scale experiments on digestion and quantitatively compare the experimental yields of methane with model estimates. The robustness of the models and their ability to be used for generalization were determined by additional analyses that included residual patterns, error distributions, uncertainty limits, and sensitivity to parameters. Finally, an implementation readiness evaluation is undertaken to provide practical suggestions on how to apply the optimized framework 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 methane content. The effects of critical operating factors, including temperature, pH, OLR, HRT, and feedstock composition, on methane production have been extensively studied [6]. Initial research was mainly based on empirical experimentation, in which operator knowledge and trial-and-error methods were used to tune the digestion process. Although the required baseline operation ranges were determined in these studies, the nonlinearity and biological complexity of AD systems hampered the scalability and accuracy of manual optimization.

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

Similar to predictive modeling, optimization techniques, such as evolutionary algorithms, grid-based algorithms, and probabilistic optimization, have been used to determine the 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 possibility of searching a 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 methods have these benefits, few attempts have been made to integrate optimization with experimental validation, leaving the question of whether optimized conditions can be adopted in the laboratory or industry.

The combination of surrogate modeling and Bayesian optimization has enhanced the optimization of methane yield in recent years by allowing quick virtual experimentation. These surrogate-assisted models provide a low-cost alternative to physical testing and can forecast the optimal conditions for digestion. Nevertheless, these optimized suggestions have not been experimentally verified [10]. Most research uses only computational performance measurements without evaluating the fact that the optimized parameters can also be applied in real-life improvements of methane productivity. Moreover, the issue of model development for 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, there is an increasing need to conduct research that not only optimally computes the yield of methane in a computational manner but also verifies the optimized results in a controlled experimental setting. Furthermore, a systematic assessment of readiness for implementation is essential for the translation of such models into useful tools that can be used by biogas plant operators. The current research is a response to these unmet needs, as it offers experimental confirmation of optimized AD conditions as well as a framework for evaluating their actual implementation. This 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 AD.

METHODOLOGY

The approach used in this study combined experimental 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: preparation of optimized operating conditions, laboratory-scale AD experiments, validation of predictive models, uncertainty and residual analysis, and implementation feasibility. The stages are described below.

Preparation of Optimized Operating Conditions

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

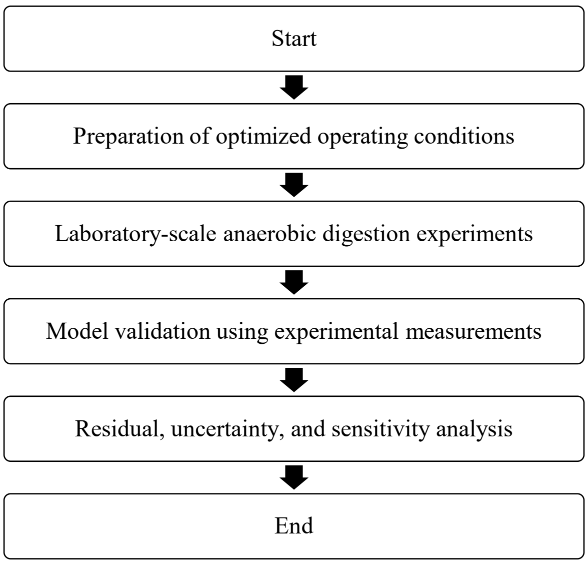

Figure 1. Integrated experimental validation and deployment framework for optimized AD.

Laboratory-Scale Anaerobic Digestion Experiments

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

Validation Using Experimental Measurements

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

Residual, Uncertainty, and Sensitivity Analysis

A residual analysis was performed by calculating the difference between the predicted and measured methane yields of all the validation samples. A residual histogram was drawn to test the symmetry of the distribution, biasing tendencies, and consistency of the prediction errors. A sensitivity analysis was performed to consider the impact of the optimized values (temperature, pH, OLR, C/N ratio, and HRT) on the experimental variability of methane. It was further determined that model stability was evaluated based on uncertainty envelopes from the component of a surrogate model that is a 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 a clear method of application in real AD plants. The strategy focuses on the precision, viability, practical applicability, and methodical analysis of operational limitations of viable, high-productive methane production using agricultural residues.

RESULTS AND DISCUSSION

This section describes and discusses the findings required for the experimental validation of the optimized prediction framework of the yield of methane produced in AD. The discussion combines quantitative comparisons of the predicted and experimental results of the methane yield, judging the operating parameter compliance, error distribution analysis, and performance enhancement. The accuracy, stability, and feasibility of the models under laboratory-scale digestion were assessed using tables and figures. This section reveals the credibility of the optimized framework and its performance in converting the results of computational optimization into performance improvements that can be 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 those of the assumed operation parameters used in the laboratory-scale AD experiments. 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 was due to the practical limitations in feeding and mixing, yet it was far below the acceptable biological range. A perfect fit in HRT indicates that the optimized retention strategy was followed to the letter. In general, the minimal deviations confirm the possibility of transferring the optimized parameters to the actual experimental context, which makes the suggested framework practical.

