Review of Explainable Machine Learning for Anaerobic Digestion

Rohit Gupta[a, b, c, #], Le Zhang[d, #], Jiayi Hou[e], Zhikai Zhang[f,g], Hongtao Liu[e], Siming You[a], Yong Sik Ok[h], Wangliang Li[f,*]

a James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK

b Nanoengineered Systems Laboratory, UCL Mechanical Engineering, University College London, London WC1E 7JE, UK

c Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University College London, London W1W 7TS, UK

d Department of Resources and Environment, School of Agriculture and Biology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

e Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

f CAS Key Laboratory of Green Process and Engineering, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China

g School of Water Resources and Environment, Hebei GEO University, Shijiazhuang 050031, Hebei, China

h Korea Biochar Research Center, APRU Sustainable Waste Management Program & Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, South Korea

# The first two authors contributed equally to this work.

*Corresponding author. Wangliang Li, wlli@ipe.ac.cn

Submitted to ***Bioresource Technology*** (November 2022)

Abstract

Anaerobic digestion (AD) is a promising technology for recovering value-added resources from organic waste, thus achieving sustainable waste management. The performance of AD is dictated by a variety of factors including system design and operating conditions. This necessitates developing suitable modelling and optimization tools to quantify its off-design performance, where the application of machine learning (ML) and soft computing approaches have received increasing attention. Here, we succinctly reviewed the latest progress in black-box ML approaches for AD modelling with a thrust on global and local model interpretability metrics (e.g., Shapley values, partial dependence analysis, permutation feature importance). Categorical applications of the ML and soft computing approaches such as *what-if* scenario analysis, fault detection in AD systems, long-term operation prediction, and integration of ML with life cycle assessment are discussed. Finally, the research gaps and scopes for future work are summarized.

Research Highlights

- Popularly used ML-based AD models are ANN, SVM, RF, and XGBOOST
- Predicted variables are biogas yield, process stability, and effluent characteristics
- Global and local model-agnostic explainability approaches are reviewed
- Potential applications are process parameter optimization, fault detection, and LCA
- It is necessary to inform ML models with biokinetic equations to improve accuracy

Keywords: Data-driven Modelling; Sustainable waste management; Renewable energy; Bioenergy; Artificial intelligence

1. Introduction

Due to the increasing industrialization worldwide, the CO_2 concentration in the atmosphere has monotonically increased, posing great threats to Earth's ecosystem and human society. A recent report reveals that about 53% of the annual global CO_2 emissions are caused by China (26.4%), United States of America (12.5%), India (7.06%), and European Union (7.03%) (Ge et al., 2020). These concerning statistics necessitates incorporation of negative CO_2 emission technologies via circular bioresource utilization, among which anaerobic digestion (AD) of organic waste is a promising alternative.

AD is a multi-step, multi-physics, biokinetic degradation process comprising of four stages: hydrolysis, acidogenesis, acetogenesis, and methanogenesis. Initiated by the hydrolysis-based breakdown of organic matter in feedstock, the AD process generates several intermediates (e.g., volatile fatty acid (VFA), H_2, CO_2, acetate, etc.), which are finally transformed to biogas in the methanogenesis stage. Since the AD process involves a series of biochemical and physical processes, its efficiency and stability are influenced by various parameters including temperature, pH, moisture content, chemical oxygen demand (COD), VFA, total ammonia nitrogen (TAN), carbon to nitrogen to phosphorous ratio (C:N:P), trace elements, and toxic substances (Ajayi-Banji & Rahman, 2022).

A wide variety of first principles (or mechanistic) models have been developed for the design and optimization of AD processes. The mechanistic models consider the conservation of mass and energy to predict the cumulative biogas yield and compositions. The most rigorous mechanistic AD model developed to date is the Anaerobic Digestion Model No. 1 (ADM1),

which can predict the time-dependent biogas production accurately (Batstone et al., 2002). However, the implementation of ADM1 model for real-time prediction and control purpose is highly impractical due to its computationally demanding characteristics. In addition, extensively performing ADM1 parameter calibration is infeasible for the full-scale industrial AD processes with varying feedstock contents (Emebu et al., 2022). To circumvent these drawbacks, machine learning (ML)-based models and soft computing techniques have emerged as an alternative method for AD process modelling (Cruz et al., 2022).

The drawbacks mitigated by ML-based AD process modelling when compared mechanistic models (e.g., ADM1) are: (a) shorter execution time, (b) not requiring the multi-disciplinary knowledge related to bio-kinetics, microbiome, heat/mass transfer, and (c) avoidance of model re-calibration if trained based on extensive datasets. A wide variety of regression and classification models such as neural network (NN), support vector machine (SVM), random forest (RF), k-nearest neighbours (KNN), gaussian process regression (GPR), and extreme gradient boosting (XGBOOST) have been developed to predict biogas yield, process stability parameters (e.g., VFA), effluent quality indicators (e.g., COD) (Cruz et al., 2022). Nevertheless, scepticism exists among researchers due to the black-box nature of the ML approaches.

Two different types of ML approaches have been developed: (a) black-box ML and (b) explainable ML, with the latter attempting to provide a deeper understanding of the functional dependence of the output variables on the input variables. It is important to note that the ML research community advocates the use of explainable (or interpretable) ML for all applications

(Rudin, 2019). Several recent works for AD process modelling have demonstrated the benefits as obtained from ML integrated with various explainability metrics such as feature importance assessment, partial dependence analysis, etc. (Choi et al., 2022; Cinar et al., 2022; Long et al., 2021; Wang et al., 2021a). The explainable ML models offer a better understanding of the representative physical processes (i.e., AD) than its black-box counterpart. However, there has been no systematic review that summarizes ML model explainability metrics and discusses the implications of these approaches for improving AD process modelling.

The present review provides a critical summary of ML models integrated with explainability approaches for AD process modelling, with a detailed discussion on various model explainability methods, process parameter optimization, and fault detection algorithms. Potential ML integration with life cycle assessment (LCA) tools is discussed. The associated challenges, opportunities, and research directions are also summarized to aid the development of explainable ML models for AD process modelling.

2. Machine learning models

2.1. *Machine learning models*

Black-box ML approach is a data-driven modelling technique, which is entirely empirical and does not include phenomenological information on AD. These models are of two types: (a) regression (e.g., NN, gaussian process regression (GPR), linear regression, logistic regression, ridge regression, lasso regression, polynomial regression, and Bayesian linear regression, etc.) and (b) classification (e.g., SVM, KNN, logistic regression, naive bayes (NB), etc.) (Asgari et

al., 2021a). A regression model predicts output variables (e.g., biogas yield) based on numerical (e.g., total solids (TS)) or categorical (e.g., reactor type) predictor variables of AD processes. In contrast, a classification model deployed for AD processes is used for faulty or anomalous state detection of the reactor such as process inhibition due to VFA accumulation. Table 1 provides a summary of the explainable/interpretable ML-based AD models. The explainability of the results obtained from black-box ML models are enabled by approaches such as correlation analysis, feature importance assessment, partial dependence analysis, etc. Some of the important ML models included this review are NN, SVM, GPR, Decision Tree (DT), and Ensembles. The abovementioned ML-based AD models have several advantages compared to physics-based models (e.g., ADM1) such as rapid implementation by open-source tools, high generalizability, uncertainty quantification capability, fast training, and short execution time, which make these methods favourable for real-world deployment (Cruz et al., 2022).

2.2. *Neural networks*

NNs are generalizable non-linear models that have the abilities for complicated data trend learning. The unit quantity of an NN is a neuron which comprises of weights, bias, and transfer function. A variety of NNs have been developed such as (a) feedforward neural network (FNN), (b) extreme learning machine (ELM) (c) recurrent neural network (RNN), and (d) convolutional neural network (CNN).

