Science of the Total Environment 712 (2020) 134574

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

# Interpretable machine learning for predicting biomethane production in industrial-scale anaerobic co-digestion

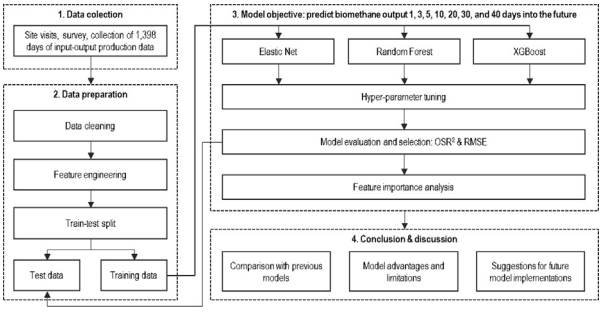
Djavan De Clercq [a], Zongguo Wen [a,*], Fan Fei [b], Luis Caicedo [c,d], Kai Yuan [c,e], Ruoxi Shang [c,f]

[a] State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, China
[b] College of Public Administration, Huazhong University of Science and Technology, China
[c] Bio-Tesseract, China
[d] EARTH University Costa Rica, Costa Rica
[e] Edinburgh Centre for Robotics, University of Edinburgh, Scotland, United Kingdom
[f] College of Engineering, University of California, Berkeley, United States

## HIGHLIGHTS

- Elastic net, random forest, and XGBoost were applied to industrial-scale biomethane prediction.
- Predictions were made over 1, 3, 5, 10, 20, 30, and 40-day time horizons to evaluate robustness.
- Depending on the time horizon, out-of-sample $R^2$ values of up to 0.88 were achieved.
- Feature importance and partial dependence were used to interpret these machine models.
- Machine learning models can be a useful supplement to experimental and mechanistic models of AD.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 9 July 2019
Received in revised form 17 September 2019
Accepted 19 September 2019
Available online 23 December 2019

*Keywords:*
Machine learning
XGBoost
Random forest
Anaerobic co-digestion
Biomethane

## ABSTRACT

The objective of this study is to apply machine learning models to accurately predict daily biomethane production in an industrial-scale co-digestion facility. The methodology involved applying elasticnet, random forest, and extreme gradient boosting to input–output data from an industrial-scale anaerobic co-digestion (ACoD) facility. The models were used to predict biomethane for 1-day, 3-day, 5-day, 10-day, 20-day, 30-day, and 40-day time horizons. These models were fit on four years of operational data. The results showed that elastic net (a model with assumptions of linearity) was clearly outperformed by random forest and extreme gradient boosting (XGBoost), which had out-of-sample $R^2$ values ranging between 0.80 and 0.88, depending on the time horizon. In addition, feature importance and partial dependence analysis demonstrated the marginal and interaction effects on biomethane of selected bio-waste inputs. For instance, food waste co-digested with percolate were shown to have strong positive interaction effects. One implication of this study is that XGBoost and random forest algorithms applied to industrial-scale ACoD data provide dependable prediction results and may be a useful complement for experimental and mechanistic/theoretical models of anaerobic digestion, especially where detailed substrate characterization is difficult. However, these models have limitations, and suggestions for deriving additional value from these methods are proposed.



* Corresponding author.
*E-mail address:* wenzg@tsinghua.edu.cn (Z. Wen).

https://doi.org/10.1016/j.scitotenv.2019.134574
0048-9697/

## 1. Introduction

### *1.1. Anaerobic digestion modeling tools have evolved significantly, but their practical application to industrial-scale facilities is limited*

There are many factors that affect biomethane output in industrial anaerobic co-digestion (ACoD) facilities. These include temperature, input substrate biodegradability, pH, auxiliary chemical inputs, and more (Hagos et al., 2017). In addition, the interaction between these factors can be complex and highly non-linear, which makes modeling anaerobic digestion (AD) phenomena a challenge, despite the relative maturity of AD technologies (Hu et al., 2018; Tan et al., 2018).

In the last few decades, many models have been developed to describe the various phenomena that occur during the AD process. These models, which were mainly derived through theoretical, empirical, and statistical approaches, describe phenomena such as methane potential, reactor stability, AD process inhibition, and liquid–gas interface mass-transfer. The comprehensive reviews by Donoso-Bravo et al. (2011), Lauwers et al. (2013), Xu et al. (2015), and (Manchala et al., 2017) provide insight into the development history of these models.

One theoretical approach which has gained considerable popularity is the Anaerobic Digestion Model No. 1 (ADM1), developed by Batstone et al. (2002). This model was developed as a unified modeling framework to simulate reactions in the AD process based on the dynamics of 24 microbial species and 19 bioconversion processes. In order to maintain simplicity (Donoso-Bravo et al., 2011) and generalizability to a broad range of experimental conditions, several species and processes were not included in the original generic model.

Despite this original generic setup, the model has been extended to broaden its applicability to novel AD scenarios (Batstone et al., 2006; Mendes et al., 2015; Poggio et al., 2016), mostly based on experimental-scale studies. For instance, (Zhao et al., 2019) modified the ADM1 structure to reflect the key kinetic parameters in lab-scale anaerobic digestion of food waste, with strong results. Other recent ADM1 modifications have focused on incorporating the effects of calcium and magnesium ions (Zhang et al., 2015), metabolic products such as lactic acid and ethanol (Antonopoulou et al., 2012), volatile fatty acids (Bai et al., 2017), trace elements such iron, nickel and cobalt (Maharaj et al., 2018), propionate oxidizing bacteria (Uhlenhut et al., 2018) and specific waste types such slaughterhouse waste (Spyridonidis et al., 2018).

ADM1 has occasionally been adapted to industrial applications (Batstone and Keller, 2003), but with more limited success. For instance Derbal et al. (2009) showed that ADM1 had acceptable performance for model fit on an industrial-scale case study for certain parameters. However, the study also concluded that the ADM1 model is relatively limited in modeling complex AD processes, and the model showed weak performance at the initial, transient stage of experimentation. A key reason for this is that one of the greatest challenges of ADM1 is the need for detailed substrate characterization (meaning determination of the specific composition of waste based on metrics such as volatile fatty acids, chemical oxygen demand, and protein content), since the model depends on a rigid stoichiometric approach.

This characterization requires a more extensive analysis of substrate input than most biogas plants are easily capable of. Although several studies have suggested ways to facilitate this characterization (Girault et al., 2012; Koch et al., 2010), these methods may still require excessive amounts of analysis. (Nordlander et al., 2017) applied the ADM1 model in an industrial-scale case study, but found model fits ranging from 0.28 to 0.78, and concluded that the extensive analysis needed for substrate characterization is a hindrance for ADM1's utilization in industrial facilities. The study recommended exploration of additional models that do not use chemical oxygen demand as a base unit for model input. Biernacki et al. (2013) also demonstrated the complexity involved in the characterization of complex substrates. For instance, they used 15 different characterization parameters such as dry mass, raw protein, lipid content, and raw fibre, which may be impractical for some industrial facilities.

Studies like this imply that while ADM1 is useful for process design and dynamic simulation, the model's fixed stoichiometric approach, which would necessitate excessive structural modifications in data collection efforts in industrial-scale facilities, may render it impractical in certain situations. Additional evidence is needed to demonstrate ADM1's practical effectiveness in modeling AD phenomena in full-scale facilities, since studies applying the ADM1 model to industrial projects – especially those where co-digestion takes place – are especially rare (Nordlander et al., 2017).

Similar limitations apply to other theoretical models. These include the two-particle model of anaerobic solid state fermentation (Kalyuzhnyi et al., 2000), which considers solid state anaerobic digestion to be heterogenous with two kinds of particles in one reactor: "seed" and "waste", based on the biodegradability and methanogenic activity. The reaction front model (Martin et al., 2003) was developed to interpret the slow transport mechanism in biogas reactors. The distributed model (Vavilin et al., 2003) is a simple one-dimensional model that involves mass transfer due to diffusion and leachate flow in solid-state AD reactors. The diffusion limitation model (Xu et al., 2014) was developed to determine the impact of total solids on lignocellulosic substrates. Limitations of other models applied to industrial-scale facilities have been described in Donoso-Bravo et al. (2011)

In addition to theoretical/mathematical models of AD such as ADM1, recent literature is rich with empirical, lab-scale studies that have investigated the impact of various parameters on anaerobic co-digestion. For instance, Algapani et al. (2017) investigated the effect of co-digesting food waste and sewage sludge on methane production. Vivekanand et al. (2018) found that co-digesting whey, manure, and fish ensilage resulted higher biomethane potential than the weighted average of individual substrates. Valenti et al. (2018a) evaluated biomethane potential from co-digestion of olive pomace, citrus pulp, poultry manure, sainfoin silage and cladodes. Many similar lab-scale studies investigating the relationship between various process parameters and methane yield have been conducted in recent years (Bohutskyi et al., 2018; Jiang et al., 2018; Jin et al., 2018; López González et al., 2017; Marques et al., 2018; Meneses-Reyes et al., 2018; Mokomele et al., 2019; Pastor-Poquet et al., 2018; Pinto et al., 2018; Rahman et al., 2018; Thorin et al., 2018; Tsapekos et al., 2019; Valenti et al., 2018b).

