Energy Advances

ROYAL SOCIETY OF CHEMISTRY

REVIEW ARTICLE

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d6ya00053c

Artificial intelligence for anaerobic digestion: advancing sustainable biogas production

Meicai Xu,[a] Carter Monson,[a] Josue Kpodo,[a] Fei Long,[b] Hong Liu,[b] Yan Liu[a] and Wei Liao *[a]

Received 25th February 2026, Accepted 8th June 2026

DOI: 10.1039/d6ya00053c

rsc.li/energy-advances

Anaerobic digestion (AD) is a technology for sustainable waste management and renewable energy production, yet its complexity and variability limit process efficiency and control. Advances in artificial intelligence (AI) offer solutions through improved predictive modeling, process optimization, and monitoring. This review provides a comprehensive and application-oriented synthesis of AI applications in AD from 2015 to 2025, covering methane yield prediction, microbial community integration, optimization, and system control. Commonly used models such as artificial neural networks, support vector machines, random forest, and hybrid frameworks are employed for predictive accuracy and design. AI-powered soft sensors enable non-invasive, real-time estimation of parameters such as volatile fatty acids, alkalinity, pH, and methane production, supporting early warning, anomaly detection, and adaptive control. Incorporating microbial data into these frameworks further enhances diagnostics and stability assessment, although such integration remains limited in current studies. Bibliometric analysis and case studies highlight research trends, innovations, and global hotspots, alongside persistent challenges in data quality, interpretability, interdisciplinary collaboration, and scale-up. Future priorities include standardized, open-access data infrastructures, explainable and robust models, sensor–AI integration, and multi-objective optimization frameworks that balance technical, environmental, and economic performance. The proposed roadmap bridges AI methodologies with AD-specific system characteristics, distinguishing this review from previous studies and providing a pathway toward practical, field-ready AD systems capable of accelerating the transition to sustainable waste-to-energy solutions.

1. Introduction

Anaerobic digestion (AD) is widely recognized as an effective technology for sustainable waste management, renewable energy production, and nutrient recovery.[1] Despite their advantages, AD systems are complex and sensitive to feedstock variability, microbial community dynamics, and operational fluctuations.[2] These factors pose challenges for real-time monitoring, process control, and overall system optimization,[3] particularly at the full scale.[4] Artificial intelligence (AI) and machine learning (ML) have recently emerged as promising tools to address these challenges.[3,5,6] Their ability to model nonlinear relationships, process high-dimensional data, and adaptively learn from operational feedback presents a valuable opportunity for AD systems. The increasing global focus on sustainable energy and waste management, coupled with rapid advances in AI, highlights the need for a comprehensive review in this field.

Several recent reviews have examined AI applications in AD, including their role in lignocellulosic biorefinery processes,[7] biochar-related applications,[8] co-digestion strategies for biomass,[9] and landfill leachate treatment.[10] Additional studies have focused specifically on AI-based optimization of biogas and methane production.[11–13] While these reviews provide useful insights, most remain limited in scope. They often emphasize AI algorithms without sufficient consideration of AD-specific challenges or focus on AD applications without integrating computational perspectives. In addition, emerging topics such as sensor-data fusion, soft sensor development, and explainable artificial intelligence are often insufficiently discussed, highlighting the need for a more integrated and forward-looking review.

To address these gaps, this review provides a structured synthesis of AI applications in AD across four major domains: biogas and/or methane yield prediction, microbial community integration, process optimization, and system monitoring and

[a] Department of Biosystem and Agricultural Engineering, Michigan State University, East Lansing, 48824, MI, USA. E-mail: liaow@msu.edu; Fax: +517-432-7205; Tel: +517-432-7205

[b] Department of Biological and Ecological Engineering, Oregon State University, Corvallis, 97331, OR, USA

control. The review combines bibliometric analysis with insights from recent case studies, offering a multidimensional perspective on the evolution, current capabilities, and key challenges of AI-enhanced AD systems. In addition to reviewing commonly used AI models – including artificial neural networks (ANNs), support vector machines (SVMs), random forest (RF), and hybrid methods like ANN-genetic algorithms (ANN–GA) and adaptive neuro-fuzzy inference system (ANFIS) – this work emphasizes practical implementation aspects, including integration with real-time sensor platforms, soft sensor development, and dynamic process control. By examining developments over the past decade (2015–2025), this review illustrates the co-evolution of AI methodologies and AD applications, revealing key methodological trends, shifting research priorities, and persistent gaps.

The novelty of this review lies in its combined focus on algorithmic development and system-level integration within AD. Rather than treating AI solely as a predictive tool, it is discussed within the broader context of AD system management, including challenges related to data quality, interdisciplinary collaboration, and real-world implementation. Three key research gaps are identified: (1) limitations in data availability, particularly the lack of standardized datasets integrating operational and microbial information; (2) the need for interpretable and robust AI models that support reliable decision-making; and (3) the importance of cross-sector collaboration to enable practical deployment in AD systems. Based on these observations, future research directions are proposed to improve the reliability, scalability, and applicability of AI-enhanced AD systems.

2. Methods

This section describes the methodological framework adopted in this review, including the literature search strategy, screening process, and data analysis approach (Fig. 1). The process began with the construction of two clusters of keywords to systematically identify literature on the application of machine learning and artificial intelligence in the field of anaerobic digestion. The first cluster comprised AD-related terms: "anaerobic digestion", "anaerobic co-digestion", "anaerobic reactor", and "anaerobic bioreactor". The second cluster focused on AI/ML concepts, including "machine learning", "artificial intelligence", "AI", "prediction", and "data-driven". A systematic search was conducted using the Web of Science Core Collection and Scopus databases, targeting peer-reviewed journal articles and reviews published between 2015 and 2025 (including early access articles available as of April 30, 2025). This search yielded 784 records from Web of Science and 946 records from Scopus. After merging the datasets and removing 543 duplicate records, 1187 unique entries remained for screening. Following manual review of titles and author keywords, 879 articles were excluded. Of the 308 reports selected for retrieval, two documents (graduate theses) were not accessible. The remaining 306 reports were assessed based on abstracts,

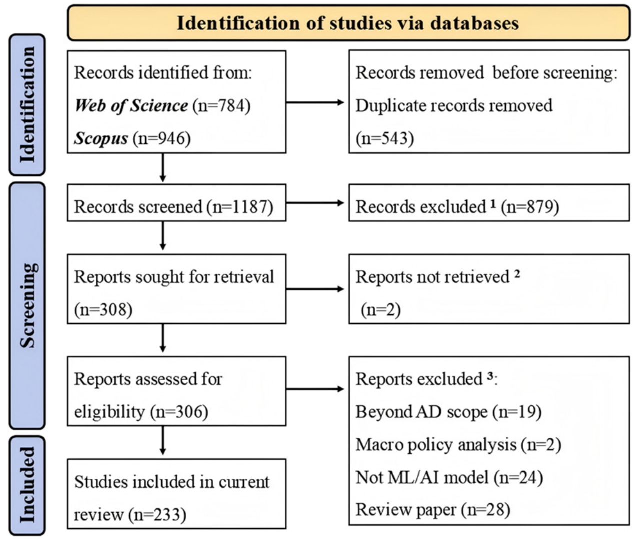

Fig. 1 PRISMA-style flow diagram of the literature screening and selection process. Notes: (1) records were excluded based on professional judgment after reviewing the title and keywords. (2) Graduate theses were not retrieved. (3) Reports were excluded after full abstract review.

leading to the exclusion of 45 articles that were either beyond the scope of anaerobic digestion (n = 19), focused on macro-level policy (n = 2), or lacking an AI/ML model (n = 24). Ultimately, 233 research articles from 100 journals were included in this review.

Excel 2019, VOSviewer©, and R (version 4.4.3) with the packages "Bibliometric",[14] "ggwordcloud", "ggplot2"[15] and "ggsankey" were used for data analysis and visualization. The analysis includes bibliometric and descriptive statistical approaches, focusing on journal distribution, publication year, keyword co-occurrence, disciplinary categorization, and geographical distribution. In addition, AI applications in AD are summarized, and key research gaps and future directions are identified through a conceptual roadmap.

3. Bibliometric analysis

3.1 Distribution by journal and publication year

Fig. 2 shows the top 10 most frequent sources of the 233 selected documents, along with their annual distribution from January 2015 to April 2025. Bioresource Technology stands out as the leading journal, publishing 28 articles (over 10% of the total), significantly exceeding Journal of Cleaner Production and Fuel, which ranked second with 10 publications each. Together, these top 10 journals account for 116 publications (45% of the total), indicating their central role in AI-related AD research. The annual distribution reveals a clear growth trend, from 3 publications in 2015 to a peak of 74 in 2024, representing a 25-fold increase. A notable surge occurred between 2022 and 2023, reflecting the rapid expansion of research interest in AI applications, following the widespread public and scientific attention garnered by the emergence of ChatGPT.[16] The apparent decline in 2025 publications does

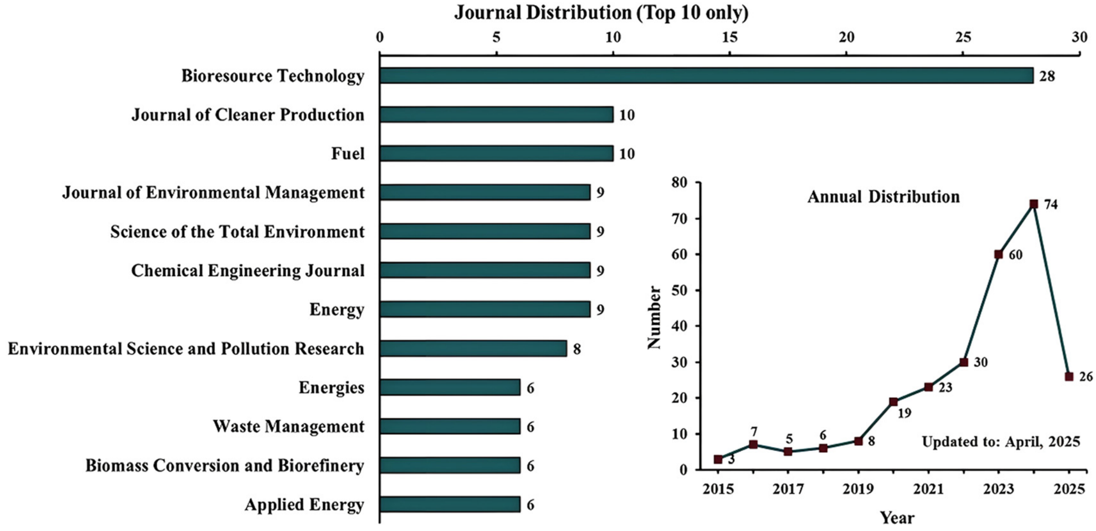

Fig. 2 Journal and annual distribution of the 233 included publications (top 10 journals, including ties).

not indicate a real decrease but reflects the data collection cutoff in April 2025. As a result, the 2025 values represent partial-year data and are not directly comparable with previous full-year values.

