STM JOURNALS

International Journal of
Energy and Thermal Applications

Volume 4, Issue 1, 2026
January–April
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 labeled 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, 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 the Bayesian optimization (BO) as an efficient instrument of AD process optimization, which can allow making operational decisions that are data driven and minimize the number of experiments in waste-to-energy framework.

***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.
DOI: https://doi.org/10.37591/IJETA.v04i01.236321

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

INTRODUCTION

AD has become an essential biological process to transform agricultural residues into renewable-rich biogas as a source of energy and to provide renewable waste management. The productivity of the methane production, though, is very sensitive to operating conditions, temperature, pH [1, 2], OLR and the C/N ratio. The interaction of these parameters is not a linear and predictable process, and optimization of AD conditions is a rather difficult task. The traditional methods of empirical tuning make use of highly time-consuming and expensive laboratory tests [3], which are not as effective as scaling around the multi-dimensional operating space. This leads to a high percentage of AD systems not performing optimally and most potential for the using methane goes to waste.

New developments in machine learning and probabilistic modeling have created possibilities in speeding up AD optimization. BO, and 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 is efficient to balance exploration and exploitation to find optimal operation parameters. This makes it particularly appropriate to biological systems where experiments are slow, costly, and prone to variation in the processes, ensuring efficiency.

In a system with anaerobic digestion, Bayesian optimization is combined to find regimes of operation with high yield on a systematic and data-driven basis. There are surrogate models that can estimate responses to methane yield with high accuracy, including GPR, Deep Neural Networks (DNN), Random Forests (RF), and Support Vector Regression (SVR), and that 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 hence ensuring that the practical implementation remains feasible.

This paper will introduce a full Bayesian optimization-based optimization strategy of advancing 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, with the help of a structured AD dataset and a search strategy that relies on surrogate assistance. The results emphasize how Bayesian optimization can be used to revolutionize the design of AD processes and present a scalable, smart, and experimentally efficient method in case of real-life biogas systems.

LITERATURE REVIEW

The optimization of AD processes has received a lot of attention both in terms of experimental research and also in relation to computer modeling of the same. Initial experiments were mainly based on empirical modification of temperature, pH, hydraulic retention time and organic loading rate to enhance the generation of methane [1]. Though these methods were useful in offering background information, they were limited by the slow reaction cycle, small parameter space, 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 interdependence between operating condition, leading to poor yield enhancement.

Machine learning methods started to be important in the process of methane yield prediction and comprehending the process variables that were of importance. Regression models, neural networks, ensemble algorithms, and support-vector-based methods were shown to have high predictive accuracy when trained on structured AD datasets [3]. These techniques provided the means of estimating intricate digestion conduct without needing the finer mechanical knowledge. But the applications of most machine-learning tried to make only prediction and not optimization, which created a gap between model outputs and operating strategy that could be implemented [4]. Also, models that are prediction oriented generally demanded large data sets and made no natural measure of uncertainty - which is vital in the case of biological systems.

In more recent times, surrogate-assisted and probabilistic optimization frameworks have appeared in the field. The use of surrogate models in methane response surfaces in the form of Gaussian processes, deep neural networks, random forests, 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 which would be otherwise impossible to experiment on. Among the others, Gaussian process-based methods can be distinguished by the fact that they directly introduce uncertainty estimates in the optimization process, and thus, allow striking the balance between exploration of unknown high-performance areas and fine-tuning of known high-performance areas [6].

BO has also become a popular method of optimization of complex systems which are costly to evaluate like anaerobic digesters [7]. BO uses probabilistic surrogates and acquisition functions to find possible operating conditions that are promising. In comparison to traditional optimization methods, which make many evaluations, BO needs a moderate number of surrogate-directed cycles to reach high-yield results [8]. This is especially appropriate to AD processes, in which an experimental run can take several weeks [9, 10]. Although there are these benefits, the use of BO in the enhancement of yield of methane is still scarce and even the existing research sometimes does not involve any constraint of search strategies to maintain biologically plausible recommendations.