FNNs, popularly referred to as artificial neural networks (ANNs), is the simplest class of NN for predicting process output variables (e.g., cumulative biogas yield and methane (CH_4)

content) based on several input variables (e.g., TS, volatile solids (VS), organic loading rate (OLR), pH, and temperature) (Cruz et al., 2022; Yi-Fan et al., 2017). Essentially, the FNNs developed for AD processes are regression models that avoid *a priori* assumption of the functional dependence between output and inputs, which is a common drawback of the mechanistic models (e.g., Gaussian, Gompertz and multi-regression) (Emebu et al., 2022). FNNs comprise of an input layer, hidden layers, and an output layer, among which the optimal number of hidden layers and number of neurons within those layers are required to be determined by model training. Notably, the number of hidden layers or neurons should be trained to prevent overfitting or underfitting.

Optimal parameters and network structures for NN models are determined using backpropagation (BP) algorithm-based training coupled with optimization algorithms (e.g., grid search, random search, and Bayesian optimization) (Cruz et al., 2022). The BP is a supervised learning algorithm that uses gradient descent technique in which the network's weights are changed along with the negative of the performance function's gradient. Another advanced class of FNN i.e., ELM does not tune the weights and bias, but rather randomly assign projections with nonlinear transformation. The ELMs possess better generalization capability and have shorter training time than the FNNs trained via the BP method (Kazemi et al., 2021; Kazemi et al., 2020b).

Nevertheless, either of the classical FNN or advanced ELM are not suitable for predicting time-dependent process variables because that do not have feedback loops or memories (e.g., transient biogas yield). The RNNs are used for predicting the transient trend of process output

variables by adding feedback loops to the FNN structure. Inclusion of the feedback loop allows RNNs to learn data sequences. However, simple RNNs with only feedback loops suffer from the problem of vanishing gradients for which the accuracy of the model is compromised. This drawback has been circumvented by the RNN architecture with long-short term memory (LSTM), which is a popular choice for learning transient process variables (Asgari et al., 2021b). The RNNs also have similar components such as input, hidden, and output layers, and must be trained in such a way that both overfitting and underfitting are minimized.

When the elements in the input layer of an NN prohibitively increases (e.g., spatiotemporal velocity or concentration field resolved via computational fluid dynamics (CFD) simulation of AD reactor), implementation of FNNs or RNNs become extremely challenging (Arnau Notari, 2022). To address this drawback, 2D-CNN has been developed, which can extract fine-grained information from high-dimensional data space and map them onto a coarse-grained low dimensional data space, enabling essential feature extraction from the input dataset. Conventional CNNs add feature extraction modules with convolution layer and max pooling layer prior to an FNN. For transient high-dimensional datasets CNNs must be combined with RNNs (or LSTMs) to add both spatial and temporal aspects to the regression model (Asgari et al., 2021a). A modified version of CNN i.e., 1D-CNN can also be used for learning time-series sequence, where the single dimension is the time-dependent output variable (e.g., biogas yield) (McCormick & Villa, 2019).

The most popular application for FNN developed for AD processes is related to prediction of biogas yield, CH_4 yield, and CH_4 content (Li et al., 2022; Wang et al., 2021a). These outputs

are predicted based on various feedstock parameters (e.g., C/N ratio, VS/TS ratio, VS, TS, COD, TSS, and OLR), AD operational variables (e.g., HRT, SRT, pH, temperature, VFA, NH_3 content, and H_2S content), and genomic data (Long et al., 2021). In some of the cases, additive materials such as zero valent iron, and biochar are used to enhance the biogas production efficiency, in which the dosage of these materials are considered as additional inputs of the NN models (Li et al., 2022; Xu et al., 2021). However, the FNN models that predict the average biogas production cannot quantify the process dynamics (time dependency) and are not applicable for process control or anomaly detection. In this regard, researchers have focused on deploying transient NN techniques such as RNN, 1D-CNN, and LSTM. Apart from biogas yield, NNs have also been applied for prediction of effluent COD (Yi-Fan et al., 2017), accumulated VFA in reactor (Kazemi et al., 2021; Kazemi et al., 2020b), NH_3 content in the effluent (Alejo et al., 2018), or relative abundance of genes (Haffiez et al., 2022). Advanced FNNs, i.e., ELM have been applied to predict biogas production and showed superior performance (up to 6.4% accuracy improvement) than conventional FNNs with fixed weights and biases (Kazemi et al., 2021; Kazemi et al., 2020b; Yan et al., 2020). However, in several instances (Li et al., 2022; Long et al., 2021; Wang et al., 2021a), advanced ML algorithms such as SVM, DT, or ensembles have significantly outperformed the NN models due to their high-dimensional adaptive feature learning capability. This suggests that the choice of optimal data-driven model can largely vary based on the nature of the dataset for AD modelling, which requires significant standardization efforts in future.

2.3. *Support vector machine*

An SVM is a non-parametric, deterministic, supervised ML model, which is frequently used for solving classification and regression problems. The input features are mapped onto a high-dimensional data space with the aid of a non-linear kernel function, which then imposes an optimal hyperplane to differentiate between different subsets. Training a conventional SVM (C-SVM) is achieved by solving a quadratic optimization problem to find a hyperplane with the maximum margin from the closest datapoints in the multi-dimensional feature space (Asgari et al., 2021c). Selection of kernel function is essential as it helps to tune the predictive accuracy, where the kernel scale plays an important role. Least-square support vector machine (LS-SVM) being another class of SVM solves a set of linear equations for model training, mitigating the computational complexity of C-SVM.

Both C-SVM and LS-SVM have been extensively developed for modelling of AD process. In several studies, the predictive accuracies of C-SVMs were compared to other models (e.g., NN, DT). A C-SVM model was developed to predict the effluent composition of the two-stage AD process with poultry manure as a feedstock (Alejo et al., 2018). The work compared the accuracy of C-SVM with other predictive models based on FNN and stoichiometric analytical methods, where C-SVM showed superior accuracy. This accuracy improvement corresponds to the capability of mapping low-dimensional features into a high-dimensional feature space. In another instance (Kazemi et al., 2020b), a C-SVM model was constructed to predict the time-dependent concentration of VFA in an AD process. The input data for the model relied on the Benchmark Simulation Model No, 2 (BSM2) model that simulates an activated sludge-

based wastewater treatment plant (WWTP) integrated with the AD process. Five different data-driven models were compared among which four models (C-SVM, FNN, ELM, and genetic programming (GP)) achieved high $R^2 \approx 0.99$ with low RMSE values.

C-SVM models have been developed by assimilating many AD data from the literature for a similar set of microbial communities and reactor configurations (Wang et al., 2020). Specifically, four different data-driven models (i.e., RF, SVM, KNN, and generalized linear model network (GLMNET)) were compared. It was concluded that the GLMNET and KNN models outperformed the RF and C-SVM models. Another novel research effort showed that by adding a time series feedback loop to the C-SVM can decrease the RSME of CH_4 yield prediction by 45% (Park et al., 2021). This approach was further used in another work for a bio-electrochemical AD process, which compared five different models including RF, XGBOOST, C-SVM, RNN, and LSTM (Cheon et al., 2022). LSTM and RNN models showed improved performances than the C-SVM model due to their superior capability in predicting time-series data. Metagenomic information (i.e., the relative abundance of various bacterial families) has also been embedded in C-SVM models in parallel to feedstock information and AD operational parameters (Li et al., 2022; Long et al., 2021). The competitive significance of bacterial communities was described, results of which can help justify the augmentation of biogas production.

In a latest effort (Ge et al., 2022), a C-SVM model has been developed to learn the kinetic parameters of ADM1 using feedstock composition and reactor temperature as input parameters.

By using SVM-assisted parameter learning for ADM1, the predictive accuracy has been significantly improved.

