STM JOURNALS

International Journal of
Energy and Thermal Applications

ISSN: 3108-1479
Volume 4, Issue 1, 2026
January–June
DOI (Journal): 10.37591/IJETA

https://journals.stmjournals.com/ijeta

Research IJETA

Bayesian Optimization-Driven Operating Parameter Tuning for Maximizing Methane Yield in Anaerobic Digestion

Asit Chatterjee[1,*], Mahim Mathur[2], Anil Pal[3], Mukesh Kumar Gupta[4], Amit Tiwari[5], Adamya Gupta[6]

Abstract

To achieve maximum methane production in an anaerobic digestion (AD) process, a combination of various operational parameters must be tuned nonlinearly in the digestion ecosystem. The traditional trial-and-error optimization methods are slow, resource-consuming, and, in most instances, cannot model the intricate parameter interaction in biogas production. The current work introduces a Bayesian optimization-based model to optimize the set of conditions to maximize the level of methane produced using the agricultural residues, in terms of temperature, pH, organic loading rate (OLR), and carbon/nitrogen (C/N) ratio. The surrogate models were trained on a structured and labelled AD dataset and were used as fast evaluators in the optimization loop. Bayesian optimization using Gaussian process regression (GPR) with upper confidence bound (UCB) acquisition was used to search and exploit the parameter space of operational constraints. Findings indicate a high increase in the yield of methane, from a baseline figure of 260 mL CH_4/g VS to 315 mL CH_4/g VS, which shows an improvement of 21.2%. Stable convergence, a high level of surrogate accuracy, and uniform identification of biologically plausible parameter combinations are found in the optimization trajectories. This article reveals the potential of Bayesian optimization (BO) as an efficient instrument for AD process optimization, which can allow making operational decisions that are data-driven and minimizing the number of experiments in a waste-to-energy framework.

Keywords: Anaerobic digestion, Bayesian optimization, constrained parameter search, increase in methane yield, surrogate modeling

***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. Bayesian Optimization–Driven Operating Parameter Tuning for Maximizing Methane Yield in Anaerobic Digestion. International Journal of Energy and Thermal Applications. 2026; 4(1): 1–8p.

INTRODUCTION

Anaerobic digestion (AD) has become an essential biological process for transforming agricultural residues into renewable-rich biogas as a source of energy and providing renewable waste management. However, the productivity of methane production is very sensitive to operating conditions, temperature, pH [1, 2], organic loading rate (OLR), and the carbon/nitrogen (C/N) ratio. The interaction of these parameters is not a linear and predictable process and optimizing AD conditions is a difficult task. Traditional methods of empirical tuning use highly time-consuming and expensive laboratory tests [3], which are not as effective when scaling across the multi-dimensional operating space. This leads to a high percentage of AD systems not performing optimally, and most of the potential for using methane goes to waste.

New developments in machine learning and probabilistic modeling have created possibilities for speeding up AD optimization. Bayesian optimization (BO), especially, has become known as an effective way of exploring complex parameter spaces with a minimum of experimental effort [4]. Through surrogate modeling and uncertainty-sensitive search algorithms, BO efficiently balances exploration and exploitation to determine optimal operation parameters. This makes it particularly appropriate for biological systems, where experiments are slow, costly, and prone to variation in the processes, ensuring efficiency.

In a system with AD, BO is combined to find regimes of operation with a high yield on a systematic and data-driven basis. Surrogate models that can estimate responses to methane yield with high accuracy, including Gaussian process regression (GPR), deep neural networks (DNN), random forests (RF), and support vector regression (SVR), can be used as rapid evaluators in the optimization cycle [5–7]. Optimization can be limited to biologically possible and operationally safe areas through constrained search mechanisms, ensuring that practical implementation remains feasible.

This study introduces a full Bayesian optimization-based strategy to advance the yield of methane using agricultural residues. The proposed method determines ideal parameter combinations, which help increase the production of methane by a significant margin, in comparison to the baselines and pre-optimization situations, using a structured AD dataset and a search strategy that relies on surrogate assistance. The results emphasize how BO can be used to revolutionize the design of AD processes and present a scalable, smart, and experimentally efficient method for real-life biogas systems.