3.2 Keyword, disciplinary, and geographic distributions

Fig. 3 presents three key dimensions of the bibliometric analysis for the 233 included publications. Fig. 3a illustrates the co-occurrence network of author keywords, with "anaerobic

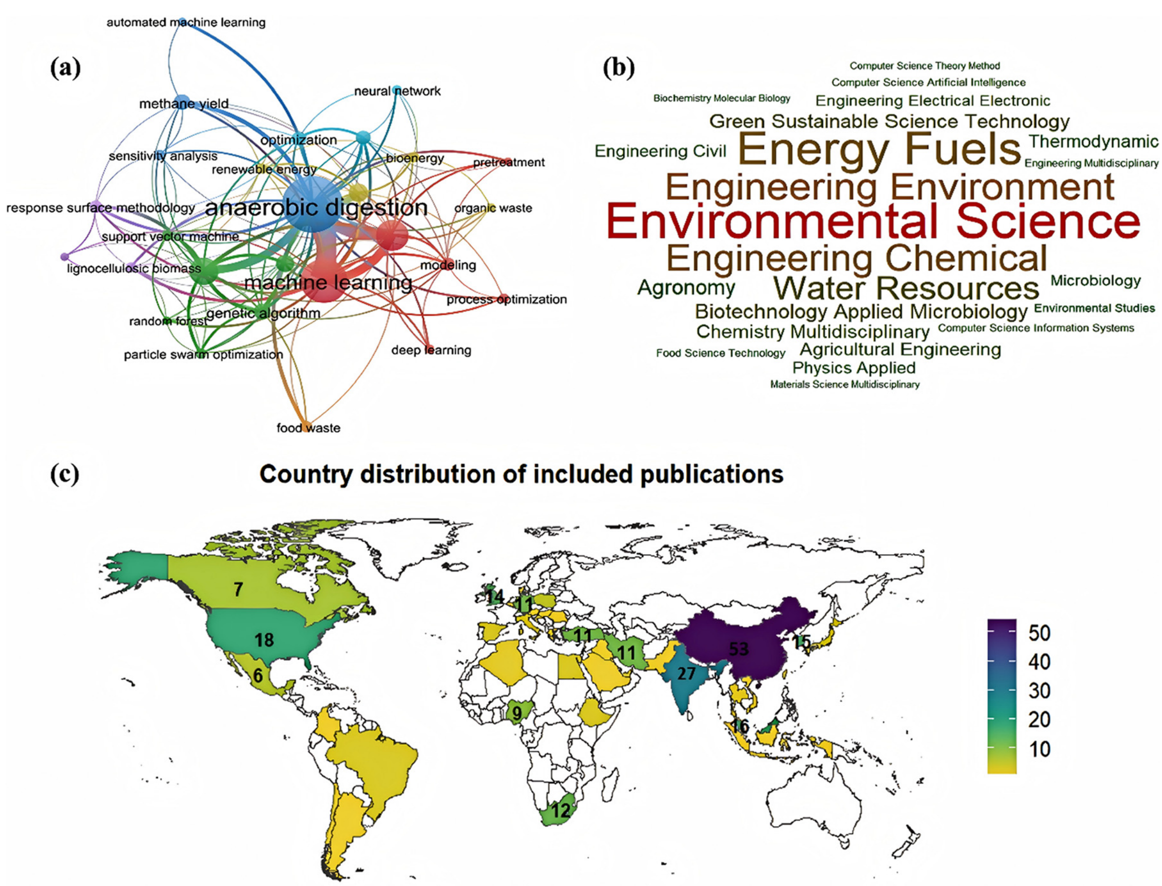

Fig. 3 Bibliometric analysis of the 233 included publications: (a) keyword co-occurrence, (b) disciplinary categorization, and (c) country distribution based on the first author's affiliation.

digestion'' and ''machine learning'' being the most dominant nodes, defining the core scope of this review. Clustered visualization highlights major thematic areas such as methane yield, renewable energy, optimization, modeling, and genetic algorithms, reflecting the methodological diversity and research focus within the field. Fig. 3b shows the disciplinary categorization of the publications based on Web of Science journal classifications. The three most prevalent categories – environmental science, engineering environmental, and energy fuels – underscore the environmental and energy-oriented nature of AI-related AD research. Fig. 3c shows the geographic distribution of publications based on the affiliation of the first author. The 233 articles originate from 42 countries, with China (n = 54), India (n = 29), and the United States (n = 18) as the leading contributors. This distribution reflects the broad global engagement in AI-integrated AD research, particularly in regions with strong energy demand and ongoing technological development. Overall, the results indicate a growing and widely distributed research interest in AI-integrated AD, accompanied by increasing methodological diversity.

4. AI methodologies and trends in anaerobic digestion research

4.1 Overview of AI model categories

Artificial intelligence models applied in AD studies can be broadly categorized into classical machine learning, tree-based ensemble methods, neural networks, advanced deep learning, and hybrid or optimization-enhanced approaches. This classification reflects differences in model complexity, data requirements, and typical applications in AD systems.

4.1.1 Classic and tree-based ML algorithms. Classic ML algorithms have been widely applied in AD research due to their simplicity, computational efficiency, and interpretability. This category includes models such as linear regression (LR), multiple linear regression (MLR), Gaussian process regression (GPR), support vector machines (SVM), *K*-nearest neighbors (KNN), and Naïve Bayes (NB), which can be grouped into: (a) regression models (LR, MLR, and GPR) and (b) classification models (SVM, KNN, and NB).[17] Regression models are typically used to predict process performance indicators (*e.g.*, biogas yield), whereas classification models are applied to identify abnormal operating conditions such as process inhibition.[18] LR and MLR serve as representative regression models, offering high interpretability and computational efficiency, particularly when linear relationships are assumed. GPR extends regression to nonlinear relationships and provides uncertainty estimates that can support process monitoring and control. KNN provides a simple and flexible approach for detecting process anomalies, although its performance depends on distance metrics and neighbor selection. SVM can address both classification and regression tasks and is particularly suitable for small datasets, capturing nonlinear relationships through kernel-based transformations.

Tree-based ensemble algorithms have attracted increasing attention in AD research due to their ability to model complex nonlinear relationships and provide robust predictive performance. Among these, random forest (RF), gradient boosting machine, and extreme gradient boosting are the most applied approaches. These models improve generalization through ensemble strategies such as bagging and boosting, enabling more accurate prediction of AD processes. In addition, other tree-based methods such as LightGBM, CatBoost, and AdaBoost have also been explored. These algorithms have been used in comparative studies with XGBoost for predicting biogas production,[19] estimating the biodegradation of dissolved organic matter,[20] and predicting volatile fatty acid concentrations in sludge AD processes.[21]

4.1.2 Neural network and deep learning approaches. Artificial neural networks are widely used ML models for capturing nonlinear, high-dimensional relationships in AD research. Among various ANN architectures, the multilayer perceptron (MLP) and backpropagation neural network (BPNN) are the most frequently applied ones in AD studies. MLP captures nonlinear relationships between process variables and target outputs, while BPNN uses backpropagation to minimize prediction errors and learn patterns in noisy datasets. Several studies have demonstrated the effectiveness of BPNN in predicting AD performance metrics such as biogas production,[22,23] methane content,[24] and COD concentrations.[23]

In addition to these traditional ANNs, advanced neural network architectures have been explored to address temporal dependencies, spatial heterogeneity, and complex feature interactions in AD systems. Long short-term memory (LSTM) networks and recurrent neural networks (RNN) are designed to capture sequential patterns and temporal dependencies in time-series data,[25] such as daily biogas yields, making them suitable for dynamic system modeling. LSTM models have been applied to predict biogas production in lab-scale AD processes treating food waste,[26] as well as in large-scale wastewater treatment plant applications,[27] achieving high predictive accuracy in both ($R^2 > 0.8$). Convolutional neural networks (CNNs), originally developed for image and signal processing tasks,[28] have been adapted to extract spatial features from multivariate AD datasets. In addition, graph convolutional networks (GCNs) have gained attention for modeling complex interactions in graph-structured data,[29] including microbial co-occurrence networks and system component relationships with AD systems.[30] Despite these advantages, advanced deep learning models require larger datasets, higher computational resources, and careful hyperparameter tuning compared to traditional ANNs.[31] Therefore, their application in AD should consider the trade-offs between model complexity, data availability, computational cost, and interpretability.