Unfortunately, such studies also suffer from limited applicability to industrial-scale facilities in that the results of lab-scale studies results do not always generalize to full-fledged industrial AD facilities. In fact certain results, such as waste substrate mixtures deemed optimal for biogas production in experimental conditions, might actually be "highly improper for industrial applications" (Matuszewska et al., 2016). Moreover, experimental determination of biomethane potential may take up to 50 days to produce results (Lauwers et al., 2013), which may be impractical for the operators of industrial facilities.

### *1.2. Traditional statistical models are helpful, but application to industrial case studies also have limitations*

According to a comprehensive review of anaerobic digestion modeling approaches conducted by Manchala et al. (2017), statistical models can be useful when exact mechanistic explanations (derived from theoretical or empirical models) of a specific AD pro-

cess are not available. In addition, these approaches may be easier to implement compared to the models above, since statistical models do not require a deep understanding of biological or physico-chemical processes (Xu et al., 2015). This especially pertinent when practical resource constraints make investigating the mechanisms behind certain phenomena infeasible.

In some cases, the predictive performance of statistical approaches is actually similar to that of mechanistic models such as ADM1. For instance, Thomsen et al. (2014) developed canonical linear and quadratic mixture models to estimating biomethane potential from lignocellulosic biomass based on experimental data, with high r-squared values above 90%. Other studies include Le Hyaric et al. (2012), which used linear regression to show the relationship of specific methane activity to moisture content during mesophilic digestion of municipal solid waste. Motte et al. (2013) used quadratic multiple linear regression to model the impact of TS content, inoculation ratio, and particle size of lignocellulosic biomass on anaerobic digestion. Similarly, multiple linear regression has been used in numerous studies to predict methane yields of different substrates based on their chemical composition (Gunaseelan, 2009; Raposo et al., 2012; Strömberg et al., 2015; Tong et al., 1990).

Unfortunately, such linear models may not be desirable for industrial processes. Dandikas et al. (2018) assessed previously published linear prediction models and found that linear models had serious limitations for generalized prediction of biomethane potential. Although experimental literature has showed that in some cases, a single process parameter can dominate the AD process and thus be modelled by linear regression, in most cases ACoD is a complex, non-linear process which may be better modelled with more complex approaches (Xu et al., 2014) that account for non-linearity (Lauwers et al., 2013).

As a result, some studies have applied machine learning techniques, which are capable of modeling non-linear relationships in various AD phenomena. However, these studies have also been largely limited to lab-scale set-ups. For example, Mahanty et al. (2013) applied an artificial neural network to biogas prediction from industrial sludge from the paper, chemical, petrochemical, automobile and food industries in a lab-scale study. Sinha et al. (2002) applied a neural network to simulating upflow anaerobic sludge blanket (UASB) performance, and trained the network on experimentally-obtained data. Jacob and Banerjee, (2016) modeled ACoD of potato waste and aquatic weed based on an artificial neural network (ANN) coupled with a genetic algorithm (GA). Wang et al. (2018) applied artificial neural networks to monitoring alkalinity in a lab-scale anaerobic co-digestion system.

The few studies that have applied machine learning techniques to full-scale facilities include Hamed et al. (2004) which used an artificial neural network to model the effluent concentrations of BOD and SS for a wastewater treatment facility in Cairo. However, the data collected in the study were limited in time and scope, and did not include important process parameters such as pH and temperature.

### 1.3. Research contributions

This study made two novel contributions. Firstly, this was the first study demonstrating the strong predictive power of machine learning algorithms such as XGBoost applied to biomethane modeling, based on industrial-scale ACoD project data. In addition, emphasis was placed on the interpretability of these models, which challenges the notion of machine learning being a purely "black box" approach when applied to modeling of the AD process. Secondly, this study used data collected over four years of project operation at a prominent biomethane facility. Such extensive time-series studies are rare in the literature. These two contributions address many of the limitations mentioned in sections 1.1. and 1.2.

## 2. Methodology

The methodology in this study is based on the flowchart in Fig. 1. Firstly, four years of operational data were collected from an ACoD facility producing vehicle-quality biomethane. Secondly, data was prepared and partitioned prior to model building. Thirdly, various machine learning models were trained in order to predict biomethane output 1, 3, 5, 10, 20, 30, and 40 days into the future based on a range of features. Lastly, the results and their implica-

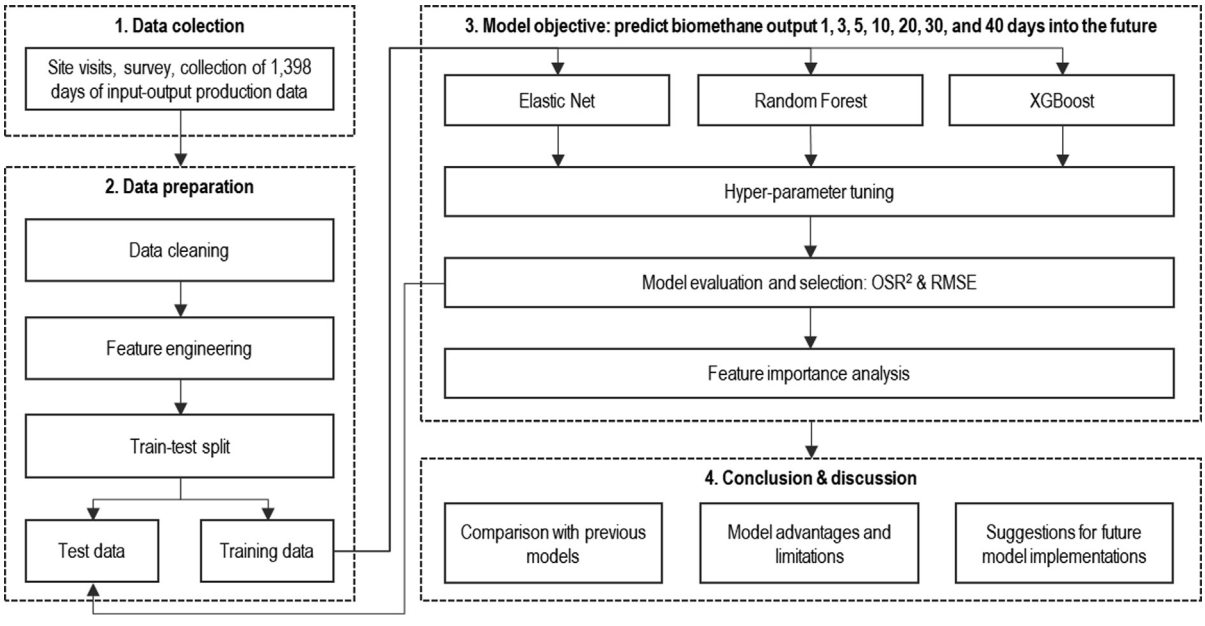



**Fig. 1.** Overview of methodology.

**Table 1**
Summary of collected data, engineered features, and features used in model training.

| Feature | Type | Notation | Raw/Derived | Model use | Units |
|---|---|---|---|---|---|
| Day number | Time | $t_1, \cdots t_{1,398}$ | Raw | N/A | |
| Pig manure | Input | $IN_1$ | Raw | Feature | t/d (wet waste) |
| Cassava | Input | $IN_2$ | Raw | Feature | t/d (wet waste) |
| Fish waste | Input | $IN_3$ | Raw | Feature | t/d (wet waste) |
| Food waste | Input | $IN_4$ | Raw | Feature | t/d (wet waste) |
| Municipal fecal waste | Input | $IN_5$ | Raw | Feature | t/d (wet waste) |
| Tea waste | Input | $IN_6$ | Raw | Feature | t/d (wet waste) |
| Chicken manure | Input | $IN_7$ | Raw | Feature | t/d (wet waste) |
| Baggasse | Input | $IN_8$ | Raw | Feature | t/d (wet waste) |
| Alcohol waste | Input | $IN_9$ | Raw | Feature | t/d (wet waste) |
| Medicine waste | Input | $IN_{10}$ | Raw | Feature | t/d (wet waste) |
| Energy grass | Input | $IN_{11}$ | Raw | Feature | t/d (wet waste) |
| Banana shafts | Input | $IN_{12}$ | Raw | Feature | t/d (wet waste) |
| Lemon waste | Input | $IN_{13}$ | Raw | Feature | t/d (wet waste) |
| Percolate | Input | $IN_{14}$ | Raw | Feature | t/d (wet waste) |
| Other waste | Input | $IN_{15}$ | Raw | Feature | t/d (wet waste) |
| Total waste | Input | $IN_{16}$ | Raw | Feature | t/d (wet waste) |
| Co-digested wastes | Input | $IN_{17}$ | Derived | Feature | Number |
| Biomethane 40 days ahead | Output | $OUT_{40}$ | Raw, time lead | Target | $m^3$ |
| Biomethane 30 days ahead | Output | $OUT_{30}$ | Raw, time lead | Target | $m^3$ |
| Biomethane 20 days ahead | Output | $OUT_{20}$ | Raw, time lead | Target | $m^3$ |
| Biomethane 10 days ahead | Output | $OUT_{10}$ | Raw, time lead | Target | $m^3$ |
| Biomethane 50 days ahead | Output | $OUT_5$ | Raw, time lead | Target | $m^3$ |
| Biomethane 3 days ahead | Output | $OUT_3$ | Raw, time lead | Target | $m^3$ |
| Biomethane 1 days ahead | Output | $OUT_1$ | Raw, time lead | Target | $m^3$ |

tions were discussed within the broader context of practical ACoD modeling in industrial-scale biogas facilities.