LITERATURE REVIEW

The optimization of AD processes has received a lot of attention in terms of experimental research and computer modeling. Initial experiments were mainly based on empirical modifications of temperature, pH, hydraulic retention time, and OLR to enhance methane generation [1]. Although these methods are useful for providing background information, they are limited by slow reaction cycles, small parameter spaces, and poor consistency in reproducibility across feedstock types [2]. Since AD systems are inherently nonlinearly coupled biochemical systems, the one factor at a time experiments traditionally used were not sufficient to measure the interdependence between operating conditions, leading to poor yield enhancement.

Machine learning methods have become important in the process of methane yield prediction and comprehending the process variables that are of importance. Regression models, neural networks, ensemble algorithms, and support vector-based methods have shown high predictive accuracy when trained on structured AD datasets [3]. These techniques provide a means of estimating intricate digestion behavior without requiring finer mechanical knowledge. However, most machine learning applications have tried to make only predictions and not optimizations, which has created a gap between model outputs and operating strategies that could be implemented [4]. In addition, prediction-oriented models generally require large datasets and do not provide a natural measure of uncertainty, which is vital in the case of biological systems.

Recently, surrogate-assisted and probabilistic optimization frameworks have emerged in this field. The use of surrogate models in methane response surfaces in the form of Gaussian processes, DNNs, RF, and kernel regressors has made it possible to conduct virtual experiments quickly by offering cheap approximations of the response surfaces [5]. These models facilitate the investigation of vast operating spaces that would otherwise be impossible to experiment on. Among these, Gaussian process-based methods can be distinguished by the fact that they directly introduce uncertainty estimates in the optimization process, thus allowing a balance between the exploration of unknown high-performance areas and the fine-tuning of known high-performance areas [6].

BO has also become a popular method for optimizing complex systems that are costly to evaluate, such as anaerobic digesters [7]. BO uses probabilistic surrogates and acquisition functions to identify promising operating conditions. Compared to traditional optimization methods, which require many evaluations, BO requires a moderate number of surrogate-directed cycles to obtain high-yield results [8]. This is especially appropriate for AD processes, in which an experimental run can take several weeks [9, 10]. Although there are these benefits, the use of BO to enhance the yield of methane is still scarce, and the existing research sometimes does not involve any constraints on search strategies to maintain biologically plausible recommendations.

METHODOLOGY

The methodology used in this study combines surrogate-assisted BO and constrained parameter search to maximize the yield of methane through AD, as shown in Figure 1. This workflow includes six primary steps: dataset preparation, feature preprocessing, surrogate model construction, BO configuration, constrained search operation, and ultimate validation of optimized operating conditions. All stages are outlined below.

Dataset Preparation and Features Selection

The input for optimization was a well-organized AD dataset that included physicochemical feedstock characteristics and working parameters. The important parameters were TS, VS, C/N ratio, lignocellulosic composition, temperature, pH, hydraulic retention time (HRT), and OLR. The target variable was the methane yield (mL CH_4/g VS). The data were cleaned by removing outliers using the interquartile range procedure (IQR) procedure, filling missing values with feature-wise Gaussian distributions, and normalizing all continuous variables with Minmax to ensure a smooth process of fitting the model.

Surrogate Model Construction

The syllabus focuses on the creation of surrogate models in the framework of international trade and finance. Four AI-ML surrogate models were created and assessed to be able to run the evaluation of methane yield in large parameter spaces efficiently:

- *Gaussian process regression* to facilitate uncertainty-aware optimization,
- *Deep neural network* to approximate nonlinear functions,
- *Random forest regressor (RFR)* for ensemble-based response estimation,
- *Support vector regression* used to predict with a margin.