4.1.3 Optimization-enhanced and other AI models. Optimization-enhanced and hybrid ML models are increasingly used in AD research to improve predictive accuracy and process optimization. In this category, GA and particle swarm optimization (PSO) are widely used as global optimization techniques for tuning model parameters and configurations.[32,33] These methods can be used independently or combined with traditional ML models, forming hybrid systems that integrate the search capabilities of GA or PSO with the predictive power of models such as ANNs.[6] Several studies have applied the ANN–PSO framework to predict and optimize biogas production from AD treatment of cattle manure,[34] sweet sorghum bagasse,[35] and sewage sludge.[36] Response surface methodology (RSM), a classical statistical tool, has been used to model and optimize AD processes by identifying optimal operating conditions.[37–39] Hybrid models, such as ANN–GA, ANN–PSO, and ANN–RSM, combined global optimization techniques with predictive modeling to address complex, multi-dimensional problems in AD research. In addition to these approaches, ANFIS provides an alternative hybrid framework that integrates neural network learning with fuzzy logic to model nonlinear and uncertain systems.[40] Recently, ANFIS has been applied to predict biogas and/or methane production from lab-scale[41] to plant-scale AD[42,43] and has been reported to achieve higher prediction accuracy than RSM in some cases.[44,45] Another emerging approach is automated machine learning (AutoML), which automates model selection, feature engineering, and hyperparameter tuning, enabling rapid development of predictive models with minimal manual intervention.[46] Popular AutoML frameworks (*e.g.*, TPOT, NNI, and H_2O AutoML) have been applied to improve efficiency and reproducibility in AD modeling.

4.2 Evolution of AI algorithms in AD

A comparative analysis of algorithm usage within each AI category reveals distinct patterns in AD research over the past decade, based on 233 peer-reviewed articles included in this review. In the classic ML category, SVM shows the highest frequency (65 counts), followed by KNN (28), LR (18), and MLR (14), reflecting their ease of implementation and broad applicability (Fig. 4a). In the tree-based Ensemble category, RF has the highest frequency (62 counts), followed by XGBoost (32), GBM (17), and DTR (16). These results indicate a growing preference for ensemble-based methods due to their strong predictive performance and ability to capture nonlinear relationships. ANNs exhibit the highest overall frequency (74 counts) among all algorithms, with common subtypes including MLP (26) and BPNN (15). They are widely applied in AD for performance prediction, process optimization, and system monitoring. In contrast, advanced deep learning models – such as LSTM (8 counts), RNN (5), and CNN (3) – are less frequently used (Fig. 4a), mainly due to higher data and computational requirements. Nevertheless, promising applications have been reported, including LSTM combined with genetic algorithms for predicting biogas production in large-scale municipal AD systems[27] and simulation data on agricultural waste.[47] In the optimization-enhanced and hybrid category, ANFIS (18 counts), RSM (17), and hybrid models such as RSM–ANN (6) show moderate usage. In the others category, applications remain limited. AutoML (8 counts), has been introduced to streamline model development, while two time-series approaches – autoregressive integrated moving average (3) and nonlinear autoregressive neural networks (3) – have also been explored (Fig. 4a). Overall, commonly used AI algorithms in AD research include ANN, SVM, RF, XGBoost, KNN, LR, ANFIS, RSM, GBM, and DTR, reflecting a combination of traditional and advanced methods for diverse modeling needs.

Fig. 4b illustrates the yearly relative distribution of six AI algorithm categories in AD research based on the 233 reviewed articles published between 2015 and 2025. A clear trend of methodological diversification is observed. While ANN models

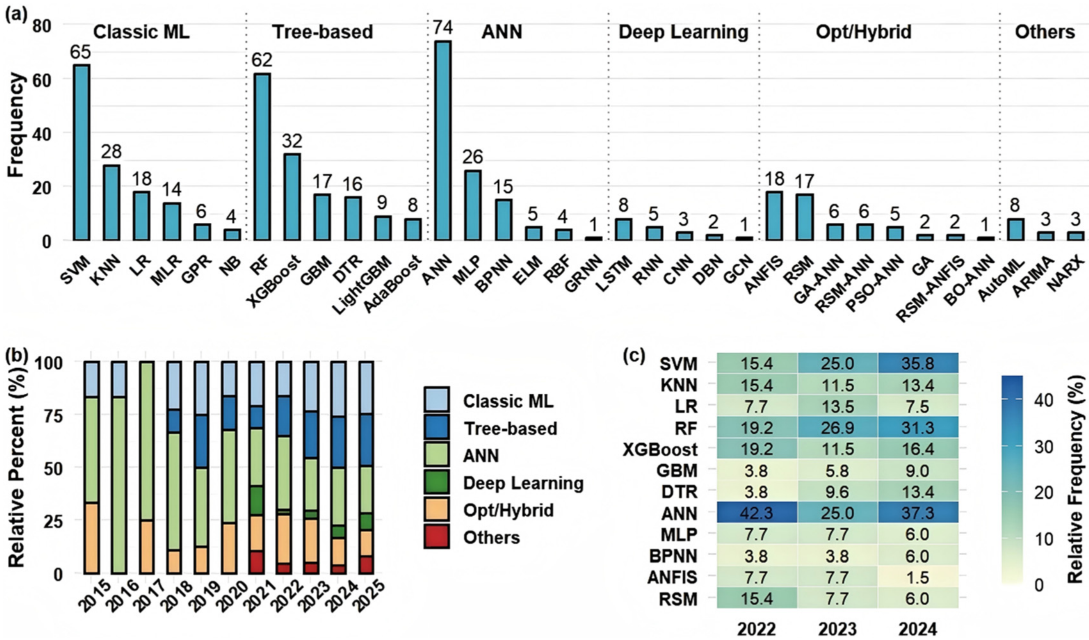

Fig. 4 AI model categories, algorithms, and temporal distribution based on 233 research articles: (a) total occurrences of individual algorithms; (b) yearly percentage distribution of six algorithm categories, normalized by article count in each year; and (c) heatmap showing the relative frequency of high-occurrence algorithms (≥15 times) in the nearest three years.

dominated earlier years, the use of tree-based, deep learning, and hybrid approaches has increased over time. For example, in 2023, the proportions – normalized to 100% based on the total number of articles and algorithms used that year – were classic ML (23.38%), tree-based (22.08%), ANN (24.48%), and optimization-enhanced and hybrid (20.78%), indicating a more balanced adoption of AI approaches. Fig. 4c presents a heatmap of high-occurrence algorithms ($\geq$15 times, as shown in Fig. 4a) during the most recent three complete years (2022–2024), highlighting recent trends in algorithm usage. The relative frequency of SVM, RF, and DTR shows a consistent upward trend, indicating increasing adoption in recent AD studies.[6] In contrast, KNN, LR, and XGBoost remain relatively stable, suggesting their continued role as baseline models in AD research[48] (Fig. 4c).

Overall, ANN and tree-based ensemble models (*e.g.*, RF and XGBoost) generally show strong predictive performance in AD applications due to their ability to capture nonlinear relationships. In contrast, models such as SVM, LR, and KNN persist as baseline approaches because of their simplicity and interpretability. Increased model complexity does not always lead to better performance, and simpler models may perform comparably under certain conditions, as reported in recent studies.

4.3 Trends in hybrid and explainable AI

In addition to the growing diversity of individual algorithm categories, recent years have seen an increasing adoption of hybrid AI models and explainable AI tools in AD research. Hybrid models – those that integrate two or more AI methods – combine the strengths of different algorithms to improve prediction accuracy, process optimization, and model adaptability. ANN has been frequently coupled with optimization algorithms such as GA, particle swarm optimization (PSO), and response surface methodology (RSM) to fine-tune model parameters and enhance generalization across AD systems. For example, one study applied ANN and SVR models optimized *via* GA and PSO to predict biogas yield from sewage sludge. The hybrid SVR–GA model achieved high accuracy while reducing computational time and improving interpretability through SHAP analysis.[36] Another study applied ANN, SVM, and RSM to model hydrogen production from organic waste, followed by GA and PSO optimization. The SVM model outperformed ANN and RSM in prediction accuracy (R^2 = 0.988, RMSE = 0.0103), and subsequent integration with GA and PSO enabled efficient parameter tuning, with PSO showing faster convergence.[49] Beyond ANN-based approaches, ANFIS has also been integrated with GA and PSO to improve modeling accuracy and support decision-making in AD applications. For instance, a PSO–ANFIS hybrid model was developed to optimize methane yield from alkali-pretreated ground nut shells, outperforming the GA–ANFIS model across multiple performance indicators.[50] Similarly, an ANFIS–PSO model for maximizing methane production from wastepaper showed better predictive performance than conventional RSM techniques.[51] These results reflect the growing use of hybrid frameworks in AD research.

At the same time, the demand for transparency and reliable AI-based decision-making has led to the increased use of explainable AI tools. While some algorithms, such as DT, KNN, and naïve Bayes, are inherently interpretable, more complex models often require model-agnostic interpretability methods.[52] These methods are used to explain model behavior at both global and local levels. Local approaches such as SHAP have been applied to interpret the outputs of black-box models like RF and deep neural networks. For example, SHAP was used in a biogas yield prediction study using SVR and ANN models, identifying influent volatile solids and temperature as key factors and supporting feature selection for optimization.[36] SHAP has also been combined with ensemble learning models to predict biogas yield and methane concentration from hydrothermal carbonation wastewater treatment. This approach enabled identification of interactions among key variables (*e.g.*, temperature, pH, and COD), supporting improved interpretation of AD performance.[53] The increasing use of explainable AI reflects a shift from performance-focused modeling toward more interpretable and application-oriented approaches in AD systems. Hybrid and explainable models are expected to play an important role in linking algorithm development with practical AD system management.