2.4. *Gaussian process regression*

GPR is a non-linear, non-parametric, Bayesian probabilistic data-driven model for regression problems (Asgari et al., 2021c). In contrast to deterministic data-driven models, GPR predicts mean values of output parameters with prediction uncertainty bands. Therefore, both the mean and variance information of the state variables could be informed to the control system. In addition, GPR is data-efficient since they balance the trade-off between model fitting and model complexity by minimizing the log-marginal likelihood function. A generalized expression for predicting a target variable y based on input x via GPR reads $y = \beta f(x) + \sigma^2 Z(x, \omega)$, where $\beta f(x)$ is the mean of GP, σ^2 is the variance, $Z(x, \omega)$ is the stochastic process with zero mean and unit variance, β is the coefficient determined via regression, and $f(x)$ is the basis function. Given a predefined choice of $f(x)$ and $Z(x, \omega)$, the values of β and σ^2 are estimated based on the input and output datasets of AD experiments. One of the most essential aspects of developing a GPR is the choice of $Z(x, \omega)$, referred to in the literature as kernel or covariance functions. The kernel functions correlate training input data points x with new testing data points x'. Popular choices of the GPR kernel functions include exponential, squared exponential, matern 5/2, matern 3/2, and rational quadratic (Schulz et al., 2018).

There have not been many efforts of utilizing GPR for ML-based AD modelling. In a prior work (Južnič-Zonta et al., 2012), GPR was used to quantify the uncertainty associated with ADM1 model parameters and time-dependent composition in COD substrates. The original simulations were replaced with meta-models, which enabled stochastic quantification of the ADM1 model parameters. The model was used to statistically approximate the pareto frontier (or trade-off) between CH_4 production and the COD of waste substrate. Recently, GPR-based surrogate modelling was used for uncertainty quantification and global sensitivity analysis of the modified ADM1 that predicts CH_4 production and VFA accumulation in AD processes (Trucchia & Frunzo, 2021). Various model parameters were ranked based on their relative impact using the data-driven surrogate model (i.e., GPR). The performance of GPR-based surrogate modelling with FNN and the polynomial chaos expansion (PCE) (a method of expressing a random variable as a polynomial function of other random variables) were compared for a WWTP with an AD unit (i.e., BSM2) (Al et al., 2019). The results revealed that GPR-based global sensitivity analysis outperformed the FNN and PCE models in terms of training time since they required lower number of datapoints during training.

2.5. *Decision trees and ensembles*

Decision trees (DT)-based ML models are constructed by nodes (or leaves) and branches, which are routinely used for building regression-based process parameter prediction or classification-based conditional decision making (Alloghani et al., 2020). Recently, the usage of DT (and associated algorithms) has gained significant attention for AD process modelling and management. The complexity of DT models is characterized by various attributes such as

the depth of the tree and the number of branches from each leaf. It has been evidenced that the models with a single DT (yet deep) lead to inferior performance due to model simplicity and overfitting. DT ensembling approach addresses this drawback, where multiple DTs are constructed in parallel based on subsets, balancing the trade-offs between model bias (performance on the training dataset) and variance (performance on the testing dataset). There are two types of ensembling methods: bagging and boosting. The bagged ensembling randomly splits the training dataset into subsets to train multiple parallel models with separate outputs (Fawagreh et al., 2014). These outputs are further unified using statistical metrics such as mean, median, or mode. A classic example of a bagged ensembling DT is the RF model which offers collective intelligence. In contrast, the boosting process converts weak learners to strong learners by averaging, weighing, or voting multiple learners (Natekin & Knoll, 2013). The main difference between bagging and boosting is that in boosting the trees are built additively to improve the model performance, while in the bagging approach the trees are built and branched in parallel independently. Thus, the boosted tree-based algorithms are much faster than bagging algorithms. Popular boosting algorithms are XGBOOST, Gradient Boosting Machine (GBM), and Adaptive Boosting (AdaBoost), which avoid overfitting by setting constraints on tree size and performing tree pruning (i.e., removing unnecessary sub-nodes of a decision node). Some other controllable hyperparameters for ensemble methods are minimum leaf size, number of learners, and learning rate.

A large segment of the data-driven AD modelling literature has used ensembled DTs for predicting AD process variables. For example, RF and XGBOOST (De Clercq et al., 2020)

have been utilized to predict transient biomethane production for an industrial-scale anaerobic co-digestion (ACoD) system, where XGBOOST performed better (R^2 = 0.88) than the RF model (R^2 = 0.80). Another work showed that retraining these two models (RF and XGBOOST) using a one-step ahead time-series method can further improve the predictive accuracy (Cheon et al., 2022; Park et al., 2021). Further, an RF-based combined regression-classification model was developed and compared to GLMNET, C-SVM, and NN (Wang et al., 2020). The findings revealed that RF models had an inferior predictive performance for both classification and regression applications when compared to all other models. A research effort (Long et al., 2021) combined genomic data, waste parameters, and operational conditions of the AD process to compare five different ML models including RF and XGBOOST, where RF showed the best performance in CH_4 yield prediction. An automatic ML algorithm selection tool was developed (Wang et al., 2021a), which compared various DT algorithms (RF, XGBOOST, AdaBoost, and GBM), C-SVM, linear models (Ridge, ElasticNet, LassoLars), and KNN for predicting the ACoD of organic waste. The ExtraTree algorithm provided the highest accuracy (R^2 = 0.72) for the testing dataset. XGBOOST algorithm also proved to be superior for predicting biogas yield for AD processes enhanced with foreign material such as zero-valent iron (Xu et al., 2021) or biochar (Li et al., 2022).

DT-based ML models have also been developed to predict critical parameters in AD process other than biogas yield, CH_4 yield, or CH_4 composition in biogas. For example, prior work used an ensemble approach to predict the transient VFA accumulation in AD reactors which is highly detrimental to biogas production. These models are either regression-based for

process variables control (Kazemi et al., 2020b) or classification-based for fault detection (Kazemi et al., 2021). In another work, the relative abundance of antibiotic-resistant genes or mobile genetic elements was predicted using RF, XGBOOST, and FNN models (Haffiez et al., 2022). However, the FNN model showed superior predictive performance (R^2 = 0.77) than both the ensemble-based DT models (RF and XGBOOST).

3. Accuracy metrics

The accuracy of the ML models discussed in Section 2 is positively correlated with its prediction capability and reliability of finding target solutions. A wide variety of statistical metrics have been utilized to evaluate the predictive accuracy of the ML models. For regression problems, some popular choices of metrics are (a) coefficient of determination (R^2), (b) root mean square error (RMSE), (c) mean absolute deviation (MAD), (d) mean absolute scaled error (MASE), (e) mean absolute error (MAE), (f) mean squared error and (g) mean absolute percentage error (MAPE). In case of classification problem, the accuracy is routinely visualized by confusion matrix and evaluated using metrics such as precision, recall, and F1-score (Jeong et al., 2021; Li et al., 2022). Two of the most widely used metrics for regression problems are RMSE and R^2, which quantify the average error of the model and the deviation from the parity line, respectively. A low value of RMSE is desirable, while an $R^2 = 1$ indicates optimal model fitting. In ML pipelines, several models are evaluated serially following which the optimal model is identified based on the abovementioned accuracy metrics. For classification models, the routinely used F1-score is the harmonic mean of precision and recall which further depend on correctly and falsely classified scenarios (De Clercq et al., 2020; Wang et al., 2020).

4. Model explainability methods

The scepticism towards ML models being entirely black-box approaches have resulted in significant research efforts towards enhancing the explainability/interpretability of the results predicted by these models. Some of the models such as GLMNET, DT, KNN, and NB are inbuilt with interpretability methods, while other models require additional integration with model-agnostic interpretability methods. The model-agnostic interpretability essentially determines the functional dependence between the output and input parameter space and is therefore independent of the choice of the ML model. Furthermore, there are global and local model-agnostic methods which explain the average and individual predictions, respectively, for the ML models. The outcome of the model explainability assessment for various research efforts in the AD literature is summarized in Table 2.