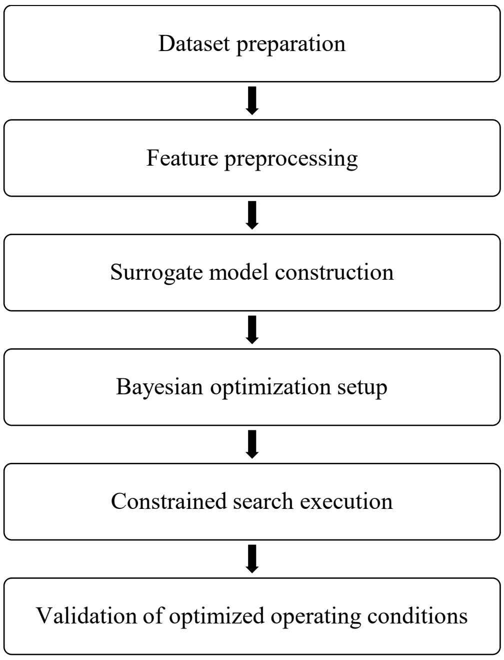

Figure 1. Framework for constrained BO of AD performance.

Each model was trained using 80% of the source data and tested using the remaining 20% of the total data using the RMSE, MAE, and R^2 statistical measures. The optimization of hyperparameters was performed using the standard Grid Search method, and the surrogate with the best performance was selected as the evaluation engine for BO.

Parameter Space Definition and Constraints

Ranges of operation were determined using pragmatic limits of AD and biological feasibility:

- *Temperature:* 30–55°C
- *pH:* 6.5–7.8
- *OLR:* 1–6 g VS/L/day
- *C/N ratio:* 20–35

Limitations such as constant pH, permissible ammonia concentration, and methane inhibition limits were imposed to guarantee that the optimization recommendations were also viable in the experiment.

Bayesian Optimization Framework

Bayesian optimization using a Gaussian process prior was utilized to effectively search the multi-dimensional search space. The upper confidence bound (UCB) acquisition function was employed to control exploration and exploitation.

The next parameter set identified by the acquisition function was the candidate parameter set. The surrogate model predicted the methane yield and uncertainty. The results were input into the GP model to refine the posterior distributions. The optimization process was repeated until convergence criteria were achieved, which was a reduction in the volatility of less than 0.5% in the yield of methane over five cycles.

Constrained Search Execution

The performance of BO was evaluated in the presence of strict process constraints to prevent the recommendation of an unstable or inhibitory condition. Candidates with candidate points that violated the constraints were eliminated, and the search was only carried out in feasible regions. The surrogate model was also able to deliver rapid virtual assessments (~milliseconds), which allowed thousands of simulated experiments to be conducted compared to weeks of experimental physical measures on digestion. Convergence was ensured by monitoring the optimization trajectories and scores of acquisitions and surrogate predictions in a biologically meaningful manner.

Validation of Optimized Operating Conditions

The best parameter combination found by BO was compared with the baseline and pre-optimization performances. The predicted yields of methane using the surrogate against the average of the dataset were compared to measure the improvement. Additionally:

Predicted vs. actual plots, error histograms, imprecision involving GPR, and 3D response surface validation were performed on the accuracy of optimization. An endpoint yield of 315 mL of CH_4/g VS, which constitutes a 21.2% growth, ensured the suitability of the BO strategy.

RESULTS AND DISCUSSION

The current section introduces and discusses the findings achieved during the Bayesian optimization-based tuning of the operating parameters of AD in an attempt to achieve the highest possible yield of methane. The analysis combines statistical summaries, convergence behavior, parameter evolution, and function dynamics of the acquisition and performance gains before and after the optimization. The outcomes were measured using quantitative metrics and visual representation to evaluate the efficiency of optimization, stability of the model, and biological viability. By systematically analyzing the optimization process and the subsequent increase in the yield of methane, this section proves that the proposed framework is efficient at determining the optimal operational regimes and ensuring a stable process and its applicability in practice.

Table 1 presents the limited search space of the BO of the parameters of AD. The lower and upper limits of temperature, pH, OLR, carbon-to-nitrogen (C/N) ratio, HRT, and mixing speed used (1. selected) were both biologically feasible and sufficiently flexible to be optimized. A temperature gradient of 30–55°C is between mesophilic and thermophilic digestion regimes, which allows the optimizer to experiment with high-activity zones without the danger of microbial inhibition. pH values in the range (6.2–7.8) indicate stable methanogenic conditions, which do not result in acidification and ammonia toxicity. Likewise, the range of 1.060 g VS/L/day g VS/L/day covers the conservative and high-throughput loading situations. Limiting factors in the C/N ratio ensure sufficient availability of nutrients, whereas the broad range of HRT ensures the investigation of both fast and slow digestion processes. The mixing speed range was chosen to compromise between improving mass transfer and preventing shear stress. In general, such a clear search space guarantees safe, efficient, and realistic optimization, which provides a strong basis for Bayesian-based parameter optimization.