5. Applications of AI across anaerobic digestion processes

To provide an overview of how AI has been applied in AD research, Fig. 5 summarizes the methodological characteristics of the 233 peer-reviewed articles included in this review. From a study purpose perspective (Fig. 5a), the majority of studies focused on performance prediction (68.60%), followed by process optimization (18.77%) and system monitoring (9.56%). A smaller proportion addresses classification and model validation tasks, reflecting the strong emphasis on predictive applications of AI in AD research.[6] These percentages are calculated based on the total number of purpose-specific entries rather than the number of unique studies, as individual articles often address multiple objectives (*e.g.*, prediction and optimization). Substrate sources are diverse, including agricultural and animal waste, food and kitchen waste, municipal sludge, and industrial waste, highlighting the broad applicability of AD for organic waste treatment[54] (Fig. 5b). Regarding experiment scale biomethane potential tests (BMP), laboratory-scale and pilot-scale studies together accounted for over 70% of all studies, indicating limited data from full-scale AD systems (Fig. 5c). In terms of data sources, lab-generated datasets dominate, with relatively few studies using public or simulated datasets, suggesting a need for greater data diversity and accessibility (Fig. 5d). Single-substrate digestion (65%) is more frequently studied than co-digestion systems (35%) (Fig. 5e), and only a small fraction of studies incorporate genus-level microbial data into AI models (Fig. 5f), despite their relevance to AD process mechanisms.[55]

Fig. 6 presents a Sankey diagram illustrating the relationships among study purposes (left), AI model categories (middle), and target variables (right). The left column

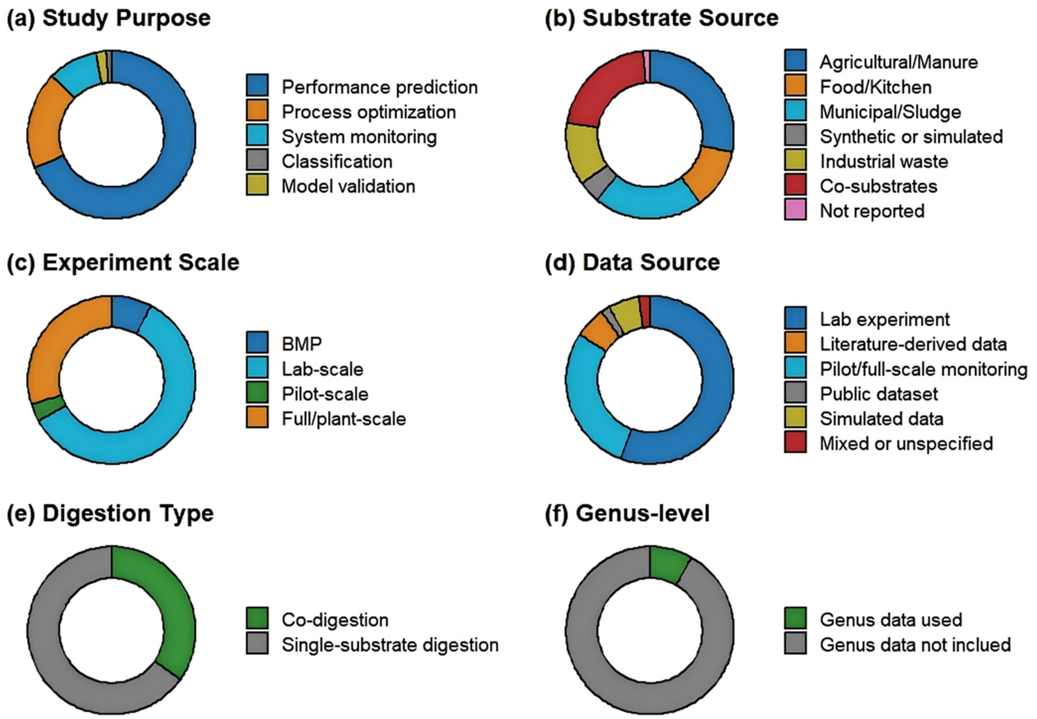

Fig. 5 Overview of anaerobic digestion studies based on 233 research articles: (a) study purpose; (b) substrate source; (c) experiment scale; (d) data source; (e) digestion type; and (f) genus-level data usage. Doughnut charts display the proportion of studies in each category, highlighting key methodological features within the AD-focused subset of literature.

corresponds to the classification of research objectives shown in Fig. 5a, while the middle column represents the six AI model categories introduced in Section 4.1. The width of each flow reflects the number of studies linking these categories across the three dimensions. Based on this framework, the following subsections (5.1–5.4) are organized into four main research themes: biogas and/or methane yield prediction, microbial community integration, process optimization, and system monitoring and control. Each section synthesizes representative studies, highlighting methodological trends and the role of AI across different stages of AD processes.

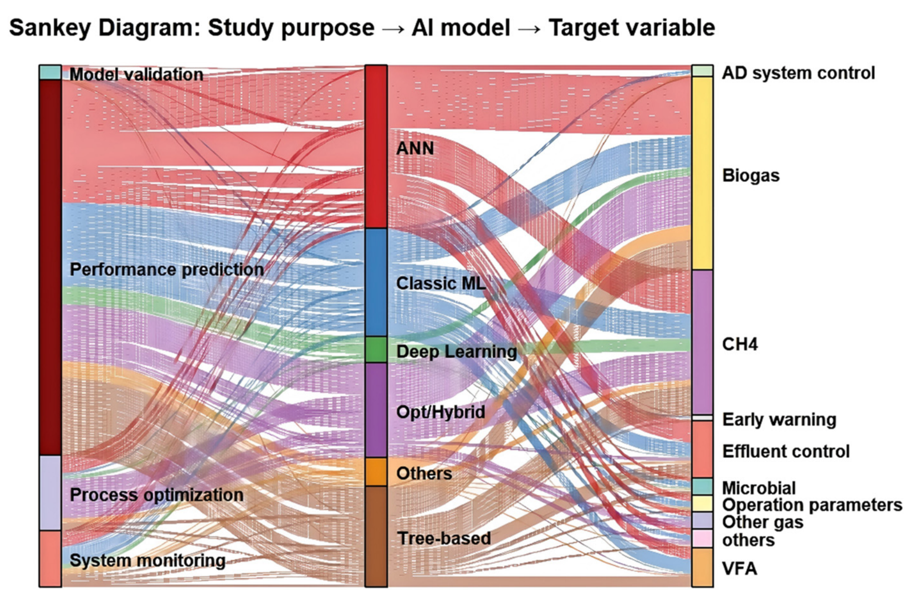

Fig. 6 The Sankey diagram illustrates the relationships among study purposes (left), AI model categories (middle), and target variables (right) in anaerobic digestion research. The left column corresponds to the classification of research objectives shown in Fig. 5(a), while the middle column represents the six AI model categories introduced in Fig. 4. The width of the flow reflects the number of studies connecting each category across the three dimensions.

5.1 Biogas and/or methane yield prediction

Biogas and/or methane yield prediction represents the most extensively studied AI application in AD research, spanning all six AI model categories (Fig. 6). Many studies employ BMP experiments as the basis for predictive modeling, as BMP provides standardized conditions for evaluating substrate biodegradability and methane yield.[56] Among the applied AI algorithms, ANNs are the most widely used and consistently show strong predictive performance, with R^2 values exceeding 0.95 across diverse substrates such as food waste,[57] cheese whey,[58] tea factory waste,[59] and activated sludge.[60] In addition, early-stage BMP data have been used to train ANN models for predicting biogas production within 3–14 days, demonstrating the potential to shorten experimental duration.[61]

Beyond BMP-based research, lab-scale studies have investigated biogas and methane prediction under more complex conditions using different data sources. Simulated datasets, particularly those generated by the AD model No. 1 (ADM1), are often used to evaluate model performance under controlled conditions. For example, ADM1-based data have been used to compare linear and nonlinear models, showing that hybrid ANN models combined with optimization algorithms significantly improve prediction accuracy (R^2 = 0.97).[62,63] In parallel, literature-derived datasets provide broader coverage of substrates and operating conditions. One study integrating long-term AD datasets reported that tree-based models outperformed ANN, SVM, and KNN in methane prediction.[64] These approaches are complementary: simulated data support controlled model evaluation, while literature-derived data enhance generalizability.

AI applications have also extended to pilot- and full-scale AD systems, where operational variability introduces additional complexity. For instance, one pilot-scale study employed fuzzy Mamdani models, ANN, and RSM to predict biogas production from poultry waste and cow dung in a modular biodigester, achieving R^2 values up to 1.0 with minimal prediction errors.[65] At full-scale, ANFIS has been widely applied. In a municipal treatment facility, ANFIS outperformed other models in predicting biogas production using five key process variables as the input (R^2 = 0.88).[66] Another study applied ANFIS to estimate biogas production rates in a cold-region wastewater treatment plant, showing strong agreement with measured data and demonstrating robustness under seasonal variations.[42]

Recent advances have also highlighted the potential of LSTM networks for full-scale AD prediction. These models have shown high accuracy, particularly when combined with data augmentation, genetic algorithms, or hyperparameter optimization, effectively capturing time-dependent process behavior under fluctuating conditions.[26,27,47] Methodological developments have further advanced AI applications in full-scale environments. One study developed an ML model incorporating optimization techniques and uncertainty quantification, achieving close agreement between predicted and measured biogas production in a full-scale wastewater treatment plant.[67] Another proposed a hybrid extreme learning machine model with data balancing and optimization, improving prediction accuracy (R^2 = 0.97) under varying loads.[68] Another study applied GA- and PSO-optimized ANN and SVR models, using SHAP analysis to identify key process variables and achieve strong predictive performance in a full-scale wastewater treatment plant.[36] Overall, biogas prediction research has evolved from controlled BMP experiments to full-scale, real-time applications. AI models increasingly support both operational optimization and long-term system planning, highlighting their expanding role in AD process modeling.