METHODOLOGY

The methodology that will be used in this study combines surrogate-assisted BO and constrained parameter search to savor maximum yield of methane through anaerobic digestion as presented in the Figure 1. This workflow includes six primary steps, including dataset preparation, feature preprocessing, construction of a surrogate model, Bayesian optimization configuration, constrained search operation, and the ultimate validation of optimized operating conditions. All of the stages are outlined below.

Dataset Preparation and Features Selection

The input in optimizing was a well-organized AD) data set that included physicochemical feedstock characteristics and working parameters. The important ones were TS, VS, C/N ratio, lignocellulosic composition, temperature, pH, HRT, and OLR. The target variable was methane yield (mL CH_4/g VS). The cleaning of the data was done by doing the following: Outliers were removed with the IQR procedure, missing values were filled with feature-wise Gaussian distributions, and all the continuous variables have been normalized with Minmax that would make the process of fitting the model to be smooth.

Surrogate Model Construction

The syllabus concentrates 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 (GPR):* to facilitate uncertainty-aware optimization,
- *Deep neural network (DNN):* to approximate nonlinear functions,
- *Random forest regressor (RFR):* for ensemble-based response estimation,
- Support vector regression (SVR): used to predict with a margin.

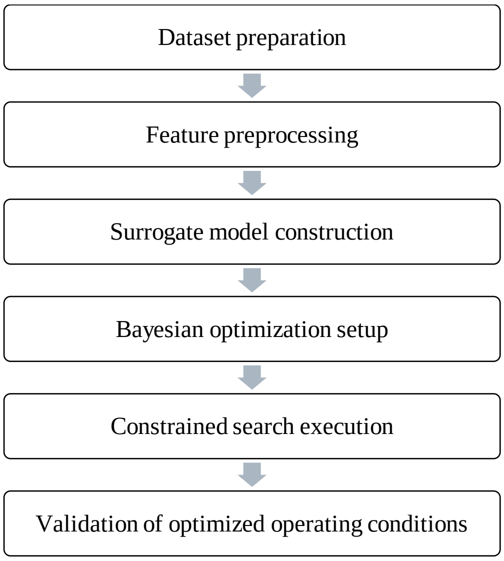

Figure 1. Framework for constrained bayesian optimization of AD performance.

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 chosen as the evaluation engine of Bayesian optimization.

Parameter Space Definition and Constraints

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

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

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

Bayesian Optimization Framework

Bayesian optimization (BO) using a Gaussian process prior, was utilized to effectively search the multidimensional search space. To control the exploration and exploitation, the Upper Confidence Bound (UCB) acquisition function was employed. At each iteration:

The next parameter set identified by the acquisition function was the set of candidate parameters. Methane yield and uncertainty were predicted by the surrogate model. Results were inputted into the GP model To refine 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 a series of five cycles.

Constrained Search Execution

The performance of BO was done in the presence of strict process constraints to prevent any recommendation of an unstable or inhibitory condition. Candidates that had candidate points that were violating 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. The convergence was ensured by monitoring optimization trajectories, the scores of acquisition and surrogate predictions in a biologically meaningful manner.

Validation of Optimized Operating Conditions

The best parameter combination that BO found was compared to baseline and pre-optimization performance. Comparison of predicted yields of methane using the surrogate against the average of the dataset was used to measure the improvement. Additionally:

Predicted vs. actual plots, error histograms, imprecision involves (GPR) and 3D response surfaces validation was done on the accuracy of optimization using were. An endpoint yield of 315 mL of CH_4/g VS that constitutes a 21.2% growth ensured the suitability of the Bayesian optimization strategy.