Table 2 shows a comparison of the predicted values of methane yield with the experimentally measured yield of five validation samples using the optimized model. The findings also show great predictive value, with the percentage errors of the percentages always being less than 4.2. The near match of the predicted and experimental results 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 natural biological variability and minor experimental uncertainties. Notably, the consistency of the errors in the 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 for realistic applications in an AD setup.

Figure 2 shows the error distribution in the percentage of the predicted and experimentally measured yields of methane across all validation samples. The findings reveal that the prediction errors are kept at a moderate level of less than 4.5, which depicts a high predictive accuracy of the optimized model. The spread of error values was quite small, which proved the lack of extreme deviations and emphasized 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 the inoculum, and inevitable experimental errors. Overall, 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 levels.

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

Figure 4 shows a scatter plot of the computed and experimental values of the 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 between the prediction of the model and the actual data. This agreement proves that the model can properly represent the non-linear relationships that control the production of methane under optimized conditions. There was no systematic overprediction or underprediction, indicating that the model was perfectly calibrated and no bias existed. This value provides pictorial support for the high values of the coefficient of determination and correlation 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 versus 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%

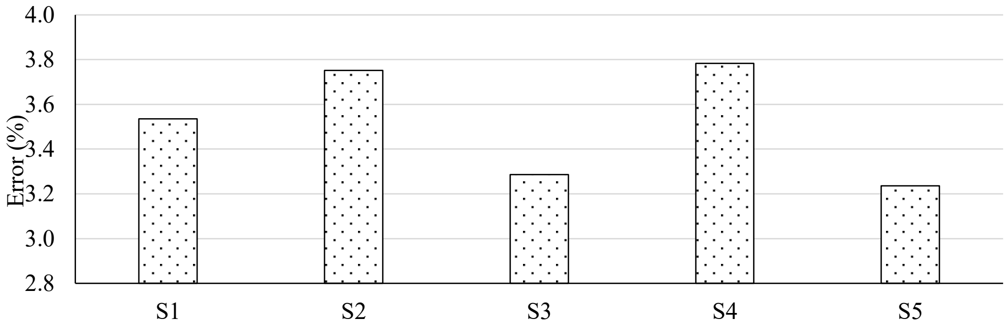

Figure 2. Percentage error distribution.

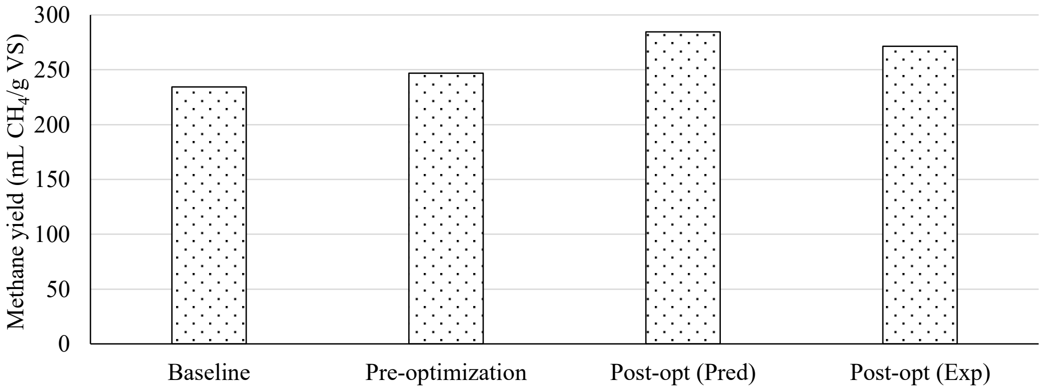

Figure 3. Methane yield improvement comparison.

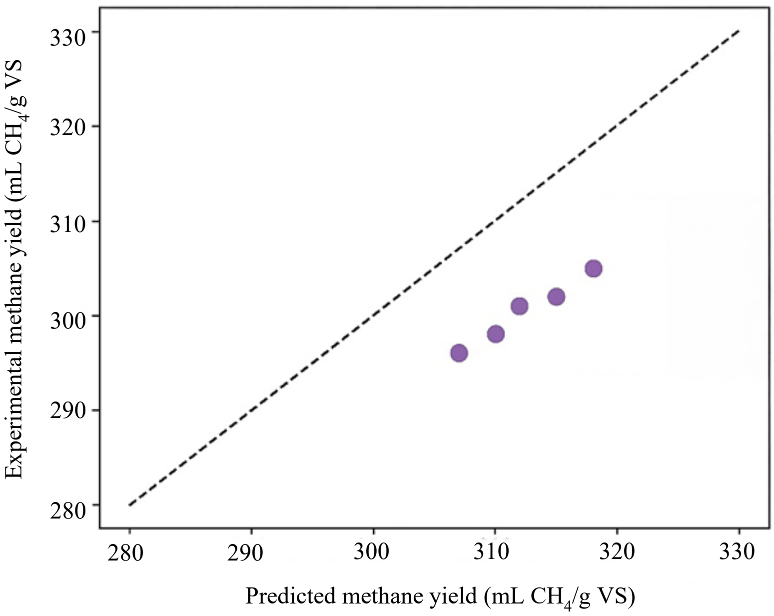

Figure 4. Predicted versus actual methane yield.