### 2.1. Data collection

The case study project is located in Hainan, a tropical island province in southern China which is a primarily agricultural economy, with half of all exports being agricultural products. Crops cultivated in Hainan include sugarcane, palm oil, rice, coffee, tea, coconuts, tropical fruit, and cashews. Livestock breeding includes chickens, cows, goats, water buffalo, geese, and ducks, and the island is also a major producer of fish products and various agriculture-derived processed goods.

The surveyed ACoD facility treats a wide variety of organic wastes including agricultural waste (livestock manure, banana straw, sugar cane bagasse, etc.) industrial waste (high-concentration organic wastewater, fish product waste, etc.), and municipal waste (food waste, human fecal waste, etc.). The project covers an area of about 33,000 square meters, and can co-digest up to 850 t/d of biowaste. The CNY160 million facility can produce 30,000 $m^3$/d of biomethane, which can meet the fuel demand of 250 buses or more than 750 taxis, replacing 8,600 t/y of diesel and reducing carbon emissions by 25,800 t/y.

The facility was continuously surveyed over a period of almost four years, from July 2014 to February 2018, equivalent or 46 months or 1,398 days. Operational input data collected included process inputs such as waste type and corresponding daily input volume, electricity consumption, water consumption, and auxiliary chemical inputs. Output data collected included daily values for raw biogas volume, biomethane volume, liquid digestate volume, and solid digestate volume. Additional information can be found in the supplementary information.

### 2.2. Data preparation

Raw data was consolidated into a unified format and thoroughly cleaned (i.e. outlier removal due to sensor failures on specific days and subsequent nearest-neighbor imputation, removal of non-numeric values, conversion to factor/numeric, etc.). In addition, additional features were generated from the initial data, such as a categorical variable indicating the count of the number of input waste types co-digested on any given day. Table 1 provides an overview of the collected data and specific features used in model training, and Fig. 2 visualizes the primary process inputs and outputs.

As shown in Fig. 2, the biomethane and fertilizer outputs fluctuated strongly throughout the operational lifetime of the project. The average values for biomethane, liquid fertilizer and solid fertilizer outputs were 5,423 $m^3$/d, 40.95 t/d, and 9.76 t/d respectively. Average waste input was 172.5 t/d over the survey period, although the specific waste composition fluctuated significantly (bottom right panel of Fig. 2)

### 2.3. Model development

#### 2.3.1. Model overview

The objective of the machine learning models (described in detail in section 2.3.2) was to predict future biomethane output at 1, 3, 5, 10, 20, 30, and 40 days into the future, corresponding to variables $OUT_1$, $OUT_3$, $OUT_5$, $OUT_{10}$, $OUT_{20}$, $OUT_{30}$, and $OUT_{40}$. These variables are represented as target output vectors $\boldsymbol{v}_1, \boldsymbol{v}_3, \boldsymbol{v}_5, \boldsymbol{v}_{10}, \boldsymbol{v}_{20}, \boldsymbol{v}_{30}$, and $\boldsymbol{v}_{40}$ in Fig. 3. Prior to model training, data was partitioned into a training set and a test set based on a 0.90 train-test split ratio, resulting in 1,222 training observations and 136 test observations. This high split ratio was chosen to reflect the scarcity of data and to maximize the models' ability to learn from the data. Models were trained to map a function $f(x)$ capable of using a $1,222 \times 17$ matrix $\boldsymbol{A}$ composed of features $IN_1, \cdots, IN_{17}$ across 1,222 days (observations) to predict biomethane output. A total of 21 models were trained; Elastic Net, Random Forest, XGBoost models were trained seven times each to predict future biomethane $OUT_1$, $OUT_3$, $OUT_5$, $OUT_{10}$, $OUT_{20}$, $OUT_{30}$, and $OUT_{40}$.

Supervised machine learning algorithms refer to methods such as classification and regression, where we have access to $p$ features $X_1, X_2, \cdots, X_p$, measured on $n$ observations, as well as a response

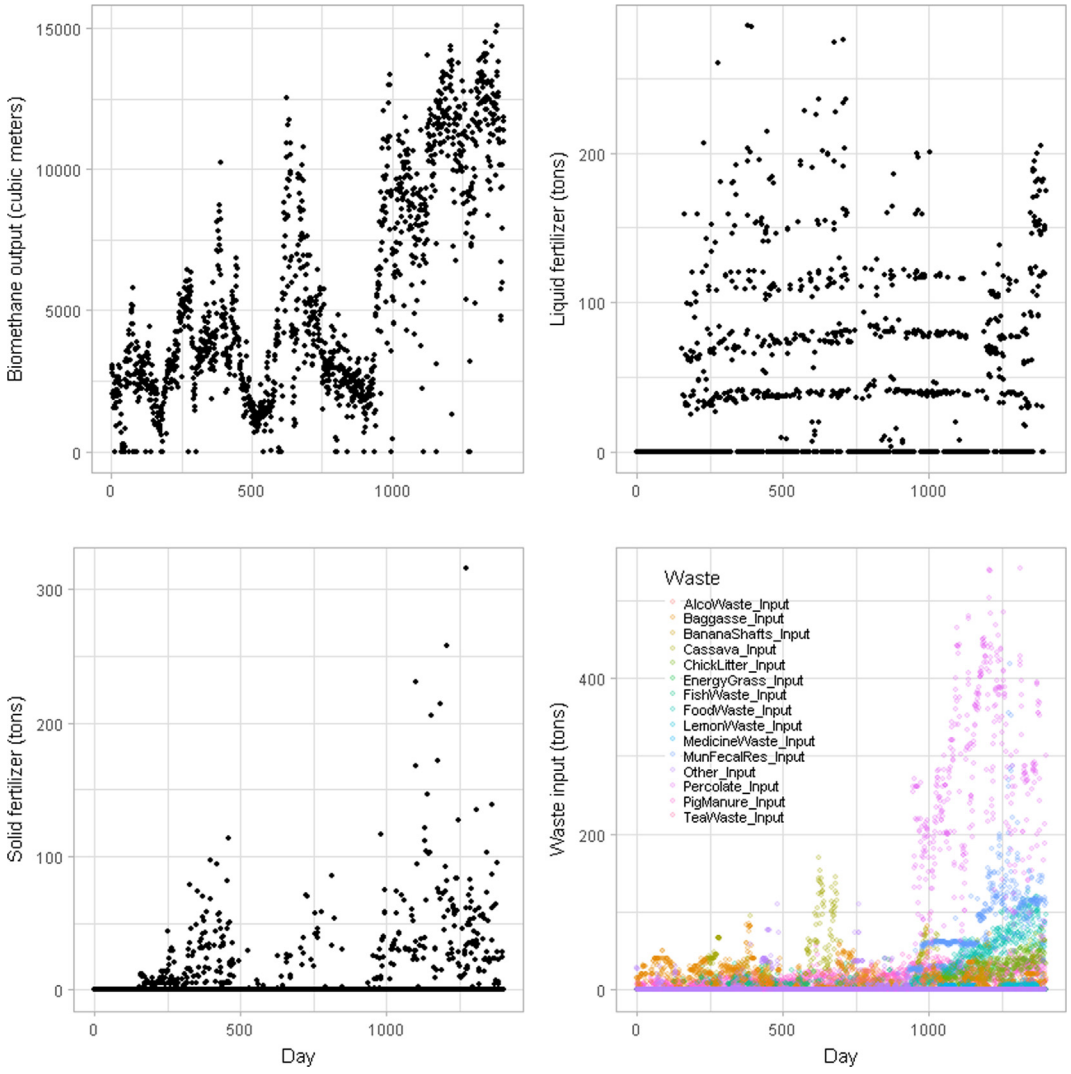



**Fig. 2.** visualization of primary inputs and outputs in the ACoD facility.

variable $Y$ also measured on $n$ observations. The objective is to predict $Y$ using $X_1, X_2, \cdots, X_p$. Supervised learning can be useful for both regression problems (where the $Y$ variable takes a continuous quantitative value) and classification problems (which occur when a response variable $Y$ takes a qualitative/categorical value). This study uses features $IN_1, \cdots, IN_{17}$ to predict the target variables $T_1$, $OUT_3$, $OUT_5$, $OUT_{10}$, $OUT_{20}$, $OUT_{30}$, and $OUT_{40}$ (future biomethane yield at different prediction horizons). The algorithms applied to this supervised learning problem included elastic net, random forest, and XGBoost.

### 2.3.2. Elastic net

The elastic net (EN) regression algorithm, proposed by Zou and Hastie (2005), is a regularized least squares regression method that has been widely used for supervised learning. The method is especially useful where the number of features is large and variable selection is required or favored. Elastic net is a hybrid approach that linearly combines the $L_1$ and $L_2$ penalties of the lasso and ridge methods. A naïve formulation can be expressed as

$$\widehat{\beta} = \arg\min_{\beta} |\boldsymbol{y} - \boldsymbol{X}\beta|^2 + \lambda_2|\beta|^2 + \lambda_1|\beta| \quad (1)$$

Where $\lambda_1$ And $\lambda_2$ are tuning parameters, $\boldsymbol{X}$ is an $n \times m$ sample matrix in which the $i$th row contains sample $\boldsymbol{x_i}$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^T$ is the response vector, $|\cdot|^2$ is the squared sum of all elements in the vector, and $|\cdot|$ is the sum of absolute values of all elements in the vector (Tan et al., 2011). Additional mathematical details can be found in Hastie et al. (2009) and Angelini (2019).