RESULTS AND DISCUSSION

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

Table 1 provides the limited search space of Bayesian optimization of the parameters of anaerobic digestion functioning. The lower and upper limits of temperature, pH, organic loading rate (OLR), carbon-to-nitrogen (C/N) ratio, hydraulic retention time (HRT), and mixing speed used (1. selected) were both biological feasible, and sufficiently flexible enough to be optimized. The 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 ranges (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 VS/L/day covers the conservative and high-throughput loading situations. Limiting factors in C/N ratio ensures sufficient availability of nutrients whereas the broad range of HRT ensures the investigation of both fast and slow digestion processes. The range of mixing speed was chosen to compromise between improving mass transfer and preventing shear stress. In general, such a clear search space guaranteed safe, efficient, and realistic optimization, which gave a strong basis in terms of Bayesian-based parameter optimization.

The convergence behavior of the Bayesian optimization process as the series of iterations progresses is shown in Figure 2. This curve demonstrates how rapidly the yield of methane grows in the initial iterations, which means that the parameter space is properly searched and promising areas of operation are identified quickly. The rate of improvement also decreases gradually as the number of iterations increases, indicating a shift in the process between exploration and exploitation. This trend of convergence proves the effectiveness of Bayesian optimization in finding near-optimal solutions at a low number of evaluations. The fact that the yield of the stabilization of methane was reached at later iterations is an indication that the optimizer has found an optimal or near-optimal operating configuration. The convergence behavior is especially useful in anaerobic digestion applications, where experimental assessments are inexpensive and time-intensive.

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

The development of the Expected Improvement (EI) acquisition function during the Bayesian optimization process is shown in Figure 4. The declining pattern of EI curve shows that the optimization framework is efficient to balance the exploration and exploitation at the successive iterations. In the early stages, the values of EI are found to be greater than the values result in more uncertainty in space of search and also a greater focus on exploration to come up with any good operating regions. The values of EI decrease slowly as the optimization continues, indicating that the surrogate model becomes confident about the predicted landscape of the yield of methane. This decrease proves that the algorithm is more oriented on the optimization of the best areas instead of trying to explore untried combinations of parameters. This kind of 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 a stable surrogate learning and the lack of the sporadic search behavior, which is pertinent to the sound optimization of the anaerobic digestion 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

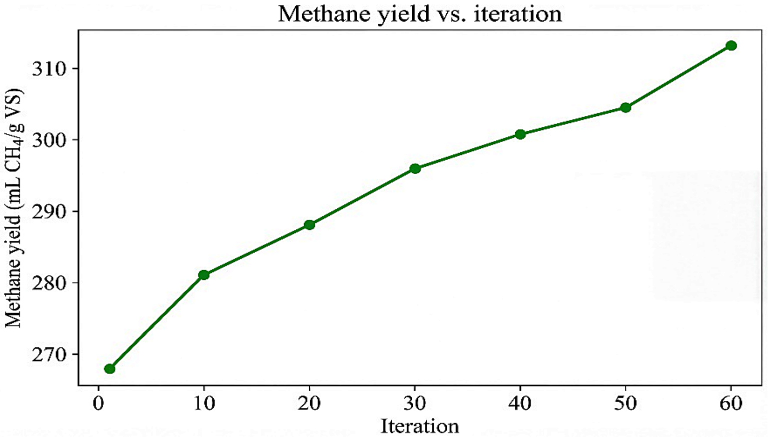

Figure 2. Bayesian optimization convergence curve.

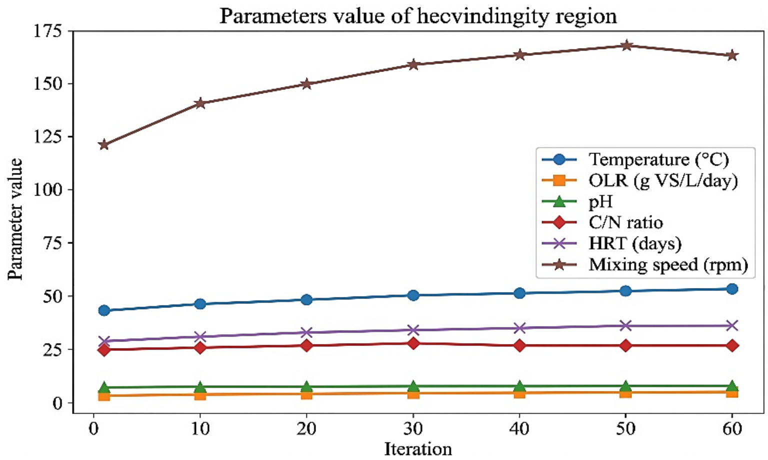

Figure 3. Parameter trajectories during optimization.