4.1. *Global model-agnostic explainability methods*

Some of the global explainability metrics are partial dependence plots (PDP), accumulated local effects (ALE), permutation feature importance, and global surrogate model. The PDP describes the marginal effect of one or two input features on the ML model outcome, where the functional dependence can be either linear or non-linear. The ALE plots are more advanced version of PDPs, which offer advantages such as lower computation time and removal of bias. The permutation feature importance is another popular method, which compares the baseline ML model predictions to a range of model predictions by shuffling individual values of an input parameter. This process is repeated for each of the input parameters, which results in a

global permutation feature importance plot. The global surrogate model approach builds an interpretability-enabled model (e.g., GLMNET, DT, KNN, NB) to explain the predictions of a non-interpretable model. A high-level investigation of Table 1 reveals that only two of the research works adopted one-way and two-way PDP analysis (De Clercq et al., 2020; Wang et al., 2021a), while quantification of permutation feature importance was more frequent (Choi et al., 2022; Cinar et al., 2022; De Clercq et al., 2020; Li et al., 2022; Wang et al., 2021a). A summary of global-model agnostic ML model explainability analysis is provided in Table 2, which reveals the functional trend of PDP plots and the most influent factors that regulate model outputs.

4.2. *Local model-agnostic explainbility methods*

Among the local model-agnostic explainability methods the Shapley additive explanation (SHAP) is a popular choice for ML model development. The SHAP is a method to explain individual predictions by an ML model and is based on the game theory-based Shapley values. The SHAP explains the contribution of each entry of an input parameter of the ML model by computing Shapley values based on coalitional game theory. The analogy between game theory and ML model is that an input feature in ML model is similar to a player in the game theory. The distribution of Shapley values provides the knowledge about how each entry of the input feature are correlated to the output. The absolute SHAP values can be further averaged to develop a global model-agnostic explainability method. This provides an alternative means to the permutation feature importance method for evaluating global feature importance. In contrast to global model-agnostic metrics, local model-agnostic metric (e.g., SHAP) are more

frequently used to explain the AD modelling results. All the research works shown in Table 1 adopted at least one of the feature importance analyses. Some of the exceptions to SHAP-based feature ranking are (a) based on Garson's method (Ghatak & Ghatak, 2018), (b) MeanDecreaseGini, (c) IncNodePurity (Long et al., 2021; Wang et al., 2020), and (d) fscaret (Kazemi et al., 2020b). A summary of the local model-agnostic explainability approaches is provided in Table 2 which reveals the most influential features that regulate output variables.

5. Applications of machine-learned models in anaerobic digestion

5.1. *Integration with optimization algorithms*

The most important application of ML-based model for AD is to understand the interplay between the input variables and output variables for a wide range of parametric scenarios, thus enabling the identification of the optimal range of different process variables. Thus, to obtain optimal values for process output variables, an ML-based predictive model pipeline can be coupled with popular heuristic optimization solvers such as genetic algorithm (GA), particle swarm optimization (PSO), simulated annealing (SA), or ant colony optimization (ACO). These optimizers can be either single- or multi-objective, based on which a pareto frontier of process output variables (e.g., biogas yield, biogas composition, and effluent characteristics) are generated. The results of such ML-coupled multi-objective optimization problems form essential look-up tables for an AD process management team.

Several important parameters that regulate the performance of AD processes are TS, VS, OLR, pH, reactor temperature, oxidation reduction potential (ORP), electrical conductivity,

alkalinity (ALK), TAN, VFA, ACoD parameters, number of reactor stages, reactor volume, scale of implementation, feedstock pre-treatment parameters, C/N ratio, C/P ratio, solid retention time (SRT), hydraulic retention time (HRT), etc. Since the ML-based algorithms do not embed any physical laws such as conservations of mass, energy, momentum, and species, the accuracy and robustness of these algorithms are largely affected by the data collected. For example, earlier ML-based AD models show high accuracy ($R^2 > 0.9$), although they are applicable to very limited number of feedstocks (Dahunsi et al., 2017; Jacob & Banerjee, 2016; Zareei & Khodaei, 2017). Therefore, the quantification of the accuracy of a model cannot provide a fair judgment of the generalizability and interpretability of these models. Moreover, integration of these models with optimization solvers cannot provide a robust *what-if* scenario analysis for a wide range of feedstock and reactor operating conditions. These drawbacks can be circumvented by generalized ML pipelines integrated with the explainability methods.

Optimization algorithms can be combined with ML models as either a pre-processor or a post-processor. In the pre-processing applications, the optimization solvers (e.g., GA and ACO) have been used to select the most influential process variables for developing an FNN model that predicted biogas yields (Beltramo et al., 2019). The work showed that the addition of GA- or ACO-based feature selection to the FNN model reduced the dimensionality of the problem by eliminating superfluous features. This resulted in a reduction of model overfitting and improved the accuracy of FNN by 6.2%. In other instances, PSO- or GA-based optimization algorithms have been used downstream to ML models (mainly FNN) for maximizing the yield of biogas or CH_4 produced by AD plants (Alrawashdeh et al., 2022; Asadi

& McPhedran, 2021; Zaied et al., 2020). Integration of FNN with a GA-based multi-objective optimization framework was also attempted to determine the pareto frontier (trade-off diagram) between biogas production and effluent COD, which revealed that maximizing the first variable inevitably minimized the latter (Huang et al., 2016). In future, ML-integrated optimization problems must be coupled with model explainability methods such as PDP and feature importance to justify the correlation between pareto frontier and multi-dimensional input space.

5.2. *Machine learning models as soft sensors*

Since AD processes involve several time-dependent output variables, obtaining precise control on these can facilitate the improvement of process efficiency and stability. In this realm, soft sensors are beneficial due to their capacity of state estimation for control application, anomaly detection, and fault identification. A few researchers have reviewed the development of dynamic soft sensor for AD processes that enhanced supervisory control actions (Cruz et al., 2021; Yan et al., 2021). Conventional methods for developing AD soft sensors are based on Kalman filters, dynamic principal component (PCA) analysis, and recursive partial least square (PLS). However, a number of works showed that ML-based time-series prediction approaches could significantly improve the accuracy, thus reducing the fluctuation of a state variable (e.g., biogas yield, VFA, COD_{eff}, etc.) as detailed subsequently.

For example, a group of literature (Kazemi et al., 2021; Kazemi et al., 2020a; Kazemi et al., 2020b) developed ML-based model-predictive control (MPC) algorithms using RF, FNN, ELM, C-SVM, and GP. The VFA predictions from these models were further coupled to

statistical control charts such as squared prediction error (SPE) and cumulative sum (CUSUM) charts to determine various anomalous events (or faults) occurring in AD reactors. These methods were compared to PCA-based fault detection, which showed up to 86.2% improvement in F1-score for various fault classification. Furthermore, the ML-based models included fscaret feature ranking method-based feature importance analysis. This in turn enhanced the explainability of the data-driven models, reduced inclusion of superfluous feature, and resulted in better model generalizability. Nevertheless, the studies lacked SHAP and PDP analysis which can provide insightful information on the granular feature importance and variational trends of output variables as the functions of input variables.

Another group of researchers (Wang & Li, 2019; Wang & Wang, 2021; Wang et al., 2021b; Yan et al., 2020) developed more advanced ML-based control algorithms to predict VFA accumulation for AD of kitchen waste. These approaches included PLS, BP-FNN, C-SVM, deep belief network (DBN), ELM, hierarchal ELM, stacked auto-encoder (SAE), SAE-ELM, CNN, graph convolutional network (GCN), gated recurrent unit (GRU), and spatiotemporal GCN (STGCN). Some of the abovementioned methods offered feature reduction of the input dataset using maximal information coefficient (MIC), minimum redundancy maximum correlation (mRMR), or fast filter-based correlation (FCBF), resulting in a higher prediction accuracy of the models. As an example, feature reduction-enabled SAE-ELM algorithm reduced the number of input variables from 9 to 4, while enhancing the accuracy by 6.4% compared to BP-FNN (Wang & Wang, 2021). These promising results show that integration

of feature reduction and model-agnostic interpretability methods are effective for the construction of data-driven AD process models.