### 2.3.3. Random forest

The random forest (RF) algorithm, first proposed by Ho (1995) and further developed by Breiman (2001) is an ensemble learning method for regression and classification. It is a substantial modification of bootstrap aggregation (bagging) which operates by constructing many de-correlated decision trees and outputs their average predictions. The key idea in bagging is to average numerous noisy but unbiased models to reduce variance. Decision trees are ideal base models for the bagging process, since they capture complex interactions in the data. However, since decision trees are noisy models, they benefit from the averaging induced by bagging.

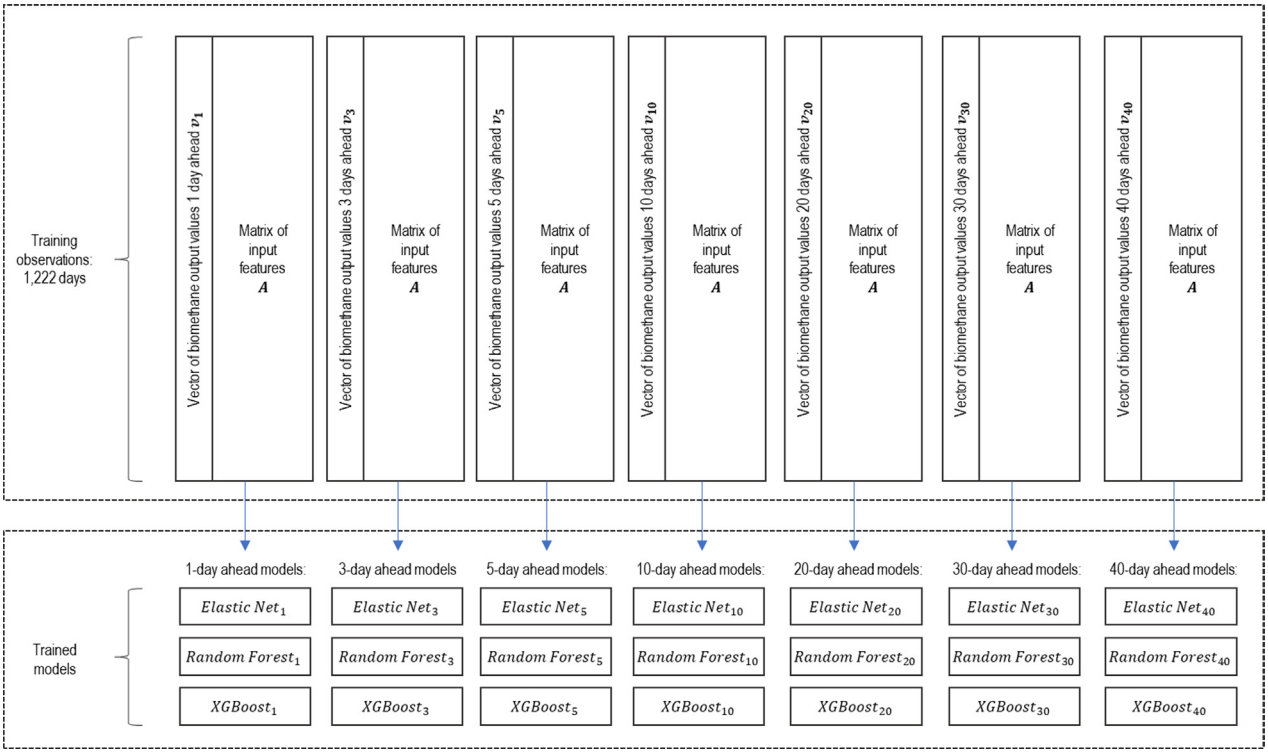



**Fig. 3.** overview of data structure and trained models.

**Algorithm 1: *Random Forest***

1. For $\boldsymbol{b} = 1$ to $\boldsymbol{B}$ :
   a. Draw a bootstrap sample $\boldsymbol{Z}^*$ of size $\boldsymbol{N}$ from the training data.
   b. Grow a random-forest tree $\boldsymbol{T_b}$ to the bootstrapped data, by recursive repetition of the following steps for each tree's terminal node, until minimum node size $\boldsymbol{n_{min}}$ is achieved.
      i. Select $\boldsymbol{m}$ variables randomly from the $\boldsymbol{p}$ variables.
      ii. Select the best variable/split-point among $\boldsymbol{m}$.
      iii. Split the node into two nodes.
2. Output the ensemble of trees $\{\boldsymbol{T_b}\}_1^{\boldsymbol{B}}$

To make a regression prediction based on a new point $x$ : $\widehat{f}_{\mathrm{rf}}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T(x;\Theta_b)$.

The key idea of random forests (see Algorithm 1) is to further enhance the variance reduction of bagging by de-correlating trees, without excessively increasing variance. This is achieved by random selection of input variables, specifically by selecting $m \leq p$ of the input variables at random as splitting candidates.

*2.3.4. Xgboost*

Extreme gradient boosting (XGBoost), proposed by Chen and Guestrin (2016), is an efficient and scalable variant of the gradient boosting machine (GBM) algorithm. Specifically, XGBoost uses a more regularized model formulation than GBM, which controls over-fitting. In addition, the XGBoost implementation has important performance enhancements including (1) the use of sparse matrices; (2) improved data structures for better processor cache utilization; and (3) support for multi-core processing which reduces training time. It commonly outperforms powerful methods such as support vector machines (SVM) and deep learning neural networks (Stojić et al., 2019).

Like random forest, the XGBoost builds an ensemble of decision trees which can capture non-linear interactions in data. A key difference, however, is that boosting grows decision trees sequentially, meaning that the growth of a particular tree takes into account the other trees that have already been grown. Algorithm 2 presents the generic gradient tree boosting algorithm which XGBoost is based on.

**Algorithm 2: Gradient Tree Boosting**

a. Initialize $\boldsymbol{f}_0(\boldsymbol{x}) = \arg\min_{\gamma} \sum_{\boldsymbol{i}=1}^{\boldsymbol{N}} \boldsymbol{L}(\boldsymbol{y_i}, \gamma)$

b. For $\boldsymbol{m} = 1$ to $\boldsymbol{M}$:

c. For $\boldsymbol{i} = 1, 2, \cdots, \boldsymbol{N}$ compute $\boldsymbol{r_{im}} = -\left[\frac{\partial \boldsymbol{L}(\boldsymbol{y_i}, \boldsymbol{f}(\boldsymbol{x_i}))}{\partial \boldsymbol{f}(\boldsymbol{x_i})}\right]_{\boldsymbol{f}=\boldsymbol{f}_{m-1}}$

d. Fit a regression tree to the targets $\boldsymbol{r_{im}}$ giving terminal regions $\boldsymbol{R_{jm}}, \boldsymbol{j} = 1, 2, .\boldsymbol{J_m}$

e. For $\boldsymbol{j} = 1, 2, .\boldsymbol{J_m}$
compute $\boldsymbol{\gamma_{jm}} = \arg\min_{\gamma} \sum_{\boldsymbol{x_i} \in \boldsymbol{R_{jm}}} \boldsymbol{L}(\boldsymbol{y_i}, \boldsymbol{f}_{\boldsymbol{m}-1}(\boldsymbol{x_i}) + \gamma)$

f. Update $\boldsymbol{f_m}(\boldsymbol{x}) = \boldsymbol{f}_{\boldsymbol{m}-1}(\boldsymbol{x}) + \sum_{\boldsymbol{j}=1}^{\boldsymbol{J_m}} \boldsymbol{\gamma_{jm}} \boldsymbol{I}(\boldsymbol{x} \in \boldsymbol{R_{jm}})$.

g. Output $\widehat{\boldsymbol{f}}(\boldsymbol{x}) = \boldsymbol{f_M}(\boldsymbol{x})$.