5.2 Microbial community integration

As shown in Fig. 5f, only a limited number of AD studies have incorporated genus-level microbial community data, indicating a clear gap in integrating microbiological information into AI-based modeling. In this review, 16 studies that incorporate microbial data into AI models were identified and are summarized in Table 1. These studies cover a range of modeling approaches and research objectives. Some are also related to biogas/methane prediction (as discussed in Section 5.1) but are included here due to their focus on microbial data integration.

Several studies have demonstrated the value of incorporating microbial community features to improve biogas and methane yield predictions. For example, one study applied an RF model and achieved high predictive accuracy (R^2 = 0.9879) by integrating bacterial genera with environmental factors, identifying *Keratinibaculum* and *Acetomicrobium* as key contributors to methane yield.[69] Another study used a multi-layer automated machine learning framework to examine interactions between archaeal genera and hydraulic retention time, showing that their combined input improved biogas prediction accuracy.[70] Similarly, an ANN–PSO model using 16S rRNA metagenomic data achieved $R^2 > 0.995$ for methane potential prediction from sweet sorghum bagasse.[35] Further work has applied microbial community network analysis to full-scale upflow anaerobic sludge reactor and continuous stirred tank reactor systems. Using RF, key taxa such as *Methanospirillum*, *Methanosphaera*, and *Methanobrevibacter* were identified as relevant to process performance under varying operational conditions.[71] These examples illustrate how genus-level microbial data can improve the accuracy, robustness, and interpretability of AI models for AD performance prediction.

Beyond biogas and methane production, several studies have used AI to assess microbial indicators related to process monitoring and environmental risk. Two studies focused on predicting antibiotic resistance genes (ARGs) dynamics in AD systems. One applied decision tree algorithm to forecast changes in ARG abundances following thermal hydrolysis-AD treatment of dairy waste, achieving a R^2 value of 0.87, demonstrating the feasibility of using ML to capture ARG dynamics.[72] Another work compared ANN, RF, and XGBoost models trained on experimental data extracted from 33 published studies to simulate ARG/MGE dynamics, with ANN showing superior predictive performance and identifying residence time and feed characteristics as dominant features.[73] These studies highlight the potential of AI to support ARG risk assessment and inform

Table 1 Summary of AI-based anaerobic digestion (AD) studies incorporating microbial community analysis (n = 16)

AI models	Microbial data source	Study objective/target variable	Experiment scale	Model performance	Year	Ref.
RF, SVM, KNN, XGBoost	16 S rRNA	Predict biogas yield	Laboratory	R^2 = 0.9879 (RF)	2025	69
GCN	High-throughput sequencing data	Predict microbial dynamics and biogas production	Laboratory	MSE = 0.11, R^2 = 0.72	2025	30
MLR, GBDT	16 S rRNA	Predict biogas production	Full-scale plant	R^2 = 0.975 (MLR)	2024	74
SVM, LR, DT	16 S rRNA	Predict antibiotic resistance gene (ARG) removal efficiency	Laboratory	R^2 = 0.87 (DT)	2024	72
GA, RF	Genomic data (SMOTER-augmented)	Predict methane production	Simulation	RMSE $<$ 0.1	2024	75
SVM, ET, XGBoost, ANN, CNN	16 S rRNA	Model relationships among microbial community, pH, and VFAs	Simulation	R^2 $>$ 0.8	2024	76
GBM, RF, XGBoost	16 S rRNA	Predict biogas production	Laboratory	RMSE = 84.21 (GBM)	2023	70
ANN–PSO	16 S rRNA	Predict methane production	BMP	R^2 $>$ 0.995 (ANN–PSO)	2023	35
DNN	nrMAGs, 16 S rRNA	Analyze microbiome structure for methanogenesis understanding	Full-scale plant	Not reported	2023	77
DaDa2	16 S rRNA	Link AD operational configurations and reactor performance with microbial community dynamics	Pilot-scale	Correlation-based	2023	78
Non-LR, MLR	16 S rRNA	Predict total viable bacterial counts	Laboratory	R^2 = 0.959, MSE = 0.18	2022	79
RF, XGBoost, KNN, ANN	Literature-derived microbial data	Predict medium-chain carboxylic acid (MCCA) performance	Laboratory	R^2 = 0.87 (RF)	2022	80
RF, XGBoost, ANN	Literature-derived microbial data	Predict ARG variation in AD process	Laboratory	R^2 $>$ 0.6 (ANN)	2022	73
SVM, KNN, RF, XGBoost	Literature-derived microbial data	Predict methane production	Laboratory	R^2 = 0.82 (RF)	2021	81
RF	16 S rRNA	Correlation microbiome with methane production	Full-scale plant	Correlation-based	2020	71
GB, DL-ANN, DRF	Flow cytometry (phenotypic microbial data)	Predict microbial functional group abundance	Laboratory	R^2 $>$ 0.9 (DL)	2018	82

process-level strategies to reduce environmental dissemination. Overall, microbial community data – particularly at the genus level – remain underutilized but show strong potential to improve AI-driven modeling of AD systems. Future research should expand their integration to improve system understanding and predictive performance across multiple operational and environmental indicators.

5.3 Process optimization

Building on the optimization-enhanced algorithms introduced in Section 4.1.3, this section examines how AI and hybrid models – such as ANN–GA, ANN–PSO, RSM–ANN, and ANFIS – are used not only to improve predictive accuracy but also to optimize key operational parameters in AD systems. These approaches are increasingly applied to identify operating conditions that enhance process efficiency, energy output, and system stability.

For example, one study integrated a deep belief network with a boosted osprey optimization algorithm to identify conditions that maximized biogas production (31.35 m^3 min^{-1}), enabling real-time adjustment of operating parameters and improving energy recovery in full-scale systems.[83] Another study compared ANFIS and RSM for modeling biogas yield from pretreated *Ulva intestinalis*, showing that ANFIS consistently achieved higher accuracy.[84] Similarly, ANFIS demonstrated strong performance in optimizing co-digestion of poultry waste and cow dung, achieving near-perfect prediction accuracy.[44] In a related study, ANN models were applied to optimize biochar dosage in cattle manure and green algae mixtures, outperforming RSM in identifying optimal conditions across multiple variables.[85]

Beyond yield-focused optimization, AI has been applied to optimize operational configurations across different stages of AD systems. For instance, ANN was used to simulate rotor placement and speed in biogas power plants, supporting efficiency estimation and system design.[86] Similarly, ANN combined with PSO was applied to reduce H_2S emissions in a full-scale AD facility, identifying operating conditions that resulted in a 49% decrease in H_2S.[87] ANN has also been used to optimize microbial electrolysis cell-assisted AD system by identifying voltage conditions that maximize energy recovery and biogas production.[88] Beyond the core AD process, AI has been extended to downstream and cross-sectoral applications. One study developed an integrated AD-gasification model using a T-ANN framework to simulate biomass conversion and determine digestion durations that maximize carbon emission reductions.[89] In agricultural applications, digestate management has been optimized using gradient boosting and principal component analysis, showing that tailored combinations of digestate and chemical fertilizer can enhance crop yield and support sustainable nutrient management.[90] Overall, AI-driven optimization is applied across multiple aspects of AD, including yield enhancement, emissions reduction, and resource recovery. Expanding these approaches to integrate energy, environmental, and resource objectives will further improve the practical value of AI in AD systems.

5.4 System monitoring and control

Anaerobic digestion is a dynamic and complex biological process involving multiple time-dependent variables. Effective monitoring and control are essential for maintaining system stability, optimizing biogas production, and ensuring robust operation under varying conditions. In this context, soft sensors have emerged as key tools for real-time estimation of process states that are difficult to measure directly.[18] These data-driven models support state estimation,[91] early warning,[74] anomaly detection,[92] and fault diagnosis,[93] thereby improving system observability and decision-making in AD operations.

Early warning is particularly important, as it enables intervention before minor fluctuations develop into severe disturbances or system failure. Recent studies demonstrate the effectiveness of soft sensors and ML algorithms in enhancing early detection. For example, a VFA-based soft sensor using SVM and principal component analysis showed strong performance in detecting small-magnitude faults.[92] Another study developed an early warning model combining an improved sparrow search algorithm with least square SVM, accurately predicting multiple operational indicators across datasets.[94] In addition, a large-scale database based on 75 AD start-up cases was analyzed using RF to identify key factors influencing start-up duration and failure risk across different systems.[95]

Beyond early warning, real-time monitoring is a major application of soft sensors in AD systems. AI-driven frameworks integrating online sensor data with predictive models enable continuous process monitoring and optimization. Examples include data-driven scheduling systems for reducing operational costs,[96] real-time data fusion platforms combining sensor signals with mechanistic insights,[97] and ANN-based estimators validated in complex wastewater systems.[98] Soft sensors have also been used to model key indicators such as alkalinity, supporting proactive control and process stability.[99] Applications have also expanded to more complex systems. For instance, ML models – including regression, tree-based, and neural network approaches – have been applied to a bio-electrochemical AD system, where one-step-ahead models using pH as an input showed effective performance for real-time methane prediction and stability control.[100] Another study using full-scale industrial data applied ANN-based sensitivity analysis to identify key operational parameters influencing biogas production, demonstrating its value for process optimization.[101] These examples illustrate how soft sensor approaches are being adapted to diverse AD configurations and real-world conditions.