The convergence behavior of the BO process as the series of iterations progressed is shown in Figure 2. This curve demonstrates how rapidly the methane yield grows in the initial iterations, indicating that the parameter space is properly searched and promising areas of operation are identified quickly. The improvement rate also decreased gradually as the number of iterations increased, indicating a shift in the process between exploration and exploitation. This trend of convergence proves the effectiveness of BO in finding near-optimal solutions with a low number of evaluations. The fact that the yield of the stabilization of methane was reached at later iterations indicates that the optimizer has found an optimal or near-optimal operating configuration. Convergence behavior is particularly useful in AD applications, where experimental assessments are expensive and time-intensive.

Figure 3 shows the evolution of the critical operating parameters during the optimization process, which provides information about the adjustments made by the BO to each variable in the maximization of the methane yield. Temperature and OLR showed noticeable positive trends, indicating that they have a strong impact on stimulating the increase of methane generation within the safety of operations. The pH curve was also fairly constant, approaching a favorable neutral pH that would support the growth of methanogens. The C/N ratio was intermediate, indicating that there was an equal amount of carbon and nitrogen constraints. HRT increases slowly and then levels off, which is the trade-off between complete digestion and process efficiency. The speed of mixing increases to enhance the contact between the substrate and microbe but becomes constant in the end to avoid shear. All these trajectories indicate that the optimization process does not violate any biological constraints and smartly adjusts the parameters to levels that are more optimal for methane.

The development of the expected improvement (EI) acquisition function during the BO process is shown in Figure 4. The declining pattern of the EI curve shows that the optimization framework is efficient in balancing exploration and exploitation in successive iterations. In the early stages, the values of EI are found to be greater than the values, resulting in more uncertainty in the search space and a greater focus on exploration to identify any good operating regions. The EI values decrease slowly as the optimization continues, indicating that the surrogate model becomes confident about the predicted landscape of the methane yield. This decrease proves that the algorithm is more oriented toward the optimization of the best areas instead of trying to explore untried combinations of parameters. This behavior shows effective convergence, whereby computational and experimental resources are used in the most efficient way. The gradual decay of the EI curve is also a further sign of stable surrogate learning and the lack of sporadic search behavior, which is pertinent to the sound optimization of AD processes.

Table 1. Optimization search space for AD operating parameters.

Parameter	Lower bound	Upper bound
Temperature (°C)	30	55
pH	6.2	7.8
OLR (gVS/L/day)	1.0	6.0
C/N ratio	18	32
HRT (days)	15	40
Mixing speed (rpm)	50	200

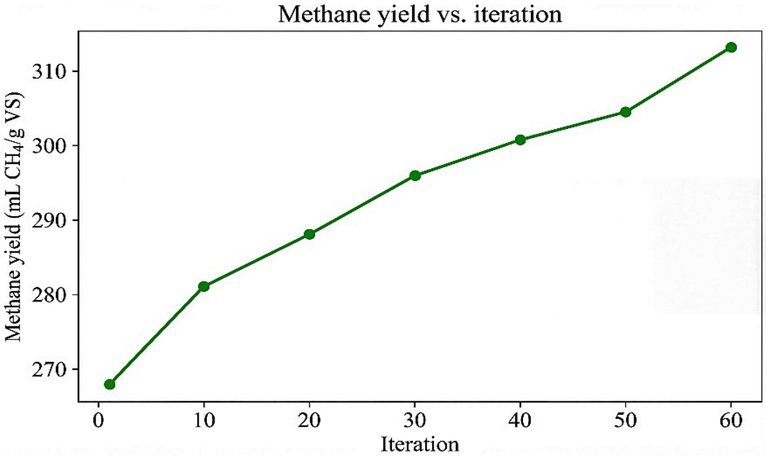

Figure 2. Bayesian optimization convergence curve.