*2.3.5. Hyperparameter tuning*

In machine learning, various models have hyperparameters which can be optimized to reduce over-fitting and enhance the models' prediction performance (Hamidieh, 2018). For elastic net, the tuned parameters included (1) fraction of full solution and (2) weight decay. For random forest, the tuning parameters included (1) number of randomly selected predictors, (2) splitting rule, and (3) minimal node size. For XGBoost, the tuning parameters included (1) number of boosting iterations, (2) max tree depth, (3) shrinkage, (4) minimum loss reduction, (5) subsample ratio of columns, (6) minimum sum of instance weight, and (7) subsample

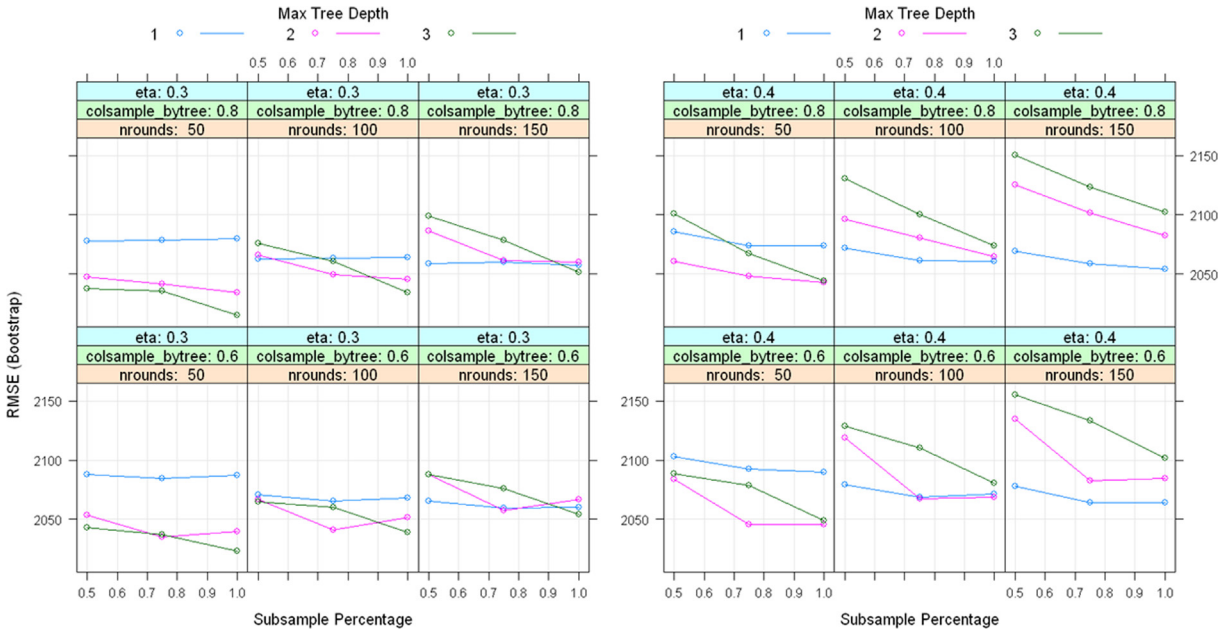



**Fig. 4.** tuned hyperparameters of 40-day prediction horizon XGBoost model.

percentage (Climent et al., 2018). An example of the tuned hyperparameters of an XGBoost model at the 40-day prediction horizon is illustrated in Fig. 4. More information on the final parameter values for the best-performing models can be found in the online code supplement.

### 2.3.6. Model evaluation

After models have been trained, it is necessary to test their generalization performance. Two metrics were used for model evaluation, out-of-sample r-squared ($OSR^2$) and root mean squared error (RMSE). RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{N}\left(\widehat{F}(x) - y\right)^2} \quad (2)$$

Where $N$ is the number of data points in the test set, $\widehat{F}$ is the output of the predictor, and $y$ is the actual output value. $OSR^2$ is computed via

$$R^2_{OS} = 1 - \frac{\sum_{t=1}^{N}\left(r_t - \widehat{r}_t\right)^2}{\sum_{t=1}^{N}\left(r_t - \overline{r_t}\right)^2} \quad (3)$$

where $\widehat{r}_t$ is the fitted value from a predictive regression estimated though period $t-1$, and $\overline{r_t}$ is the average of the regressand through period $t-1$.

### 2.3.7. Feature importance and partial dependence

It is useful to interpret the derived approximation of $f(x)$ in this study to better understand the impact of particular input variables that are most influential in contribution to its variance (Friedman, 2001). This study computed and visualized feature importance plots to better understand the features driving variance in the target variable. Feature importance can be understood as the increase in model prediction error from removing certain variables.

However, insights provided from ranking features by their importance is limited. Once important features have been identified, it is often necessary to evaluate the relationship between them and the response variable. This can be achieved by constructing partial dependence plots (PDPs), which visualize the relationship between a subset of model features (typically one to three features) and the target variable, while accounting for the average effect of other model predictors. The partial dependence of a target variable on input feature $\boldsymbol{z}_s$ can be defined as

$$f_s(\boldsymbol{z}_s) = \boldsymbol{E}_{\boldsymbol{z}_c}\left[\widehat{f}(x)\right] = \boldsymbol{E}_{\boldsymbol{z}_c}\left[\widehat{f}(\boldsymbol{z}_s, \boldsymbol{z}_c)\right] = \int \widehat{f}(\boldsymbol{z}_s, \boldsymbol{z}_c) p_c(\boldsymbol{z}_c) d\boldsymbol{z}_c \quad (4)$$

where $x = \{x_1, x_2, \cdots, x_p\}$ represents predictors in a model where the prediction function is $\widehat{f}(x)$, $\boldsymbol{x}$ is partitioned into an interest set $\boldsymbol{z_s}$ and its compliment $\boldsymbol{z}_c = \boldsymbol{x}\{\boldsymbol{z}_s, \boldsymbol{p}_c$ is the marginal probability density of $\boldsymbol{z}_c : \boldsymbol{p}_c(\boldsymbol{z}_c) = \int p(\boldsymbol{x}) d\boldsymbol{z}_s$. Eq. 4 can be estimated on training data by

$$\overline{f}_s(\boldsymbol{z}_s) = \frac{1}{n}\sum_{i=1}^{n}\widehat{f}\left(\boldsymbol{z}_s, \boldsymbol{z}_{i,c}\right) \quad (5)$$

where $z_{i,c}$ $(i = 1, 2, \cdots, n)$ are values of $z_c$ found in the training sample, meaning that effects of all the other features in the model are averaged out. In the context of ACoD facilities, partial dependence plots are especially useful to better understand the effects of various waste inputs on the AD process. In addition, they can also reveal interactions between various waste types to shed light on the co-digestion dynamics between different wastes.

### 2.3.8. Computation

All computation was conducted in the statistical programming language R. Various packages were employed in the analysis for model building (caret, xgboost, ranger, etc.), data manipulation (tidyverse, reshape2, etc.) data visualization (ggplot2, ggpubr), and parallel computation for enhanced speed in model training (doParallel). To ensure reproducibility, the full dataset is available upon request, and code used in this study is open-source and can be found at: https://github.com/djavandeclercq/BiomethanePredictionML.

## 3. Results and discussion

### 3.1. Comparison of model performance

The tree-based models (random forest and XGBoost) performed much better than elastic net. They were able to consistently predict biomethane output at various time leads with an $OSR^2$ above 0.80. Elastic net predicted biomethane one day with an $OSR^2$ of 0.85. However, as the time lead increased, the $OSR^2$ value declined consistently to 0.50 for biomethane predictions 40 days ahead. Random forest had an $OSR^2$ value of 0.88 for predictions one day

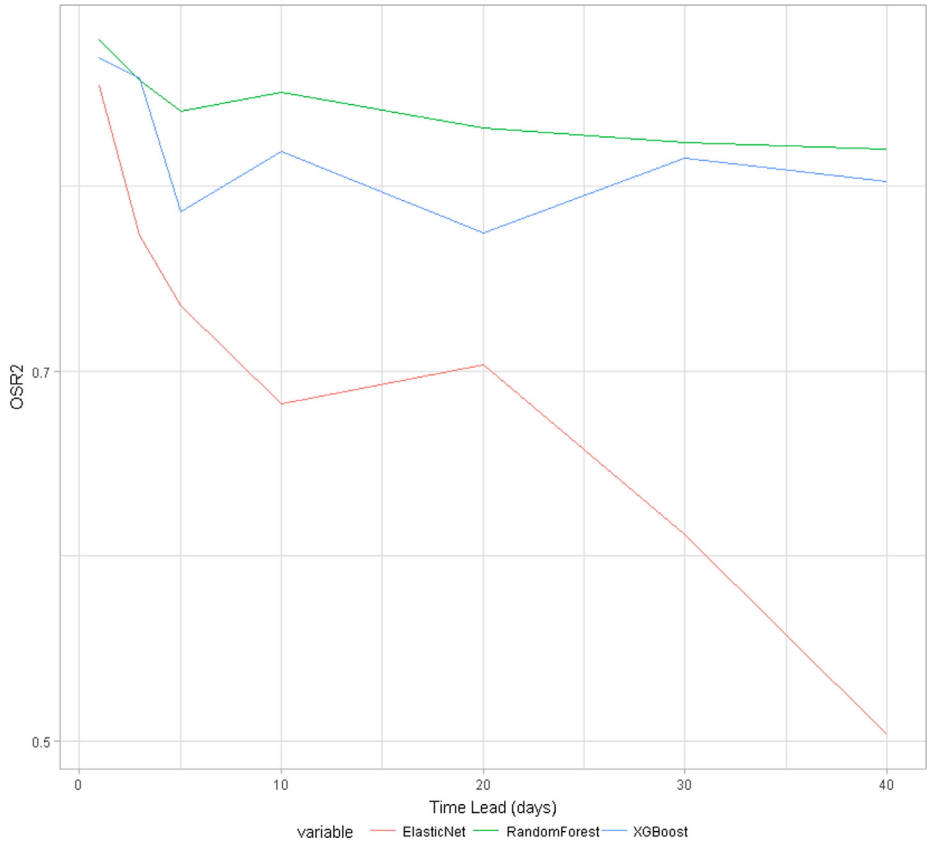



**Fig. 5.** Comparison of model out-of-sample performance ($n$ days ahead).

ahead, and 0.82 for 40 days ahead. XGBoost managed an $OSR^2$ of 0.87 for predictions one day ahead, and 0.80 for predictions 40 days ahead. These performance values are summarized in Fig. 5.