Most monitoring studies have been conducted at pilot or full-scale,[78,91,102] emphasizing practical applicability. Among monitored variables, volatile fatty acids remain the most widely used due to their sensitivity to process fluctuations and ease of measurement.[76,92,103–105] Continued development of AI-enhanced soft sensors and integrated monitoring frameworks will be critical for improving the reliability and adaptability of AD systems. Collectively, these applications highlight the expanding role of AI across AD processes, particularly in prediction, optimization, and monitoring. This trend reflects a shift from proof-of-concept studies toward more application-oriented implementations in AD systems.

6. Research gaps and future directions

Artificial intelligence has shown strong potential to advance AD research and practice, particularly in process modeling, optimization, and system monitoring.[106] However, most existing studies remain at the proof-of-concept or laboratory scale, where conditions are controlled and datasets are relatively clean. In contrast, real-world AD systems operate under variable feedstock characteristics, fluctuating environmental conditions, and imperfect sensor networks, which limit the transferability of lab-trained models. These differences highlight persistent methodological, practical, and operational challenges that hinder the broad adoption of AI in AD systems.

To address these challenges, this section synthesizes key research gaps and proposes a roadmap for advancing intelligent, reliable, and sustainable AD systems (Fig. 7). The framework is organized into four interconnected stages: data foundation, model development, decision support, and deployment, which provides practical guidance for overcoming key barriers to AI adoption in industrial AD systems.

6.1 Data foundations

High-quality, context-rich data are essential for AI-driven AD systems. Currently, AI-based AD research is constrained by fragmented, inconsistent, and often incomplete datasets,[3] most of which originate from laboratory-scale experiments under simplified and controlled conditions (Fig. 5c). While such data provide useful baseline insights, they do not fully represent the variability and complexity of full-scale AD systems, limiting the robustness and generalizability of models across different feedstocks, reactor configurations, and environmental conditions.[107] To address these limitations, future work should focus on developing real-time sensor networks, IoT-enabled monitoring platforms,[108] and standardized data acquisition protocols. These approaches enable continuous measurement of key process variables, including temperature, pH, volatile fatty acids, methane concentration, and ammonia, thereby improving system observability and control responsiveness (Fig. 7).

Beyond operational metrics, genomic-scale microbial datasets, such as those derived from 16S rRNA sequencing and metagenomic analyses, provide additional insight into microbial dynamics and functional stability in AD systems.[109–112] However, these datasets remain underutilized due to analytical complexity and the lack of consistent linkage to process metadata. Addressing this gap requires the development of open-access, cross-scale databases that integrate biological, chemical, and engineering data under standardized formats. Standardized protocols for data collection, annotation, and interoperability, together with closer collaboration between academia and

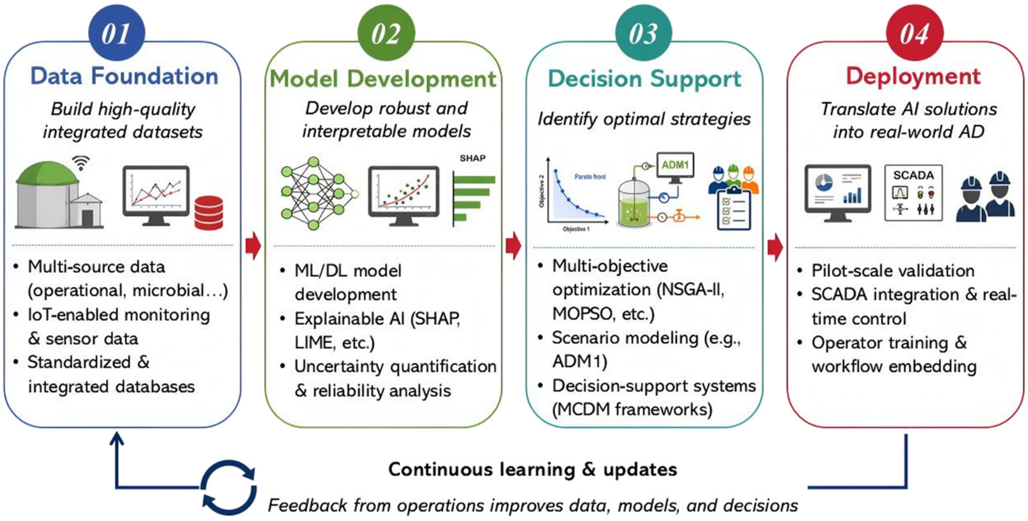

Fig. 7 Stage-based roadmap for advancing AI-integrated AD systems from laboratory concepts to field-ready applications. The framework outlines a sequential workflow from data foundation and model development to decision support and real-world deployment. A feedback loop connects all stages, enabling continuous learning and iterative system improvement based on operational data.

industry, will improve the quality and usability of data for subsequent model development.

6.2 Model development and decision support

AI models show strong capability in predicting methane yield,[113–115] detecting process anomalies,[92] and optimizing AD performance.[90,116–118] However, their application in industrial settings requires a shift from purely predictive models to more robust, interpretable, and decision-oriented frameworks (Fig. 7). This transition involves developing ML models that capture the nonlinear dynamics of AD processes, followed by integration of explainable AI tools to improve transparency.[119] Techniques such as SHAP and LIME enable identification of key variables and relationships, helping stakeholders understand model behavior and build confidence in AI-assisted decisions. In addition to interpretability, uncertainty quantification is important for improving model reliability under real-world conditions characterized by noise, missing data, and operational variability. Together, these elements provide a basis for AI systems that support decision-making rather than prediction alone.

Building upon this foundation, AI frameworks should further evolve toward multi-objective optimization and scenario-based decision support, reflecting the multi-dimensional nature of AD systems (Fig. 7). In practice, AD operation requires balancing multiple objectives, including methane production, greenhouse gas emissions, economic cost, and process stability.[120] Optimization algorithms such as NSGA-II and MOPSO can generate pareto-optimal solutions, while coupling AI models with process-based simulations (*e.g.*, ADM1) enables scenario analysis under different operational conditions. These results can then be translated into actionable insights through multi-criteria decision-making (MCDM) frameworks and interactive visualization tools, allowing stakeholders to evaluate trade-offs and select appropriate strategies. In addition, increasing reliance on data-driven decision-making raises considerations related to transparency, data governance, and regulatory compliance. Responsible AI deployment requires models to be interpretable, results to be reported in a standardized method, and applications to comply with environmental regulations.

6.3 Deployment and continuous learning

Successful deployment of AI in AD systems depends on translating model outputs into operational practice. This process typically begins with pilot-scale validation, where AI models are tested under realistic conditions to evaluate performance, robustness, and scalability. Following validation, integration with supervisory control and data acquisition (SCADA) systems enables real-time monitoring, automated control, and continuous optimization in full-scale facilities. Embedding AI tools into existing operational infrastructures allows model outputs to support process adjustment and decision-making. Technological integration alone, however, is insufficient without interdisciplinary collaboration. Co-design involving engineers, plant operators, and data scientists is essential to ensure that AI solutions align with practical constraints, including maintenance requirements, budget limitations, and regulatory frameworks (Fig. 7). In addition, capacity building supports this transition by enabling stakeholders to interpret and effectively use AI-driven systems.

Real-world deployment also creates a closed-loop learning system, where operational data are fed back into data collection and model development. This iterative process allows models to adapt to changing conditions, improve over time, and maintain performance in dynamic environments (Fig. 7). Institutional support, including pilot projects, funding mechanisms, and industry partnerships, can accelerate this process by providing testbeds for validation and refinement. Through this progression – from data foundation to model development, decision support, and deployment – AI-integrated AD systems can move from laboratory studies to field applications, supporting long-term sustainability and circular bioeconomy objectives.

7. Conclusions

This review synthesizes recent advances in AI applications for anaerobic digestion from 2015 to 2025, covering prediction, optimization, monitoring, and model validation. By integrating bibliometric analysis with methodological and case-study insights, it identifies both the significant progress achieved and the persistent gaps that must be addressed to transition from laboratory proof-of-concept to operational AI–AD systems. Further development of AI-integrated AD systems requires coordinated improvements in data quality, model reliability, and system-level integration. In particular, standardized datasets, explainable modeling approaches, and closer integration with monitoring and control systems are essential for practical deployment. The roadmap presented in this review provides a structured pathway to bridge the gap between simulation and real-world implementation, supporting the development of robust and field-ready AD systems for sustainable waste-to-energy applications.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

During the preparation of this work, the author(s) used ChatGPT to improve grammatical accuracy and word choice. After using this tool/service, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the published article.