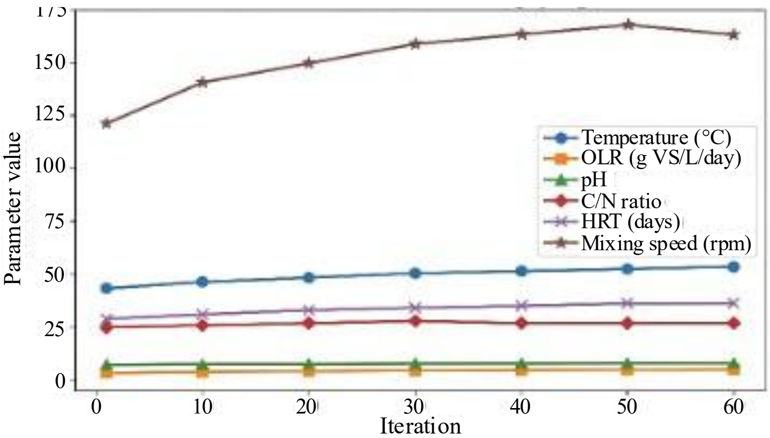

Figure 3. Parameter trajectories during optimization.

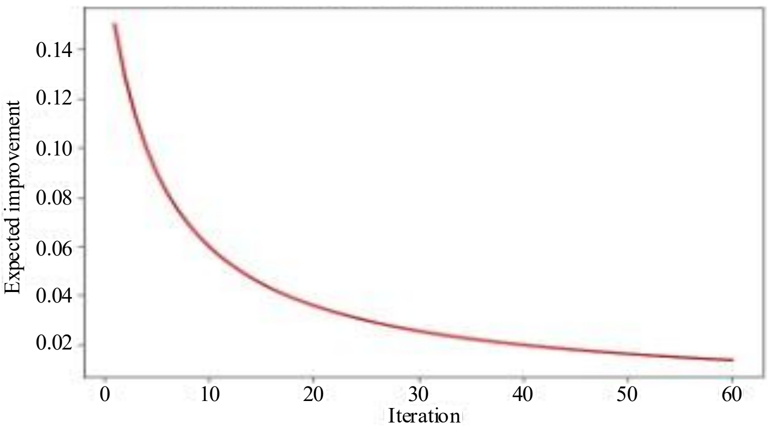

Figure 4. Acquisition function (EI) during optimization.

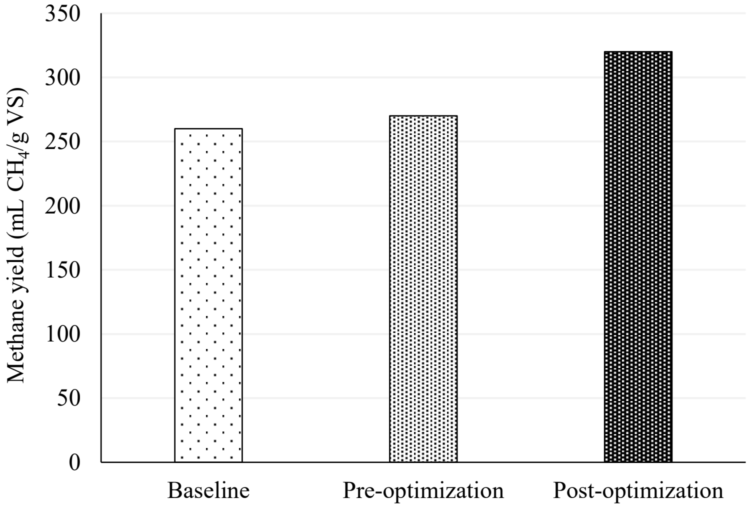

Figure 5. Methane yields before and after optimization.