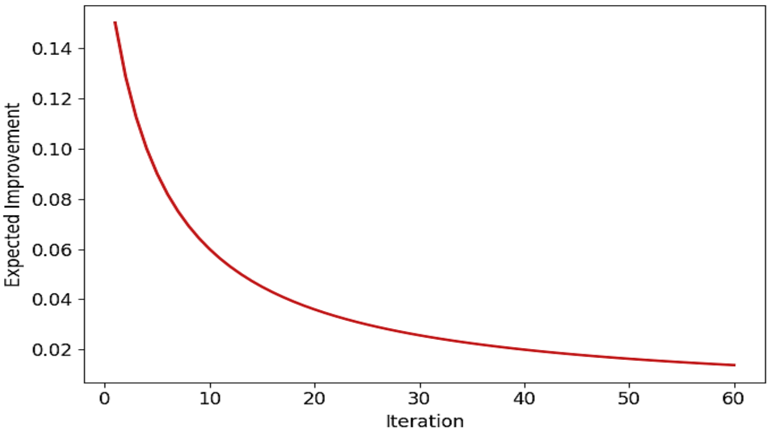

Figure 4. Acquisition function (EI) during optimization.

Comparison of the yield of methane at the baseline conditions, pre-optimization average conditions and post-optimization conditions derived using Bayesian optimization is presented in Figure 5. It is evident that there was a significant increase in the production of methane after optimization. Although the anaerobic digestion apparatus on the baseline conditions provides relatively lower levels of methane production, a progressive change is recorded when switching to dataset-average situations. It is optimized and after that, the enhancement is 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 in the maximization of biogas productivity. Yield improvement is also observed which confirms predictive accuracy of the surrogate-assisted optimization model which proves to be practical. In general, Figure 5 presents compelling facts that Bayesian optimization can generate practical performance objectives in an anaerobic digestion mechanism and, therefore, the use of the same in real world optimization of biogas plants. The results above corroborate that Bayesian optimization with surrogate modeling and constrained search space can be a potent and effective solution to the optimization of methane generation in a system of anaerobic digestion. The convergence behavior observed, the stable parameter curves, and the significant gains made by the proposed optimization framework all justify the strength of the proposed optimization framework. Notably, the optimized operational conditions are kept within biologically safe limits and operationally practical limits, which supports the practical use of the methodology. The research results can be used to form a solid base to experimentally validate and practically apply 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.

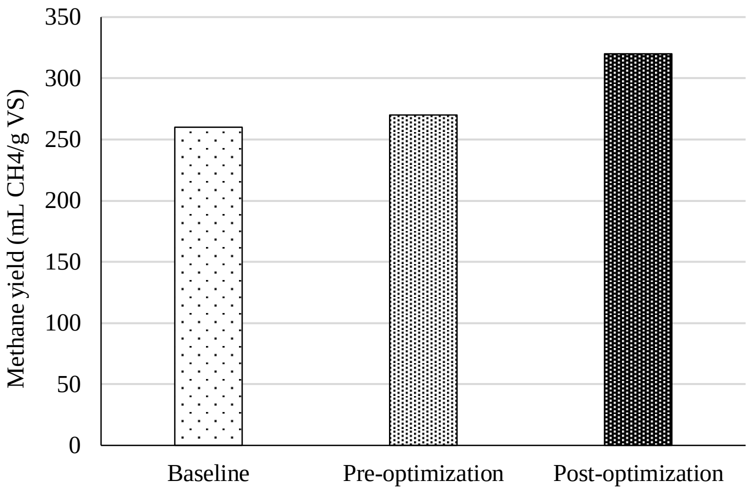

Figure 5. Methane yield before and after optimization.