5.3. *Machine learning models for long term operation*

Long-term AD operation includes a start-up period and a stable operation process, the study of which can support the development of effective start-up strategies, AD performances, reliable data to describe system conditions, stability control methods, and microbial community information. In contrast to the previously discussed works in Section 5.1 which are suitable for *what-if* scenario and parametric optimization, deploying ML algorithms for supporting long-term AD operation requires extensive training dataset and sophisticated time series techniques (with feedback loops), a few of which are similar to those used in AD control applications (see Section 5.2).

Combining model interpretability methods with time-series AD modelling techniques offers significant benefits such as *a priori* prediction of biogas yield or VFA accumulation based on historical information, or dynamical tuning of (or re-train) model parameters based on changes in input datasets. Researchers have shown that the use of static feature importance analysis for ML model based on which dimensionality reduction is performed can lead to exclusion of important features in transient scenarios. For example, a recent study compared the feature importance maps of 4 days and 40 days from the AD start-up point, and showed that the dominant features were drastically different (De Clercq et al., 2020). Therefore, for

such dynamic process models, it is instructive to investigate time-dependent change of SHAP, permutation importance, and PDP.

In another instance, popular time series forecasting techniques such as LSTM were further improved by integrating with dual-stage attention (DA) and variable selection network (VSN) (Jeong et al., 2021). This hybrid model exhibited up to 36% relative accuracy improvement when compared to an LSTM for predicting two-year biogas generation for ACoD. The study enabled model interpretability for continuous and discontinuous datasets, revealing the essential feature that regulated biogas yields. Subsequent research efforts on ML-based AD model indicated that 1-step ahead retaining method could further improve accuracy, especially for the OLR transition periods during long-time operation (Cheon et al., 2022; Park et al., 2021). Moreover, using the 1-step ahead method, pH could be only used as the input parameter for real-time prediction of CH_4 yield, which indicated the promising potential towards reducing model training and execution time.

In a latest study, a tree-based ML pipeline optimization tool (TPOT), was developed to simulate the impact of organic waste and operating parameters on biogas yields using 8 years data of an industrial-scale WWTP (Wang et al., 2021a). The data included daily input of 31 waste stream compositions (such as brine, dairy, fats, oils, greases, primary sludge, thickened waste activated sludge, etc) and 5 operating parameters. The robust predictive power of TPOT for ACoD process modelling showed superior predictive performance when compared to a FNN model. The combination of SHAP, permutation feature importance, and PDP analyses showed the functional dependence of the most significant parameters towards regulating biogas

yields and CH_4 contents. The TPOT method was further used to investigate the decomposition mechanisms of different wastes streams within the digester. The results showed that the predictive models were powerful tools for supporting efficient operation in long-term scenarios, understanding microbial dynamics, and balance the interplay of operating parameters.

5.4. *Life cycle assessment informed by machine-learning models*

A potential application of explainable ML models is in facilitating LCA for AD systems, which improves the quantification of the environmental impact abatement potential offered by biogas and digestate utilization. As per ISO, an LCA framework constitutes four phases: (a) setting goal and scope (system boundary, functional unit, etc), (b) process model development and generation of life cycle inventory (LCI), (c) life cycle impact assessment (LCIA), and (d) interpretation of the results (Gupta et al., 2022). During the second phase of LCA (i.e., LCI development), data from relevant literature are collected, which raises several questions towards the generalizability of the LCA framework.

First, the LCI data used in the AD literature are about a limited number of scenarios based on the input waste type, ACoD strategy, operation mode, reactor design, and operation parameters, ultimately limiting the usage and generalization of LCA results obtained. Second, the LCI data are often subject to uncertainties that are difficult to quantify, therefore adversely affecting the confidence of policymakers in trusting the LCA results. Third, even though intricate physics-based models (e.g., ADM1) can ensure better generalizability of LCA

framework (Meneses et al., 2015), they require extensive model calibration and require larger computation time, making them less attractive for whole-system modelling.

LCA framework integrated with ML models for AD systems can circumvent these drawbacks. For example, explainable ML models based on extensive data collection (waste compositions, operating modes, microbiome abundance data and operating condition) can ensure better generalizability of a unit process in LCA and allow the prediction of biogas/CH_4 yield for a wide range of scenarios (De Clercq et al., 2020; Li et al., 2022; Long et al., 2021; Zhao et al., 2021). Probabilistic ML models such as GPR enable the quantification of both the mean and the uncertainty associated with predicted variables (e.g., biogas yield, CH_4 yield, VFA, etc.), which can facilitate the evaluation of the uncertainty propagation towards calculating the LCIA metrics (e.g., global warming potential) (Al et al., 2019; Južnič-Zonta et al., 2012; Trucchia & Frunzo, 2021). Due to lower computation time required by ML model-based predictions, they serve as a promising alternative to the models involving solving a large number of differential equations (e.g., ADM1). Although there have been rare attempts of integrating LCA with ML models to date, the discussion indicates the promising potential of developing a holistic whole-system model for the environmental accounting of AD development in future.

6. Challenges and perspectives

Despite the significant efforts toward developing ML models for AD processes in recent years, the relevant development is yet in its initial stage. Majority of the prior works treat ML

modelling of AD as a “black-box approach” with limited (or zero) physical understanding of process phenomena, which poses several challenges.

- First, black-box models are mostly on basis of experimental data from prototypical lab-scale or pilot-scale reactors, leading to limited generalizability for industrial scale development. In other words, it is potentially problematic to make extrapolation of these models for predicting a full-scale system.
- Second, except for a few recent works, most of the ML modelling applied to AD processes lacks the presentation of SHAP, permutation feature importance, and partial dependence quantification. These metrics are extremely important for understanding the correlation and variational relationship between predictors (input) and predicted (output) variables. When several types of ML algorithms are to be compared for automatic optimal algorithm selection, most of the prior works did in terms of their predictive accuracies. However, model explanatory metrics which can resolve the reasons behind the performance improvement are mostly overlooked. In future, it will be valuable to deploy and examine the other model model-agnostic explainability approaches such as functional decompositions, counterfactual explanations, and scoped rules (or anchors).
- Third, most of the ML models for AD processes are either based on metagenomics data or operational parameters. Nevertheless, examples of unifying metagenomic data and operation parameters are rare, for which the generalizability of the model is compromised. Moreover, despite the promising potential of GPRs for model uncertainty quantification,

they have not been extensively used in ML modelling of AD processes, which is an essential scope for future work.

- Additionally, efforts to integrate between existing mechanistic AD models and the data-driven models are limited, which has led to unrealistic predictions by some of the data-driven models. This drawback can be mitigated in two ways: (a) using ML models to learn parameters of the mechanistic models from experimental data or (b) embedding the residuals of the ADM1 differential equations in the loss function of the ML models.
- Finally, coupling the vector and scalar fields from CFD simulation of AD reactor can improve the predictive performance of the ML models. Significant efforts are required toward fast and robust control models (or soft sensors) for enabling real-time model predictive control of AD reactors. In this realm, exploration of the feature importance of a data-driven model can eliminate unimportant features and reduce the execution time of online model training (for re-enforcement learning) and evaluation.

7. Conclusions

This review comprehensively summarized the state-of-the-art of black-box ML approaches integrated with model interpretability methods for AD processes. Applications of these models included process optimization and control, *what-if* scenario investigation, process fault identification, carbon footprint assessment, and kinetic parameter learning for ADM1. In addition, model-agnostic explainability metrics that describe the correlations between predicted variables and input features are critically discussed for AD process modelling.

Despite the progress in ML modelling of AD processes, this field is still in its early stage due to insufficient (or unstandardised) data, and lack of consistent principles for model selection.

Declaration of Competing 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.