The higher performance of tree-based models is intuitive, as they are more capable of capturing more complex interactions in the data than standard linear models.

Fig. 6 shows the predictions of the three models compared to the actual biomethane output values for six different time leads (1, 5, 10, 20, 30, and 40 days). The 3-day predictions can be found in the supplementary information. As can be inferred from the $OSR^2$ performance values above, elastic net tracked actual output closely with a short time lead (1, 5 days), but became increasingly inaccurate as the prediction horizon extended. The tree-based models, on the other hand, showed a steadier performance as the prediction horizon extended. While the models were incapable of matching the exact peaks and troughs of actual biomethane output, they performed well in capturing the overall movement of biomethane output.

### 3.2. Feature importance

Fig. 7 and Fig. 8 show the feature importance (the factor by which the error is increased compared to the original model if a certain feature is permuted) charts generated for the random forest and XGBoost models trained for six different prediction horizons. The figures illustrate how the importance values of different features evolve as the prediction horizon extends.

In Fig. 7 (random forest) for instance, for the one-day biomethane prediction horizon, important features included total input, percolate, and the number of co-digested wastes. For the 40-day prediction horizon, however, the top three features were food waste input, followed by total input and percolate input.

Fig. 8 shows that the feature importance values computed via XGBoost are concentrated among a smaller subset of variables. This is intuitive, since when there are several correlated features, boosting tends to choose one and use it in several trees. For random forest however, there is a random selection of features used in each tree, meaning that each correlated feature has the chance to be selected in one of these trees. Random forest's parallel tree construction means that each tree is not aware of which features have been used in other trees. Hence, for XGBoost, total input is highly correlated with other input features, making it an important feature based on the boosting method.

### 3.3. Partial dependence

The partial dependence plots in Fig. 9 show the marginal effect of a feature on the predicted outcome of a previously fit model (Friedman, 2001). The prediction function is fixed at certain values of the chosen features and averaged over the other features, allowing us to isolate the effect of changing the value of a feature of interest on future biomethane output. For example, the bottom right panel of Fig. 9 demonstrates marginal effect of increasing food waste on biomethane output, holding all other waste inputs at a fixed amount. The top right panel of Fig. 9 shows that higher percolate input has a strong positive effect on biomethane output until about 175 tons, when the curve starts to flatten. The blue line is LOESS (Locally Estimated Scatterplot Smoothing) smoothed to facilitate interpretability.

Fig. 10 shows two-variable partial dependence plotted with a convex hull, which outlines the region of the selected predictor space (in this case food waste and municipal fecal residue) that the model was trained on. This negates the possibility of over-interpreting partial dependence beyond a reasonable region.

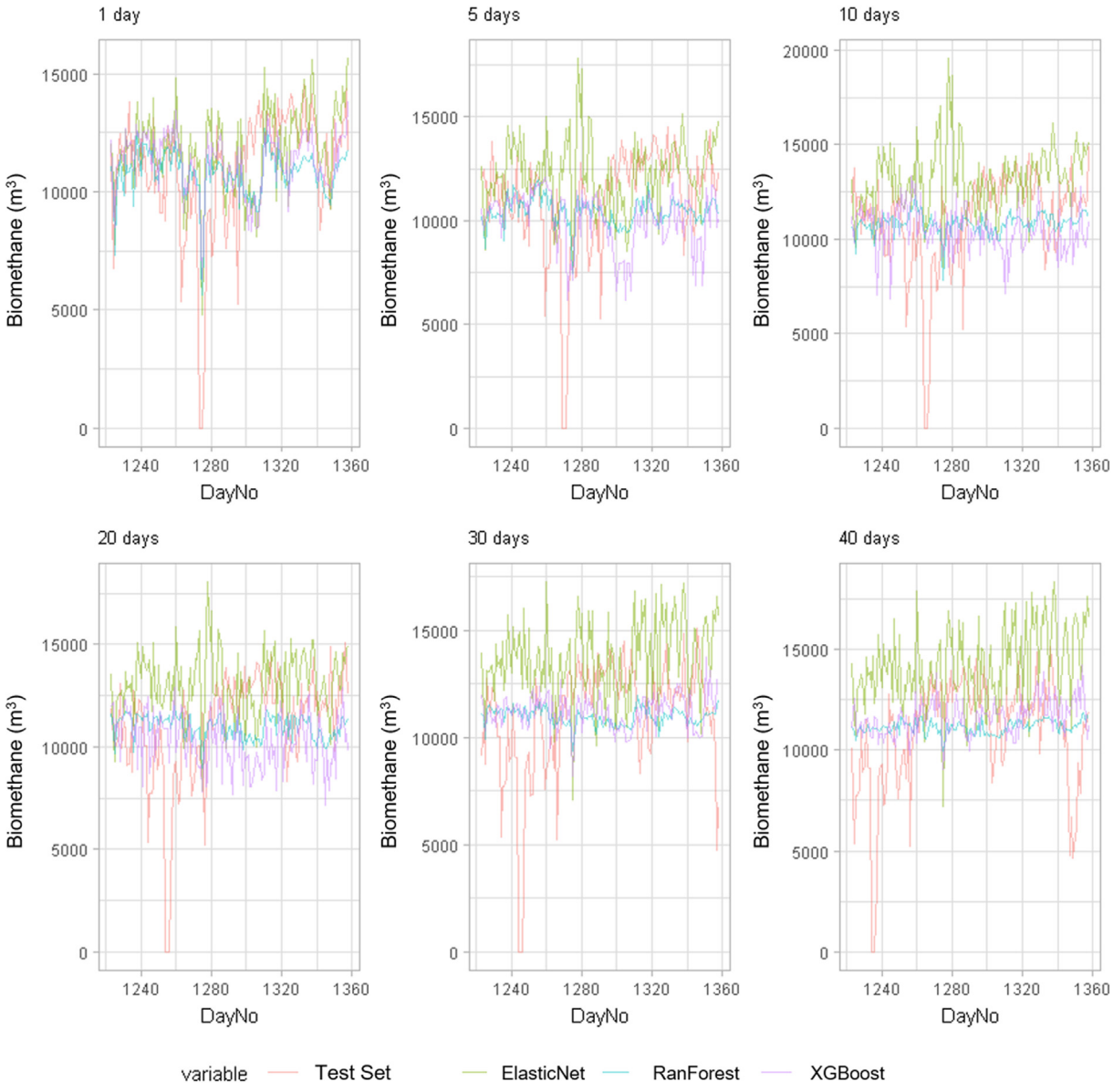



**Fig. 6.** Comparison of model predictions (*n* days ahead).

Two-variable partial dependence is useful for showing the interaction effect between two variables on biomethane output. Fig. 10 provides the 2D two-way partial dependence visualizations for ease of interpretability. For instance, the left panel shows partial dependence for both food waste input and percolate input, and it can be inferred that there exists a positive interaction between food waste input and percolate input, since increasing both features leads to increased biomethane production. The positive interaction flattens out at approximately 40 tons of food waste input and 100 tons of percolate input.

### 3.4. Discussion

#### 3.4.1. Limitations of machine learning applied to ACoD

Machine learning approaches, including artificial neural networks, are often criticized for their "black box" nature (Xu et al., 2015), and may be problematic from an interpretability perspective when applied to ACoD. For instance, while linear model coefficients and tree-based model feature importance scores might provide some indication of variable importance, they do no give a clear view on the underlying mechanisms governing the ACoD process (Lauwers et al., 2013). Another problem is that some machine learning models may be prone to overfitting data and hence simulating non-informative noise.

Another crucial limitation is that machine learning approaches require a substantial amount of data. This might not be available in the early stages of an industrial-scale facility's operation. In this context, it might be necessary to conduct experimental studies to determine optimal digester conditions. For instance, Mirmasoumi et al. (2018) investigated the effect of pre-treatment, digestion temperature increase, and co-digestion in a WWTP to enhance biomethane productivity. These kinds of experiments, where certain conditions are isolated, can be difficult to implement with machine learning in the absence of historical process data.