A comparison of the methane yield under baseline, pre-optimization average, and post-optimization conditions derived using BO is presented in Figure 5. There was a significant increase in methane production after optimization. Although the AD apparatus under baseline conditions provides relatively lower levels of methane production, a progressive change is recorded when switching to dataset-average situations. It was optimized, and thereafter, the enhancement was the greatest, with a significant increase in the yield of methane, which demonstrates the usefulness of the optimization framework suggested. This enhancement proves that the coordination of operation parameter changes, as opposed to a singular change, is highly important for maximizing biogas productivity. Yield improvement was also observed, which confirms the predictive accuracy of the surrogate-assisted optimization model which proves to be practical. In general, Figure 5 presents compelling evidence that BO can generate practical performance objectives in an AD mechanism and, therefore, its use in real-world optimization of biogas plants. The results above corroborate that BO with surrogate modeling, and a constrained search space can be a potent and effective solution for optimizing methane generation in an AD system. The convergence behavior observed, stable parameter curves, and significant gains made by the proposed optimization framework all justify the strength of the proposed optimization framework. Notably, the optimized operational conditions were maintained within biologically safe and operationally practical limits, which supports the practical use of the methodology. The research results can be used to form a solid base for the experimental validation and practical application of AI-based optimization methods, contributing to the creation of intelligent and high-performance biogas generation systems and the further evolution of sustainable waste-to-energy conversion.

CONCLUSION

This study demonstrated that BO is an efficient and data-driven model for maximizing the yield of methane in an AD system. By combining structured AD data with surrogate modeling and uncertainty-sensitive Bayesian search, the proposed framework was able to identify the optimal operating conditions that significantly boosted methane production. The parameter set with the optimum boosted the yield of methane to 315 mL CH_4/g VS, which is a 21.2% higher yield than the baseline and achieves better success than the conventional manual techniques of tuning. GPR, as part of the optimization loop, allowed for the effective exploration of the operating space and provided biologically reasonable recommendations using constraints. In addition, surrogate models offer the speed of virtual experimentation, which greatly minimizes the power of widely spread laboratory experimentation. Overall, the results prove that BO is an efficient method for the process of AD improvement that provides a scalable, interpretable, and experimentally viable path to performance improvement. This framework may be expanded in the future by researchers to multi-objective optimization, dynamic operating control, and real-time digital twin integration to manage intelligent biogas plants.

REFERENCES

1. Zhang Y, Yang X, Feng Y, Dai Z, Jing Z, Li Y, et al. Accelerating integrated prediction, analysis and targeted optimization for anaerobic digestion of biomass after hydrothermal pretreatment using automated machine learning. Renew Sustain Energy Rev. 2024;202:114688. doi:10.1016/j.rser.2024.114688.
2. Li Y, Chen Y, Wu J. Enhancement of methane production in anaerobic digestion process: A review. Appl Energy. 2019;240:120-137. doi:10.1016/j.apenergy.2019.01.243.
3. 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 (Basel). 2025;17(6):2688. doi:10.3390/su17062688.
4. Rutland H, You J, Liu H, Bull L, Reynolds D. A systematic review of machine-learning solutions in anaerobic digestion. Bioengineering (Basel). 2023;10(12):1410. doi:10.3390/bioengineering10121410. PMID:38136001.
5. Rashwan AK, Bai H, Osman AI, Eltohamy KM, Chen Z, Younis HA, et al. Recycling food and agriculture by-products to mitigate climate change: A review. Environ Chem Lett. 2023;21:3351-3375. doi:10.1007/s10311-023-01639-6.
6. Zang Y, Yang Y, Hu Y, Ngo HH, Wang XC, Li YY. Zero-valent iron enhanced anaerobic digestion of pre-concentrated domestic wastewater for bioenergy recovery: Characteristics and mechanisms. Bioresour Technol. 2020;310:123441. doi:10.1016/j.biortech.2020.123441. PMID:32361204.
7. Chan PC, Lu Q, de Toledo RA, Gu JD, Shim H. Improved anaerobic co-digestion of food waste and domestic wastewater by copper supplementation: Microbial community change and enhanced effluent quality. Sci Total Environ. 2019;670:337-344. doi:10.1016/j.scitotenv.2019.03.081. PMID:30904647.
8. Meola A, Wolf K, Weinrich S. Meta-tuning and fast optimization of machine learning models for dynamic methane prediction in anaerobic digestion. Bioresour Technol. 2025;432:132654. doi:10.1016/j.biortech.2025.132654. PMID:40355002.
9. 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.
10. Yu D, Liang Y, Thejani Nilusha R, Ritigala T, Wei Y. Prediction of the long-term effect of iron on methane yield in an anaerobic membrane bioreactor using Bayesian network meta-analysis. Membranes (Basel). 2021;11(2):100. doi:10.3390/membranes11020100. PMID:33572581.