Acknowledgments

Wangliang Li would like to thank the financial support from the National Natural Science Foundation of China (No. 21878313). Siming You would like to acknowledge the financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) Programme Grant (EP/V030515/1), Supergen Bioenergy Hub Rapid Response Funding (RR 2022_10), and Royal Society Research Grant (RGS\R1\211358). Rohit Gupta gratefully acknowledges the Royal Society Newton International Fellowship (NIF\R1\211013). Yong Sik Ok acknowledges the support of the Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01475801) from Rural Development Administration, the Republic of Korea. This work was also partly supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A6A1A10045235).

Author Contributions

Rohit Gupta: Conceptualization, Methodology, Data Curation, Writing – Original Draft, Writing – Review & Editing; **Le Zhang**: Writing – Original Draft, Methodology, Writing –

Review & Editing; **Jiayi Hou**: Writing – Original Draft; **Zhikai Zhang**: Writing – Original Draft; **Hongtao Liu**: Writing – Original Draft, Supervision; **Siming You**: Writing – Review & Editing, Funding Acquisition, Supervision; **Yong Sik Ok**: Writing – Review & Editing, Supervision; **Wangliang Li**: Conceptualization, Writing – Original Draft, Writing – Review & Editing, Funding Acquisition, Supervision.

Table 1. Summary of research works related to ML-based AD integrated with model explainability assessment.

Reference	**Models explored**	**Optimal Model**	**Inputs**	**Outputs**	**R^2**	**RMSE**	**Exaplainability Methods**
(Yi-Fan et al., 2017)	FNN	---	COD_{inf}, HRT, ALK, pH, VFA, ORP	COD_{eff}	0.87	0.38 mg/l	Statistical feature ranking using correlation analysis
(Alejo et al., 2018)	C-SVM, FNN	C-SVM	TAN, VS, COD_{inf}, TS	TAN_{eff}	0.90	0.31 gTAN/l	Statistical feature ranking using correlation analysis
(Ghatak & Ghatak, 2018)	FNN	---	Feedstock composition, HRT, temperature	Biogas yield	0.99	43 ml/gVS	Feature importance using Garson metric
(Oloko-Oba et al., 2018)	FNN with GA	---	Reactor shape, feedstock composition	Biogas yield	0.85	---	Feature importance assessment
(De Clercq et al., 2020)	RF, XGBOOST, ElasticNet	RF	Feedstock composition	CH_4 yield	0.80	---	Permutation feature importance, one-way and two-way PDPs

(Kazemi et al., 2020b)	RF, FNN, ELM, C-SVM, GP	GP	pH, TAN, pressure, TS, COD, ALK, gas flow, mole fractions of CH_4, H_2 and CO_2	VFA	0.99	---	Statistical feature ranking using fscaret
(Wang et al., 2020)	RF, GLMNET, C-SVM, KNN	KNN	Feedstock composition and temperature	CH_4 yield	0.73	26.7 ml/(l.day)	MeanDecreaseGini and IncNodePurity feature importance
(Jeong et al., 2021)	LSTM, DA-LSTM, DA-LSTM-VSN	DA-LSTM-VSN	Sludge inflow and outflow, temperature, SRT. VS/TS, BOD, COD, SS, TN, TP				
(Long et al., 2021)	GLMNET, RF, XGBOOST FNN, KNN, C-SVM	RF	Feedstock composition, operational conditions, and genomic data.	CH_4 yield	0.82	40 ml/gVS	MeanDecreaseGini feature importance

(Park et al., 2021)	XGBOOST, C-SVM, RNN, RF	RNN	pH, alkalinity, COD removal efficiency, VFA	CH_4 yield	0.97	23 ml/gCOD	Principal component analysis, PDP
(Wang et al., 2021a)	DT, AdaBoost, XGBOOST, RF, ExtraTrees, GBM, Ridge, ElasticNet, LassoLars, LS-SVM, SGD, and KNN	ExtraTrees	Waste property, TS, VS, VFA, ALK	CH_4 yield	0.72	247 scfm	Permutation feature importance, one-way PDP
(Xu et al., 2021)	RF, XGBOOST, FNN	XGBOOST	TS, VS, COD, ZVI concentration and size, pH, temperature	CH_4 yield	---	21 ml/gVS	Surrogate-based feature importance analysis
(Cheon et al., 2022)	RF, XGBOOST, C-SVM, LSTM, RNN	RNN	OLR, pH, alkalinity, VFA, and COD removal efficiency	CH_4 yield	0.97	20 ml/gCOD	Principal component analysis, PDP

(Cinar et al., 2022)	Linear regression, GLMNET, KNN, C-SVM, DT, RF, XGBOOST	C-SVM	Temperature, pressure, feed, volume, and nutrient solution usage	CH_4 yield	0.85	---	Permutation feature importance
(Haffiez et al., 2022)	RF, XGBOOST, FNN	FNN	Operating mode, feedstock pre-treatment, additives, temperature, and HRT	Relative abundance of ARG and MGE	0.79	---	SHAP feature importance
(Li et al., 2022)	KNN, RF, GBM, C-SVM, FNN	GBM with C-SVM	C/N, VS/TS, HRT, OLR, pH, biochar dosage, temperature, and COD	CH_4 yield and percentage	0.84	68.04 ml/gVS and 5.84%	SHAP feature importance, permutation feature importance, correlation coefficient analysis
(Choi et al., 2022)	C-SVM, RF, AdaBoost, XGBOOST	XGBOOST	Feedstock composition, VS, TS, HRT, pH, ALK, COD	VFA	0.64	---	SHAP feature importance, permutation feature importance

Abbreviations- FNN: feedforward neural network, COD: chemical oxygen demand, HRT: hydraulic retention time, ALK: alkalinity, VFA: volatile fatty acid, ORP: oxidation reduction potential, TAN: total ammonium nitrogen, VS: volatile solid, TS: total solids, RF: random forest, XGBOOST: extreme gradient boosting, GP: genetic programming, C-SVM: conventional support vector machine, ELM: extreme learning machine, KNN: k-nearest neighbours, GLMNET: generalized linear model network, PDP: partial dependence plot, DT: decision tree, GBM: gradient boosting machine, SGD: stochastic gradient descent, LS-SVM: least square support vector machine, LSTM: long-short term memory, ZVI: zero-valent iron, AdaBoost: adaptive gradient boosting, SHAP: shapley additive explanations, OLR: organic loading rate, ARG: antibiotic resistant genes, MGE: mobile genetic elements, BOD: biological oxygen demand, SRT: solid retention time, SS: suspended solids, TN: total nitrogen, TP: total phosphorus.

Table 2. Results of various ML model explainabiliy-based research works for AD systems. This table is to be interpreted in accordance with Table 1.

Reference	**Outcome of model explainability assessment**
(Yi-Fan et al., 2017)	• Feature importance analysis revealed that VFA and COD_{inf} had 50.4% and 24.7% for predicting COD_{eff} using FNN.
(Alejo et al., 2018)	• Feature importance analysis revealed that TAN_{inf} and VS_{inf} were top two features for predicting TAN_{eff}.
(Ghatak & Ghatak, 2018)	• Garson's metric-based feature importance analysis suggested that waste composition, HRT and temperature had 78%, 17%, and 5%, respectively for predicting cumulative biogas yield.
(Oloko-Oba et al., 2018)	• Feature importance analysis across different types of input wastes indicated that cow dung, piggery waste, poultry dropping, and plantain peels had 45%, 25%, 15% and 10% contributions, respectively, for predicting biogas yield.
(De Clercq et al., 2020)	• XGBOOST model was dependent on a higher number of features than the RF model. • The permutation feature importance distribution for each model was also time dependent. • Top five features for RF model are: total waste input, percolate input, co-digested amount, food waste input, and cassava input, while those for XGBOOST are: food waste input, bagasse input, cassava input, total waste input, and pig manure input.