References

1 S. Kumar Khanal, F. Lü, J. W. C. Wong, D. Wu and H. Oechsner, *Bioresour. Technol.*, 2021, **337**, 125378.
2 A. Roopnarain, H. Rama, B. Ndaba, M. Bello-Akinosho, E. Bamuza-Pemu and R. Adeleke, *Renew. Sustain. Energy Rev.*, 2021, **152**, 111717.
3 H. Rutland, J. You, H. Liu, L. Bull and D. Reynolds, *Bioengineering*, 2023, **10**, 1410.
4 S. Sidi Habib, S. Torii and K. S. Mol, *J. Renewable Energy Environ*, 2024, **11**, 9–27.
5 V. G. Sharmila, S. P. Shanmugavel and J. R. Banu, *Biomass Bioenergy*, 2024, **180**, 106997.
6 I. Andrade Cruz, W. Chuenchart, F. Long, K. C. Surendra, L. Renata Santos Andrade, M. Bilal, H. Liu, R. Tavares Figueiredo, S. K. Khanal and L. Fernando Romanholo Ferreira, *Bioresour. Technol.*, 2022, **345**, 126433.
7 X. Y. Huang, X. Zhang, L. Xing, S. X. Huang, C. Zhang, X. C. Hu and C. G. Liu, *Bioresour. Technol.*, 2025, **428**, 132434.
8 P. Zhang, T. Zhang, J. Zhang, H. Liu, C. Chicaiza-Ortiz, J. T. E. Lee, Y. He, Y. Dai and Y. W. Tong, *Carbon Neutrality*, 2024, **3**, 2.
9 M. O. Fajobi, O. A. Lasode, A. A. Adeleke, P. P. Ikubanni and A. O. Balogun, *Energy Sources, Part A*, 2022, **44**, 5314–5339.
10 N. A. F. Zamrisham, A. M. A. Wahab, A. Zainal, D. Karadag, D. Bhutada, S. Suhartini, M. A. Musa and S. Idrus, *Water*, 2023, **15**, 1303.
11 L. Zhang, K. C. Loh and J. Zhang, *Bioresour. Technol. Rep.*, 2019, **5**, 280–296.
12 A. M. Enitan, J. Adeyemo, F. M. Swalaha, S. Kumari and F. Bux, *Rev. Chem. Eng*, 2017, **33**, 309–335.
13 K. C. Obileke, G. Makaka, S. Tangwe and P. Mukumba, *Environ. Dev. Sustain.*, 2025, **27**, 15025–15051.
14 M. Aria and C. Cuccurullo, *J. Informetr.*, 2017, **11**, 959–975.
15 H. Wickham, *Wiley Interdiscip. Rev. Comput. Stat.*, 2011, **3**, 180–185.
16 J. J. Zhu, M. Yang, J. Jiang, Y. Bai, D. Chen and Z. J. Ren, *Environ. Sci. Technol. Lett.*, 2024, **11**, 1327–1333.
17 S. Asgari, R. Gupta, I. K. Puri and R. Zheng, *Appl. Soft Comput.*, 2021, **110**, 107638.
18 R. Gupta, L. Zhang, J. Hou, Z. Zhang, H. Liu, S. You, Y. Sik Ok and W. Li, *Bioresour. Technol.*, 2023, **369**, 128468.
19 A. Ahmad, A. K. Yadav, A. Singh and D. K. Singh, *Biomass Bioenergy*, 2024, **180**, 106995.
20 J. Liu, C. Wang, J. Zhou, K. Dong, M. Elsamadony, Y. Xu, M. Fujii, Y. Wei and D. Wang, *Bioresour. Technol.*, 2024, **412**, 131382.
21 U. A. Abubakar, G. S. Lemar, A. A. D. Bello, A. Ishaq, A. A. Dandajeh, Z. T. Jagun and M. R. Houmsi, *Environ. Sci. Pollut. Res.*, 2025, **32**, 28239–28252.
22 F. Tufaner, Y. Avşar and M. T. Gönüllü, *Clean Technol. Environ. Policy*, 2017, **19**, 2255–2264.
23 H. Y. Zhao, F. L. Huang, L. Li and C. Y. Zhang, *Desalination Water Treat.*, 2018, **122**, 30–35.
24 D. Dominguillo-Ramírez, J. Aburto, H. Hugo Leon-Santiesteban and E. Martinez-Hernandez, *Fuel*, 2023, **344**, 128053.
25 G. Van Houdt, C. Mosquera and G. Nápoles, *Artif. Intell. Rev.*, 2020, **53**, 5929–5955.
26 Z. Geng, X. Shi, B. Ma, C. Chu and Y. Han, *Environ. Sci. Pollut. Res.*, 2024, **31**, 9121–9134.
27 M. M. Salamattalab, M. Hasani Zonoozi and M. Molavi-Arabshahi, *Waste Manage.*, 2024, **175**, 30–41.
28 H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar and P. A. Muller, *Data Min. Knowl. Discov.*, 2019, **33**, 917–963.
29 T. N. Kipf and M. Welling, in 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings, 2017.
30 H. G. Kim, S. Il Yu, S. G. Shin and K. H. Cho, *Water Res.*, 2025, **274**, 123144.
31 G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken and C. I. Sánchez, *Med. Image Anal.*, 2017, **42**, 60–88.

32 D. Wang, D. Tan and L. Liu, *Soft Comput.*, 2018, **22**, 387–408.
33 K. Deb, Sadhana - Academy Proceedings in Engineering Sciences, DOI: 10.1007/BF02823145.
34 B. K. Zaied, M. Rashid, M. Nasrullah, B. Sama Bari, A. W. Zularisam, L. Singh, D. Kumar and S. Krishnan, *Biomass Conv. Bioref.*, 2023, **13**, 73–88.
35 L. Machineni and G. R. Anupoju, *Environ. Sci. Pollut. Res.*, 2023, **30**, 114095–114110.
36 F. Farzin, S. S. Moghaddam and M. Ehteshami, *Renew. Energy*, 2024, **227**, 120554.
37 S. Krushna Bhujbal, P. Ghosh and V. Kumar Vijay, *Sustain. Energy Technol. Assess.*, 2024, **71**, 104006.
38 Y. Zhan and J. Zhu, *Appl. Energy*, 2024, **355**, 122336.
39 J. W. Chen, Y. J. Chan, S. K. Arumugasamy and S. K. Yazdi, *J. Water Process Eng.*, 2023, **52**, 103493.
40 J. S. R. Jang, *IEEE Trans. Syst. Man Cybern.*, 1993, **23**, 665–685.
41 K. O. Olatunji, D. M. Madyira, N. A. Ahmed, O. Adeleke and O. Ogunkunle, *Waste Biomass Valorization*, 2023, **14**, 1123–1141.
42 M. Asadi, H. Guo and K. McPhedran, *J. Environ. Manage.*, 2020, **253**, 109708.
43 R. Flores-Asis, J. M. Méndez-Contreras, A. Alvarado-Lassman, G. Fernández-Lambert, D. Villanueva-Vásquez and A. A. Aguilar-Lasserre, *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.*, 2019, **54**, 582–592.
44 M. O. Okwu, O. D. Samuel, D. R. E. Ewim and Z. Huan, *Int. J. Energy Environ. Eng.*, 2021, **12**, 353–363.
45 K. O. Olatunji, N. A. Ahmed, D. M. Madyira, A. O. Adebayo, O. Ogunkunle and O. Adeleke, *Renew. Energy*, 2022, **189**, 288–303.
46 X. He, K. Zhao and X. Chu, *Knowl. Based. Syst.*, 2021, **212**, 106622.
47 A. Meola, M. Winkler and S. Weinrich, *Bioresour. Technol.*, 2023, **372**, 128604.
48 S. Ö. Cinar, S. Cinar and K. Kuchta, *Fermentation*, 2022, **8**, 65.
49 C. Mahata, S. Ray and D. Das, *Energy Convers. Manag.*, 2020, **219**, 113047.
50 K. O. Olatunji, S. O. Oladipo, D. M. Madyira and Y. Sun, *Waste Biomass Valorization*, 2025, **16**, 423–440.
51 A. G. Olabi, A. M. Nassef, C. Rodriguez, M. A. Abdelkareem and H. Rezk, *Int. J. Energy Res.*, 2020, **44**, 9598–9608.
52 A. Adadi and M. Berrada, *IEEE Access*, 2018, **6**, 52138–52160.
53 Z. Ma, R. Wang, G. Song, K. Zhang, Z. Zhao and J. Wang, *Sci. Total Environ.*, 2024, **908**, 168279.
54 P. Weiland, *Appl. Microbiol. Biotechnol.*, 2010, **85**, 849–860.
55 K. Venkiteshwaran, B. Bocher, J. Maki and D. Zitomer, *Microbiol. Insights*, 2015, **8**, 33593.
56 C. Holliger, M. Alves, D. Andrade, I. Angelidaki, S. Astals, U. Baier, C. Bougrier, P. Buffière, M. Carballa, V. De Wilde, F. Ebertseder, B. Fernández, E. Ficara, I. Fotidis, J. C. Frigon, H. F. De Laclos, D. S. M. Ghasimi, G. Hack, M. Hartel, J. Heerenklage, I. S. Horvath, P. Jenicek, K. Koch, J. Krautwald, J. Lizasoain, J. Liu, L. Mosberger, M. Nistor, H. Oechsner, J. V. Oliveira, M. Paterson, A. Pauss, S. Pommier, I. Porqueddu, F. Raposo, T. Ribeiro, F. R. Pfund, S. Strömberg, M. Torrijos, M. Van Eekert, J. Van Lier, H. Wedwitschka and I. Wierinck, *Water Sci. Technol*, 2016, **74**, 2515–2522.
57 D. Singh, M. Tembhare, N. Machhirake and S. Kumar, *Energy*, 2023, **263**, 126138.
58 V. E. Cordoba, J. Mussi, M. De Paula and G. G. Acosta, *IEEE Lat. Am. Trans.*, 2023, **21**, 1032–1039.
59 S. Özarslan, S. Abut, M. R. Atelge, M. Kaya and S. Unalan, *Fuel*, 2021, **306**, 121715.
60 P. Kianmehr and F. Kfoury, *Ozone Sci. Eng.*, 2016, **38**, 465–471.
61 S. M. Hunter, E. Blanco and A. Borrion, *Bioresour. Technol. Rep.*, 2024, **26**, 101845.
62 T. Beltramo, C. Ranzan, J. Hinrichs and B. Hitzmann, *Biosyst. Eng.*, 2016, **143**, 68–78.
63 T. Beltramo and B. Hitzmann, *Eng. Agric. Environ. Food*, 2019, **12**, 397–403.
64 M. Ghazizade Fard and E. H. Koupaie, *Bioresour. Technol.*, 2024, **394**, 130255.
65 M. O. Okwu, O. D. Samuel, O. B. Otanocha, L. K. Tartibu, H. O. Omoregbee and V. M. Mbachu, *Biomass Conv. Bioref.*, 2023, **13**, 917–926.
66 K. Yetilmezsoy, K. Karakaya, M. Bahramian, S. A. Abdul-Wahab and B. İ. Goncaloğlu, *Neural Comput. Appl.*, 2021, **33**, 11043–11066.
67 M. Samkhaniani, S. S. Moghaddam, H. Mesghali, A. Ghajari and N. Gozalpour, *Waste Manage.*, 2025, **197**, 14–24.
68 Y. Zhang, L. Li, Z. Ren, Y. Yu, Y. Li, J. Pan, Y. Lu, L. Feng, W. Zhang and Y. Han, *Bioresour. Technol.*, 2022, **363**, 127899.
69 Y. Guo, Y. Zhao, Z. Li, Z. Wang, W. Zhang, K. Lin and T. Zhou, *Bioresour. Technol.*, 2025, **416**, 131762.
70 Y. Zhang, Z. Jing, Y. Feng, S. Chen, Y. Li, Y. Han, L. Feng, J. Pan, M. Mazarji, H. Zhou, X. Wang and C. Xu, *Chem. Eng. J.*, 2023, **475**, 146069.
71 E. C. G. Vendruscolo, D. Mesa, D. V. Rissi, B. H. Meyer, F. de Oliveira Pedrosa, E. M. de Souza and L. M. Cruz, *Sci. Total Environ.*, 2020, **742**, 140314.
72 H. Su, T. Zhu, J. Lv, H. Wang, J. Zhao and J. Xu, *Bioresour. Technol.*, 2024, **399**, 130536.
73 N. Haffiez, T. H. Chung, B. S. Zakaria, M. Shahidi, S. Mezbahuddin, R. Maal-Bared and B. R. Dhar, *Sci. Total Environ.*, 2022, **839**, 156211.
74 Y. Sato, K. Hasemi, K. Machikawa, H. Kinjo, N. Yashiro, Y. Iimura, H. Aoki and H. Habe, *Bioresour. Technol.*, 2024, **402**, 130766.
75 Z. Wang, F. Wu, N. Hao, T. Wang, N. Cao and X. Wang, *J. Clean. Prod.*, 2024, **466**, 142909.
76 S. Il Yu, H. Jeong, J. Shin, S. G. Shin, A. Abbas, D. Yun, H. Bae and K. H. Cho, *J. Water Process Eng.*, 2024, **60**, 105225.
77 R. Wirth, Z. Bagi, P. Shetty, M. Szuhaj, T. T. S. Cheung, K. L. Kovács and G. Maróti, *ISME J.*, 2023, **17**, 1326–1339.