Based on this observation, further research needs to be conducted on the minimum amount of data required for training useful models. This study used four years of data for model

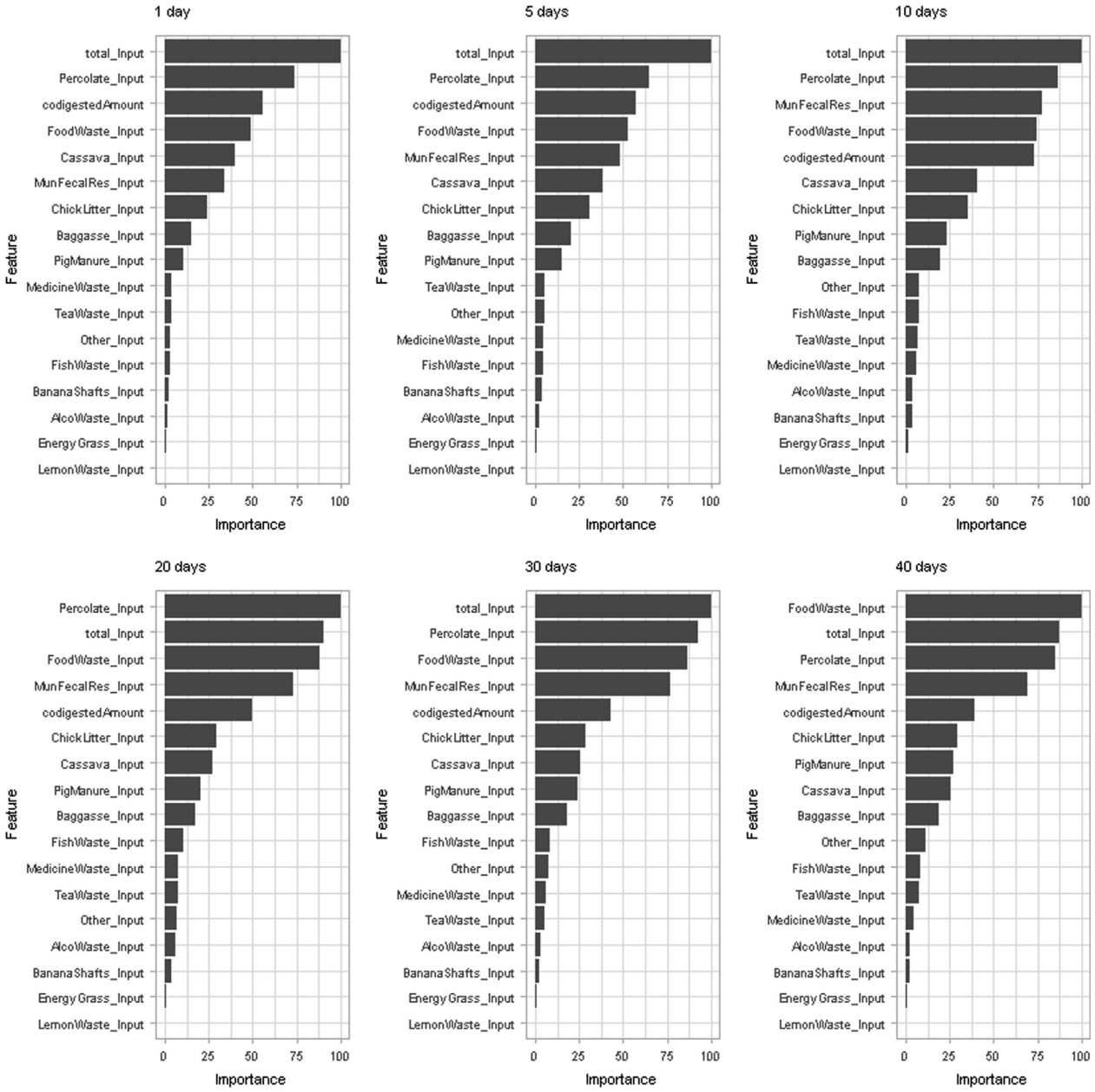



**Fig. 7.** Random forest feature importance.

training; future research may look into using training data sets restricted to 2-months, 4-months, or 6-months (and so on) of observations.

### 3.4.2. Advantages of machine learning applied to ACoD

One great advantage of machine learning applied to ACoD is that they do not require the extensive substrate characterization that is a prerequisite for models such as ADM1 (Lauwers et al., 2013). In addition, these models do not require a deep understanding of biological or physicochemical processes (Xu et al., 2015), and as a result can be much easier to develop and apply to practical problems encountered in operational facilities. This allows industrial ACoD project managers greater flexibility in deriving insights from their operational data, regardless of the structure this data might have.

Moreover, machine learning models are increasingly interpretable, as demonstrated by the application of feature importance analysis in conjunction with partial dependence analysis. Additional model-agnostic methods for better interpretation include Individual Conditional Expectation (Goldstein et al., 2015), Accumulated Local Effects (ALE), Global Surrogate, Local Surrogate, and Shapley Values (Papadopoulos and Kontokosta, 2019). This is a large boon for ACoD facility operators, as machine learning models on operational big data can be used as a precursor to experimental studies or mechanistic model development.

### 3.4.3. Suggestions for enhancing the value derived from applying machine learning to ACoD

To implement out-of-box machine learning models effectively, ACoD facilities could implement a comprehensive strategy which encompasses the entire data analytics pipeline, including sensor integration (e.g. monitoring temperature and pH), automatic data collection (which minimizes the need for error-prone human data input), dashboard analytics (visualization of process data in

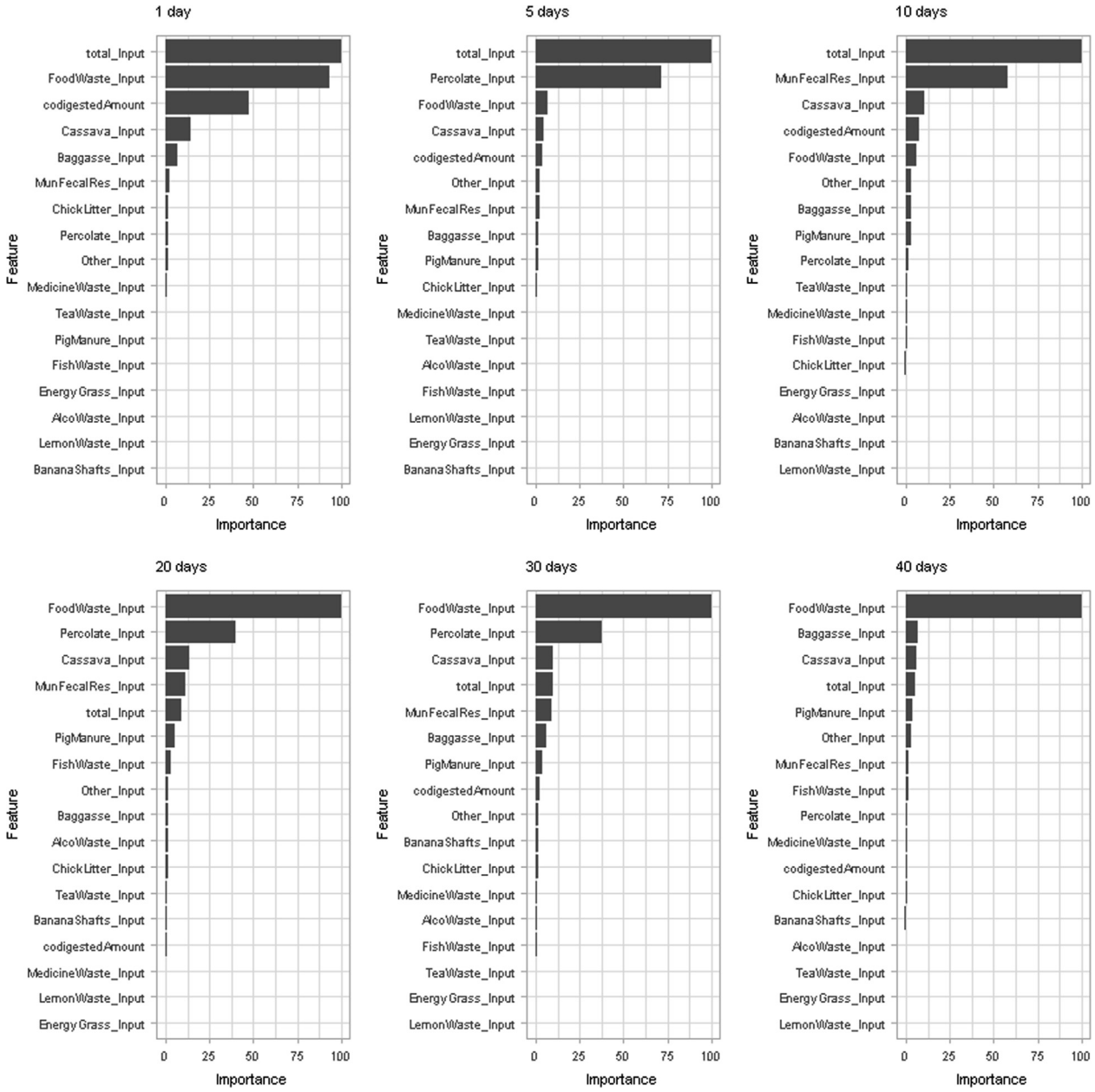



**Fig. 8.** XGBoost feature importance.

real-time), and predictive machine learning (e.g. predicting biomethane output, digestate production, and occurrence of process bottlenecks such as foaming). However, such initiatives would have to be carefully considered from an economic perspective; comparing the extra costs of such installations to the benefits that dynamic process optimization can offer requires further research.

A positive finding is that high accuracies were achieved with just 17 model input features, corresponding to the input volume of different waste types. However, higher accuracy might be achieved by using sensors to incorporate additional online parameter measurements (pH, temperature, VFA, TOC, HRT, etc.). This would further reduce the need for lab analysis and provide additional information to boost machine learning models' predictive power and interpretability. However, additional research is still required regarding the efficacy of machine learning approaches in the case of facilities with a low level of automation, a limited number of installed sensors, and spare data.