	• One-way PDP analysis shows and exponential increasing trend of biomethane output upon increasing the important input features. In addition, two-way PDP also suggested positive interaction effects.
(Kazemi et al., 2020b)	• As per fscaret-based feature ranking the top five features for regulating VFA are: pH, TAN, pressure, CO_2 mole fraction, and TS_{eff}.
(Wang et al., 2020)	• Top five features for RF-based regression model are: total carbon, TAN, lignin content, C/N ratio, and xylan content. • For RF-based classification model, the top five features are: temperature, glucan content, cellulose content, total carbon, and C/N ratio.
(Jeong et al., 2021)	• For continuous time series of input variables, the influential feature are sludge loading, HRT, and temperature. • For discontinuous time series of input variables, SS, VS/TS, BOD, TN, and TP.
(Long et al., 2021)	• Feature importance analysis for RF model revealed that genomic abundance data influences the CH_4 to a larger extent than the process operating condition and feedstock compositions. • The top five features are abundances of proteobacteria, chloroflexi, fibrobacteres, actinobacteria, and spirochaete.

(Wang et al., 2021a)	• Top five important features for the tree-based automatic ML pipeline are: general waste content with COD > 20000 mg/l, dairy waste content, content of fat, oil, and gas (FOG), content of rendering waste, and amount of poultry blood. • PDP shows linear and exponential increasing trends of CH_4 yield when the most influential parameters are altered.
(Xu et al., 2021)	• The top five parameters obtained from feature importance that regulate CH_4 yield are: TS_{eff}, sCOD, ZVI dosage, ZVI particle size, and TS_{inf}.
(Cinar et al., 2022)	• The top five important feature for KNN model obtained via permutation feature importance are biogas temperature, reactor temperature, nutrient solution usage, reactor pressure, and OLR.
(Haffiez et al., 2022)	• For predicting the abundance of ARG and MGE, the top five SHAP features are: feedstock type, mode of operation, HRT, adoption of feedstock pre-treatment, and temperature.
(Li et al., 2022)	• As per the SHAP-based feature importance analysis, the top five feature for regulating CH_4 yield and content are: sCOD, VS/TS, OLR, biochar dosage, and temperature. The 4^{th} and 5^{th} important features changed to pH and C/N, when the permutation importance was utilized.
(Choi et al., 2022)	• As per SHAP-based feature importance, the top five important parameters for predicting VFA are: ALK, content of FW, TKN of FW, SO_4 of FW, and lipid content in FW.

Abbreviations - VFA: volatile fatty acid, COD: chemical oxygen demand, TAN: total ammonium nitrogen, VS: volatile solid, XGBOOST: extreme gradient boosting, HRT: hydraulic retention time, RF: random forest, PDP: partial dependence plot, TS: total solids, C/N ratio: carbon to nitrogen ratio, ALK: alkalinity ZVI: zero-valent iron, ARG: antibiotic resistant gene, MBE, mobile genetic element, OLR: organic loading rate, SHAP: shapley additive explanations, FW: food waste, TKN: total Kjeldahl nitrogen, sCOD: soluble chemical oxygen demand, KNN: k-nearest neighbour.

References

1. Ajayi-Banji, A., Rahman, S. 2022. A review of process parameters influence in solid-state anaerobic digestion: Focus on performance stability thresholds. *Renewable and Sustainable Energy Reviews*, **167**, 112756.

2. Al, R., Behera, C.R., Zubov, A., Gernaey, K.V., Sin, G. 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants–An application to the BSM2 model. *Computers & Chemical Engineering*, **127**, 233-246.

3. Alejo, L., Atkinson, J., Guzmán-Fierro, V., Roeckel, M. 2018. Effluent composition prediction of a two-stage anaerobic digestion process: machine learning and stoichiometry techniques. *Environmental Science and Pollution Research*, **25**(21), 21149-21163.

4. Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., Aljaaf, A.J. 2020. A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised and unsupervised learning for data science*, 3-21.

5. Alrawashdeh, K.A.B., Al-Samrraie, L., Al Issa, H.A., Qasem, I., Hussien, A., Al-Zboon, K. 2022. Prediction and optimization of biogas production from OMW digestion using fenton pre-treatment process with particle swarm optimization. *International Journal of Design, Nature & Ecodynamics*, **17**(2), 157-168.

6. Arnau Notari, M.R. 2022. Hydrodynamic and biochemical Computational Fluid Dynamic modelling of full-scale anaerobic digesters for wastewater treatment, Universitat Jaume I.

7. Asadi, M., McPhedran, K. 2021. Biogas maximization using data-driven modelling with uncertainty analysis and genetic algorithm for municipal wastewater anaerobic digestion. *Journal of Environmental Management*, **293**, 112875.

8. Asgari, S., Gupta, R., Puri, I.K., Zheng, R. 2021a. A data-driven approach to simultaneous fault detection and diagnosis in data centers. *Applied Soft Computing*, **110**, 107638.

9. Asgari, S., MirhoseiniNejad, S., Moazamigoodarzi, H., Gupta, R., Zheng, R., Puri, I.K. 2021b. A gray-box model for real-time transient temperature predictions in data centers. *Applied Thermal Engineering*, **185**, 116319.

10. Asgari, S., Moazamigoodarzi, H., Tsai, P.J., Pal, S., Zheng, R., Badawy, G., Puri, I.K. 2021c. Hybrid surrogate model for online temperature and pressure predictions in data centers. *Future Generation Computer Systems*, **114**, 531-547.

11. Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S., Rozzi, A., Sanders, W., Siegrist, H., Vavilin, V. 2002. The IWA anaerobic digestion model no 1 (ADM1). *Water Science and technology*, **45**(10), 65-73.

12. Beltramo, T., Klocke, M., Hitzmann, B. 2019. Prediction of the biogas production using GA and ACO input features selection method for ANN model. *Information Processing in Agriculture*, **6**(3), 349-356.

13. Cheon, A., Sung, J., Jun, H., Jang, H., Kim, M., Park, J. 2022. Application of Various Machine Learning Models for Process Stability of Bio-Electrochemical Anaerobic Digestion. *Processes*, **10**(1), 158.

14. Choi, S., Kim, S.I., Yulisa, A., Aghasa, A., Hwang, S. 2022. Proactive Prediction of Total Volatile Fatty Acids Concentration in Multiple Full-Scale Food Waste Anaerobic Digestion Systems Using Substrate Characteristics with Machine Learning and Feature Analysis. *Waste and Biomass Valorization*, 1-16.

15. Cinar, S.Ö., Cinar, S., Kuchta, K. 2022. Machine Learning Algorithms for Temperature Management in the Anaerobic Digestion Process. *Fermentation*, **8**(2), 65.

16. Cruz, I.A., Andrade, L.R., Bharagava, R.N., Nadda, A.K., Bilal, M., Figueiredo, R.T., Ferreira, L.F. 2021. An overview of process monitoring for anaerobic digestion. *Biosystems Engineering*, **207**, 106-119.

17. Cruz, I.A., Chuenchart, W., Long, F., Surendra, K., Andrade, L.R.S., Bilal, M., Liu, H., Figueiredo, R.T., Khanal, S.K., Ferreira, L.F.R. 2022. Application of machine learning in anaerobic digestion: Perspectives and challenges. *Bioresource Technology*, **345**, 126433.

18. Dahunsi, S.O., Oranusi, S., Efeovbokhan, V.E. 2017. Cleaner energy for cleaner production: Modeling and optimization of biogas generation from Carica papayas (Pawpaw) fruit peels. *Journal of cleaner production*, **156**, 19-29.

19. De Clercq, D., Wen, Z., Fei, F., Caicedo, L., Yuan, K., Shang, R. 2020. Interpretable machine learning for predicting biomethane production in industrial-scale anaerobic co-digestion. *Science of the Total Environment*, **712**, 134574.

20. Emebu, S., Pecha, J., Janáčová, D. 2022. Review on anaerobic digestion models: Model classification & elaboration of process phenomena. *Renewable and Sustainable Energy Reviews*, **160**, 112288.

21. Fawagreh, K., Gaber, M.M., Elyan, E. 2014. Random forests: from early developments to recent advancements. *Systems Science & Control Engineering: An Open Access Journal*, **2**(1), 602-609.