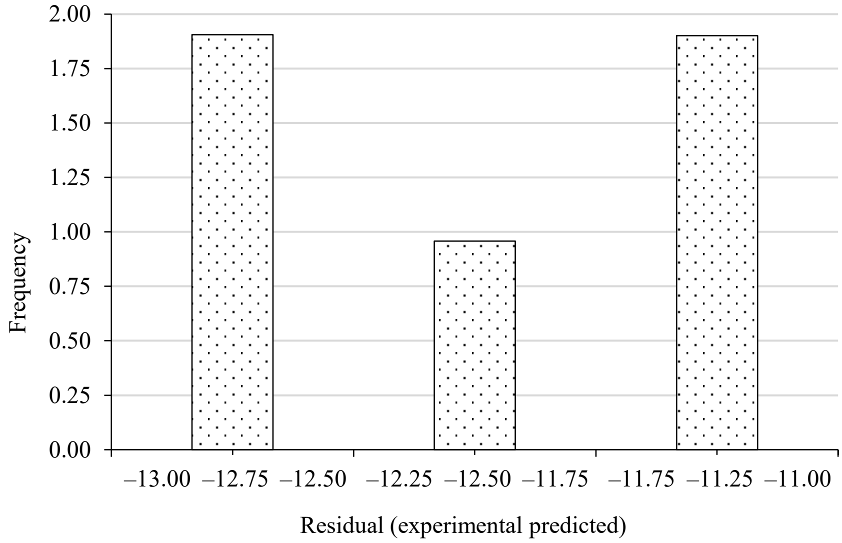

Figure 5. Distribution of residuals.

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 symmetrically in shape, which means that there is not much bias in the prediction process. The skewness and long tails are not present and indicate 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 residuals further supports confidence in the robustness of the model and allows its application in operational decisions and optimization in real-time in AD 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 for the practical implementation and give assurance for the implementation of the AI-based optimization strategies to improve the performance of AD. The verified framework is an important milestone towards intelligent, data-driven control and optimization of sustainable biogas production systems.

CONCLUSION

This publication provided a general testing and experimental application plan for testing the efficacy of an optimized model for predicting the yield of methane from AD. The model predictions were found to be in good agreement with the experimental results using laboratory-scale experiments under the operating conditions of the Bayesian optimization, with an error averaged below 5% and a 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 and capacity to be generalized and used in practice. The residual and sensitivity analyses also proved that the model's behavior was stable, and the implementation readiness analysis revealed that the model had high feasibility in terms of 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 for optimizing methane yield in AD systems in real life. This research is a valid starting point for pilot-scale applications, process automated control, and integration into digital twins to control intensive biogas plants. The optimization of multi-objectives, dynamic real-time regulation, and merging in IoT-enabled monitoring systems can provide further directions to allow fully autonomous and highly productive processes of AD.

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. Case Stud Chem Environ Eng. 2024;10:100912. doi:10.1016/j.cscee.2024.100912.
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. Sustainability. 2025;17:2688. doi:10.3390/su17062688.
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. J Mater Cycles Waste Manag. 2022;24:97-110. doi:10.1007/s10163-021-01298-1.
4. Andrade Cruz I, Chuenchart W, Long F, Surendra KC, Andrade LR, Bilal M, et al. Application of machine learning in anaerobic digestion: Perspectives and challenges. Bioresour Technol. 2022;345:126433. doi:10.1016/j.biortech.2021.126433. PMID: 34848330.
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. Results Eng. 2024;22:102292. doi:10.1016/j.rineng.2024.102292.
6. Valenti F, Porto SMC. Net electricity and heat generated by reusing Mediterranean agro-industrial by-products. Energies. 2019;12:470. doi:10.3390/en12030470.
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. Bioresour Technol. 2025;435:132940. doi:10.1016/j.biortech.2025.132940. PMID: 40618907.
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. Biomass Bioenergy. 2020;138:105601. doi:10.1016/j.biombioe.2020.105601.
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. Energies. 2025;18(22):5914. doi:10.3390/en18225914.
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. J Clean Prod. 2022;380:135074. doi:10.1016/j.jclepro.2022.135074.