CONCLUSION

This paper has shown that the Bayesian optimization is an efficient and data-driven model to maximize the yield of methane in an anaerobic digestion system. Combining a structured AD-data with surrogate modeling and uncertainty-sensitive Bayesian search, the proposed framework was able to identify the optimal operating conditions that boosted the production of methane significantly. 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 strives better success than the conventional manual techniques of tuning. The Gaussian process regression as a part of the optimization loop allowed to effectively explore the operating space and provide biologically reasonable recommendations by using constraints. In addition, the surrogate models offered the speed of virtual experimentation, which greatly minimized the power of widely spread laboratory experimentation. Altogether, the results prove that the Bayesian optimization is an efficient method of the process of anaerobic digestion improvement that provides a scalable, interpretable, and experimentally viable path to performance improvement. This framework may be expanded in the future by researchers to the multi-objective optimization, dynamic operating control, and real-time digital twin integration to manage intelligent biogas plants.

REFERENCES

1. Yi Zhang, Xingru Yang, Yijing Feng, Zhiyue Dai, Zhangmu Jing, Yeqing Li, Lu Feng, Yanji Hao, Shasha Yu, Weijin Zhang, Yanjuan Lu, Chunming Xu, Junting Pan, Accelerating integrated prediction, analysis and targeted optimization for anaerobic digestion of biomass after hydrothermal pretreatment using automated machine learning, Renewable and Sustainable Energy Reviews, Volume 202, 2024, 114688, https://doi.org/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.
3. Bensegueni, C., Kheireddine, B., Khalfaoui, A., Amrouci, Z., Bouznada, M. O., & Derbal, K. (2025). Optimization of Biogas and Biomethane Yield from Anaerobic Conversion of Pepper Waste Using Response Surface Methodology. Sustainability, 17(6), 2688. https://doi.org/10.3390/su17062688
4. Rutland, H., You, J., Liu, H., Bull, L., & Reynolds, D. (2023). A Systematic Review of Machine-Learning Solutions in Anaerobic Digestion. Bioengineering (Basel, Switzerland), 10(12), 1410. https://doi.org/10.3390/bioengineering10121410
5. Rashwan, A.K.; Bai, H.; Osman, A.I.; Eltohamy, K.M.; Chen, Z.; Younis, H.A.; Al-Fatesh, A.; Rooney, D.W.; Yap, P.-S. Recycling Food and Agriculture By-Products to Mitigate Climate Change: A Review. Environ. Chem. Lett. 2023, 21, 3351–3375.
6. Zang, Y.; Yang, Y.; Hu, Y.; Ngo, H.H.; Wang, X.C.; Li, Y.Y. Zero-valent iron enhanced anaerobic digestion of pre-concentrated domestic wastewater for bioenergy recovery: Characteristics and mechanisms. Bioresour. Technol. 2020, 310, 123441.
7. Chan, P.C.; Lu, Q.; de Toledo, R.A.; Gu, J.D.; 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.
8. Alberto Meola, Klara Wolf, Sören Weinrich, Meta-tuning and fast optimization of machine learning models for dynamic methane prediction in anaerobic digestion, Bioresource Technology, Volume 432, 2025, 132654, https://doi.org/10.1016/j.biortech.2025.132654.
9. Jeppu, G.P.; 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.
10. Yu, D., Liang, Y., Thejani Nilusha, R., Ritigala, T., & Wei, Y. (2021). Prediction of the Long-Term Effect of Iron on Methane Yield in an Anaerobic Membrane Bioreactor Using Bayesian Network Meta-Analysis. Membranes, 11(2), 100. https://doi.org/10.3390/membranes11020100