22. Ge, M., Friedrich, J., Vigna, L. 2020. World Resources Institute, Charts Explain Greenhouse Gas Emissions by Countries and Sectors, https://www.wri.org/insights/4-charts-explain-greenhouse-gas-emissions-countries-and-sectors#:~:text=China%20is%20the%20biggest%20emitter,the%20European%20Uni

on%20at%207.03%25.&text=Most%20of%20the%20top%2010,tCO2e%20per%20person.

23. Ge, Y., Tao, J., Wang, Z., Chen, C., Mu, L., Ruan, H., Yon, Y.R., Su, H., Yan, B., Chen, G. 2022. Modification of Anaerobic Digestion Model No. 1 with Machine Learning Models towards Applicable and Accurate Simulation of Biomass Anaerobic Digestion. *Chemical Engineering Journal*, 140369.

24. Ghatak, M.D., Ghatak, A. 2018. Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates. *Fuel*, **232**, 178-189.

25. Gupta, R., Miller, R., Sloan, W., You, S. 2022. Economic and environmental assessment of organic waste to biomethane conversion. *Bioresource Technology*, **345**, 126500.

26. Haffiez, N., Chung, T.H., Zakaria, B.S., Shahidi, M., Mezbahuddin, S., Maal-Bared, R., Dhar, B.R. 2022. Exploration of machine learning algorithms for predicting the changes in abundance of antibiotic resistance genes in anaerobic digestion. *Science of The Total Environment*, 156211.

27. Huang, M., Han, W., Wan, J., Ma, Y., Chen, X. 2016. Multi-objective optimisation for design and operation of anaerobic digestion using GA-ANN and NSGA-II. *Journal of Chemical Technology & Biotechnology*, **91**(1), 226-233.

28. Jacob, S., Banerjee, R. 2016. Modeling and optimization of anaerobic codigestion of potato waste and aquatic weed by response surface methodology and artificial neural network coupled genetic algorithm. *Bioresource technology*, **214**, 386-395.

29. Jeong, K., Abbas, A., Shin, J., Son, M., Kim, Y.M., Cho, K.H. 2021. Prediction of biogas production in anaerobic co-digestion of organic wastes using deep learning models. *Water Research*, **205**, 117697.

30. Južnič-Zonta, Ž., Kocijan, J., Flotats, X., Vrečko, D. 2012. Multi-criteria analyses of wastewater treatment bio-processes under an uncertainty and a multiplicity of steady states. *Water research*, **46**(18), 6121-6131.

31. Kazemi, P., Bengoa, C., Steyer, J.-P., Giralt, J. 2021. Data-driven techniques for fault detection in anaerobic digestion process. *Process Safety and Environmental Protection*, **146**, 905-915.

32. Kazemi, P., Giralt, J., Bengoa, C., Steyer, J.-P. 2020a. Data-driven fault detection methods for detecting small-magnitude faults in anaerobic digestion process. *Water Science and Technology*, **81**(8), 1740-1748.

33. Kazemi, P., Steyer, J.-P., Bengoa, C., Font, J., Giralt, J. 2020b. Robust data-driven soft sensors for online monitoring of volatile fatty acids in anaerobic digestion processes. *Processes*, **8**(1), 67.

34. Li, J., Zhang, L., Li, C., Tian, H., Ning, J., Zhang, J., Tong, Y.W., Wang, X. 2022. Data-driven based in-depth interpretation and inverse design of anaerobic digestion for CH4-rich biogas production. *ACS ES&T Engineering*, **2**(4), 642-652.

35. Long, F., Wang, L., Cai, W., Lesnik, K., Liu, H. 2021. Predicting the performance of anaerobic digestion using machine learning algorithms and genomic data. *Water Research*, **199**, 117182.

36. McCormick, M., Villa, A.E. 2019. LSTM and 1-D convolutional neural networks for predictive monitoring of the anaerobic digestion process. *International Conference on Artificial Neural Networks*. Springer. pp. 725-736.

37. Meneses, M., Concepción, H., Vrecko, D., Vilanova, R. 2015. Life cycle assessment as an environmental evaluation tool for control strategies in wastewater treatment plants. *Journal of Cleaner Production*, **107**, 653-661.

38. Natekin, A., Knoll, A. 2013. Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, **7**, 21.

39. Oloko-Oba, M.I., Taiwo, A.E., Ajala, S.O., Solomon, B.O., Betiku, E. 2018. Performance evaluation of three different-shaped bio-digesters for biogas production and optimization by artificial neural network integrated with genetic algorithm. *Sustainable Energy Technologies and Assessments*, **26**, 116-124.

40. Park, J.-G., Jun, H.-B., Heo, T.-Y. 2021. Retraining prior state performances of anaerobic digestion improves prediction accuracy of methane yield in various machine learning models. *Applied Energy*, **298**, 117250.

41. Schulz, E., Speekenbrink, M., Krause, A. 2018. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, **85**, 1-16.

42. Trucchia, A., Frunzo, L. 2021. Surrogate based global sensitivity analysis of ADM1-based anaerobic digestion model. *Journal of Environmental Management*, **282**, 111456.

43. Wang, L., Long, F., Liao, W., Liu, H. 2020. Prediction of anaerobic digestion performance and identification of critical operational parameters using machine learning algorithms. *Bioresource technology*, **298**, 122495.

44. Wang, Y., Huntington, T., Scown, C.D. 2021a. Tree-based automated machine learning to predict biogas production for anaerobic co-digestion of organic waste. *ACS Sustainable Chemistry & Engineering*, **9**(38), 12990-13000.

45. Wang, Y., Li, X. 2019. Soft measurement for VFA concentration in anaerobic digestion for treating kitchen waste based on improved DBN. *IEEE Access*, **7**, 60931-60939.

46. Wang, Y., Wang, S. 2021. Soft sensor for VFA concentration in anaerobic digestion process for treating kitchen waste based on SSAE-KELM. *IEEE Access*, **9**, 36466-36474.

47. Wang, Y., Yan, P., Gai, M. 2021b. Dynamic soft sensor for anaerobic digestion of kitchen waste based on SGSTGAT. *IEEE Sensors Journal*, **21**(17), 19198-19208.

48. Xu, W., Long, F., Zhao, H., Zhang, Y., Liang, D., Wang, L., Lesnik, K.L., Cao, H., Zhang, Y., Liu, H. 2021. Performance prediction of ZVI-based anaerobic digestion reactor using machine learning algorithms. *Waste Management*, **121**, 59-66.

49. Yan, P., Gai, M., Wang, Y., Gao, X. 2021. Review of Soft Sensors in Anaerobic Digestion Process. *Processes*, **9**(8), 1434.

50. Yan, P., Shen, B., Wang, Y. 2020. Soft sensor for VFA concentration in anaerobic digestion process for treating kitchen waste based on DSTHELM. *IEEE Access*, **8**, 223618-223625.

51. Yi-Fan, H., Chang-Zhu, Y., Jin-Feng, D., Wen-Hong, P., Jia-Kuang, Y. 2017. Modeling of expanded granular sludge bed reactor using artificial neural network. *Journal of environmental chemical engineering*, **5**(3), 2142-2150.

52. Zaied, B., Rashid, M., Nasrullah, M., Bari, B.S., Zularisam, A., Singh, L., Kumar, D., Krishnan, S. 2020. Prediction and optimization of biogas production from POME co-digestion in solar bioreactor using artificial neural network coupled with particle swarm optimization (ANN-PSO). *Biomass Conversion and Biorefinery*, 1-16.

53. Zareei, S., Khodaei, J. 2017. Modeling and optimization of biogas production from cow manure and maize straw using an adaptive neuro-fuzzy inference system. *Renewable Energy*, **114**, 423-427.

54. Zhao, B., Shuai, C., Hou, P., Qu, S., Xu, M. 2021. Estimation of unit process data for life cycle assessment using a decision tree-based approach. *Environmental Science & Technology*, **55**(12), 8439-8446.