78 H. Dang, N. Yu, Y. Zhang, L. Zhang and Y. Liu, *J. Environ. Eng.*, 2023, **149**, 04023014.
79 S. Tangwe, P. Mukumba and G. Makaka, *Energies*, 2022, **15**, 7407.
80 F. Long, J. Fan, W. Xu and H. Liu, *J. Clean. Prod.*, 2022, **377**, 134223.
81 F. Long, L. Wang, W. Cai, K. Lesnik and H. Liu, *Water Res.*, 2021, **199**, 117182.
82 A. S. Dhoble, P. Lahiri and K. D. Bhalerao, *J. Biol. Eng.*, 2018, **12**, 19.
83 J. Duan, G. Cao, G. Ma and B. Yazdani, *Sci. Rep.*, 2025, **15**, 4814.
84 U. O. Aigbe, K. E. Ukhurebor, A. O. Osibote, M. A. Hassaan and A. El Nemr, *Renew. Energy*, 2025, **3**, 3483–3498.
85 H. Şenol, E. Çolak, E. A. Elibol, M. A. Hassaan and A. El Nemr, *Chem. Eng. J.*, 2024, **493**, 152750.
86 A. Heller, H. Pomares and P. Glösekötter, *Meas. Sci. Technol.*, 2026, **37**, 195903.
87 N. Hasanpour Seyedlar, S. M. Zamir, M. Nosrati and E. R. Rene, *Renew. Energy*, 2024, **232**, 121016.
88 V. T. Nguyen, Q. T. H. Ta and P. K. T. Nguyen, *Biochem. Eng. J.*, 2022, **187**, 108670.
89 Y. Ge, J. Tao, Z. Wang, C. Chen, R. Liang, L. Mu, H. Ruan, Y. Rodríguez Yon, B. Yan and G. Chen, *Bioresour. Technol.*, 2023, **369**, 128420.
90 B. K. Show, S. Panja, R. GhoshThakur, A. Basu, A. Koley, A. Ghosh, K. Pramanik, S. Chaudhury, A. K. Hazra, N. Dey, A. B. Ross and S. Balachandran, *Sustainability*, 2023, **15**, 13706.
91 M. Kim and F. Cui, *J. Environ. Manage.*, 2023, **347**, 119153.
92 P. Kazemi, J. Giralt, C. Bengoa and J. P. Steyer, *Water Sci. Technol*, 2020, **81**, 1740–1748.
93 I. Juárez-Barojas, R. Posada-Gómez, A. Alvarado-Lassman and J. P. Rodríguez-Jarquín, *Electronics*, 2023, **12**, 799.
94 Y. Chen, Z. Huang, C. Ma, Z. Li, Z. Zhang, T. Tan and Y. Chen, *Chem. Eng. J.*, 2024, **490**, 151743.
95 A. Hmaissia, Y. Bareha and C. Vaneeckhaute, *J. Environ. Manage.*, 2024, **359**, 121068.
96 F. Piadeh, I. Offie, K. Behzadian, A. Bywater and L. C. Campos, *Bioresour. Technol.*, 2024, **392**, 130017.
97 X. Wang, I. Rashid, Z. Zhao, M. Oladele, W. Xiang, Y. Huang, E. Wazer, J. McCutcheon, G. Bollas, J. Contreras and B. Li, *ACS ES&T Water*, 2024, **4**, 1061–1072.
98 L. Dewasme, *Water Sci. Technol.*, 2020, **80**, 1975–1985.
99 X. Wang, X. Bai, Z. Li, X. Zhou, S. Cheng, J. Sun and T. Liu, *Biochem. Eng. J.*, 2018, **140**, 85–92.
100 A. Cheon, J. Sung, H. Jun, H. Jang, M. Kim and J. Park, *Processes*, 2022, **10**, 158.
101 P. Sakiewicz, K. Piotrowski, J. Ober and J. Karwot, *Renew. Sustain. Energy Rev.*, 2020, **124**, 109784.
102 S. Il Yu, C. Rhee, K. Hwa Cho and S. G. Shin, *Environ. Eng. Res.*, 2022, **28**, 220037.
103 P. Kazemi, C. Bengoa, J. P. Steyer and J. Giralt, *Process Saf. Environ. Prot.*, 2021, **146**, 905–915.
104 L. A. Putra, M. Köstler, M. Grundwürmer, L. Li, B. Huber and M. Gaderer, *Appl. Energy*, 2025, **377**, 124447.
105 P. Kazemi, J. P. Steyer, C. Bengoa, J. Font and J. Giralt, *Processes*, 2020, **8**, 67.
106 P. Ganeshan, A. Bose, J. Lee, S. Barathi and K. Rajendran, *Bioresour. Technol.*, 2024, **400**, 130665.
107 M. Khan, W. Chuenchart, K. C. Surendra and S. Kumar Khanal, *Bioresour. Technol.*, 2023, **370**, 128501.
108 M. Logan, M. Safi, P. Lens and C. Visvanathan, *Process Saf. Environ. Prot.*, 2019, **127**, 277–287.
109 S. Uludag-Demirer, M. Xu, A. Marks, Y. Liu, C. Saffron and W. Liao, *Biomass Bioenergy*, 2025, **198**, 107891.
110 N. Tao, M. Xu, X. Wu, Z. Pi, C. Yu, D. Fang and L. Zhou, *Fuel*, 2021, **299**, 120883.
111 M. Xu, S. Uludag-Demirer, D. Fang, L. Zhou, Y. Liu and W. Liao, *Energy Fuels*, 2021, **35**, 2282–2292.
112 M. Xu, S. Uludag-Demirer, Y. Liu and W. Liao, *Agronomy*, 2025, **15**, 305.
113 C. Song, Z. Zhang, X. Wang, X. Hu, C. Chen and G. Liu, *Waste Manage.*, 2024, **187**, 235–243.
114 H. Şenol, *Energy*, 2021, **215**, 119173.
115 M. M. Ali, M. Ndongo, K. Yetilmezsoy, M. Bahramian, B. Bilal, I. Youm and B. İ. Goncaloğlu, *J. Mater. Cycles Waste Manag.*, 2021, **23**, 301–314.
116 M. Saghouri, R. Abdi, M. Ebrahimi-Nik, A. Rohani and M. Maysami, *Energy Sources, Part A*, 2024, **46**, 8564–8580.
117 V. W. G. Tan, Y. J. Chan, S. K. Arumugasamy and J. W. Lim, *J. Cleaner Prod.*, 2023, **414**, 137575.
118 A. Reza and L. Chen, *Sci. Total Environ.*, 2022, **851**, 158321.
119 V. Hassija, V. Chamola, A. Mahapatra, A. Singal, D. Goel, K. Huang, S. Scardapane, I. Spinelli, M. Mahmud and A. Hussain, *Cog. Comput.*, 2024, **16**, 45–74.
120 F. Piadeh, I. Offie, K. Behzadian, J. P. Rizzuto, A. Bywater, J. R. Córdoba-Pachón and M. Walker, *J. Environ. Manage.*, 2024, **349**, 119458.