Machine learning models are often used empirically, and it can be tempting to adopt an "as long as it works" mentality when applying these models. As a result, researchers may be tempted to engage in complex feature engineering in order to squeeze slightly higher accuracies out of these models. A downside of this is that it may become difficult to easily interpret the trained models. As a result, if interpretability is favored over accuracy, easily understandable variables should be selected as inputs to machine learning models so that feature importance and partial dependence results can be interpreted intuitively. For instance, it would be easier to interpret the partial dependence plot of a simple organic waste input $x$ or chemical de-foaming agent $y$ rather than inter-

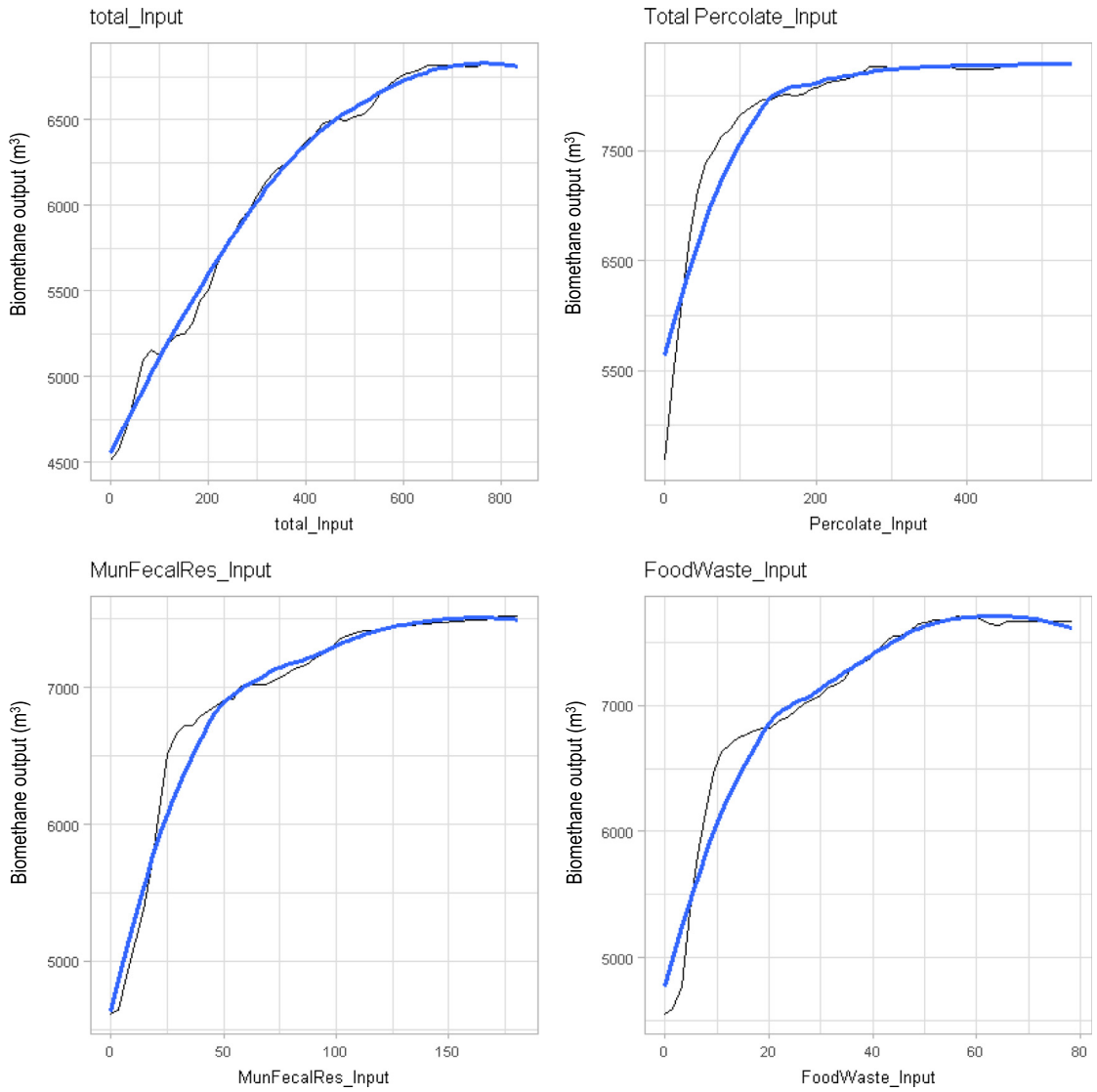



**Fig. 9.** Partial dependence plots for the random forest model at 20-day prediction horizon.

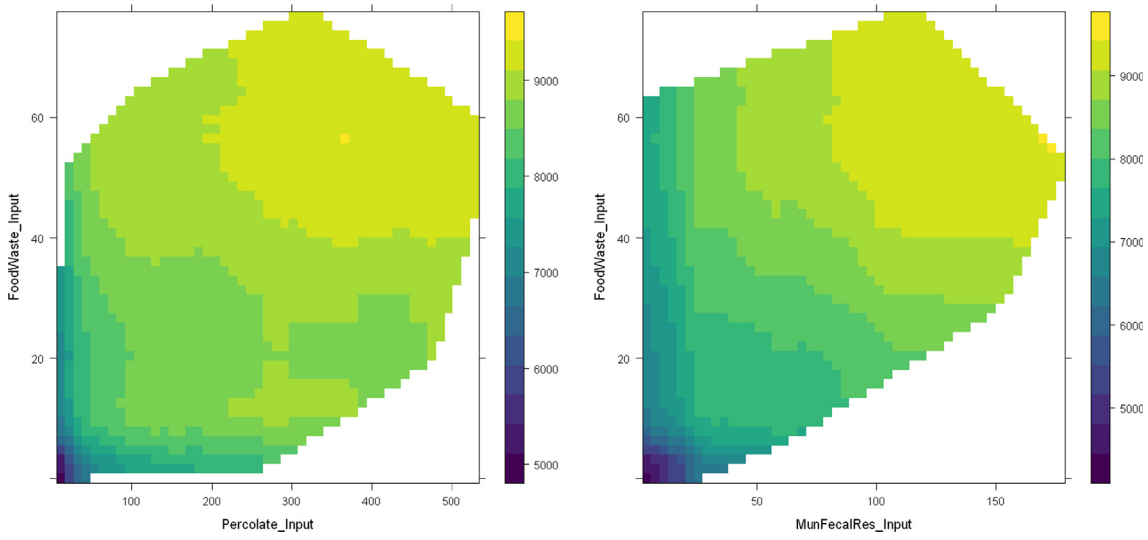



**Fig. 10.** Two-variable partial dependence plots for the random forest model at 20-day prediction horizon.

preting the plot for a running cumulative sum of waste input $x$ or the logarithm of de-foaming agent $y$.

## 4. Conclusion

This study addressed the issue of interpretable ACoD modeling in industrial-scale ACoD facilities by combining XGBoost and random forest algorithms with interpretability tools such as partial dependence analysis to accurately predict biomethane output. These models, trained on four years of process data, shed light on the usefulness of applying machine learning techniques to industrial-scale anaerobic digestion. The results resolved some of the limitations of previous literature pertaining to the generalizability of lab-scale, theoretical, and linear statistical models to large-scale ACoD projects.

This study applied three machine learning regression models to modeling biomethane output at as function of various organic waste inputs in an industrial-scale anaerobic co-digestion facility. The models were used to predict biomethane output at various time horizons: 1 day, 3 days, 5 days, 10 days, 20, days, 30 days, and 40 days.

Elastic net (a model with assumptions of linearity) was clearly outperformed by random forest and XGBoost, which had out-of-sample $R^2$ values ranging between 0.80 and 0.88, depending on the time horizon. In addition, feature importance and partial dependence analysis was used to demonstrate the marginal effects and interaction effects selected biowaste features.

Machine learning is not a silver bullet for modeling complex interactions during anaerobic co-digestion in industrial-scale facilities. However, this study demonstrated that it can be a highly valuable tool with strong predictive power and good interpretability. Used in conjunction with empirical lab-scale studies and mechanistic models, machine learning can offer valuable insights to the operators of industrial-scale ACoD facilities. It is especially useful in the short term if the plant operator possesses daily input–output data and would like to derive rapid insight from this data prior to more costly experimentation.

Additional value can be derived from machine learning applied to industrial-scale anaerobic digestion by (1) enhancing end-to-end data integration in ACoD facilities; (2) incorporating online sensor measurement of important process parameters; and (3) training models on a set of features than can be intuitively analyzed in subsequent feature importance and partial dependence analyses.

## Acknowledgements

The authors gratefully acknowledge the financial support from the “Thirteenth Five-Year” National Key R&D Proram of China (2018YFC190080), National Science Fund for Distinguished Young Scholars of China (71825006) and General Programs (71774099). The responsibility for any error rests solely with the authors. The contents of this paper reflect the views of the authors and do not necessarily indicate acceptance by the sponsors.

## References

Algapani, D.E., Wang, J., Qiao, W., Su, M., Goglio, A., Wandera, S.M., Jiang, M., Pan, X., Adani, F., Dong, R., 2017. Improving methane production and anaerobic digestion stability of food waste by extracting lipids and mixing it with sewage sludge. Bioresour. Technol. 244, 996–1005. https://doi.org/10.1016/j.biortech.2017.08.087.

Angelini, C., 2019. Regression Analysis, in: Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C.B.T.E. of B. and C.B. (Eds.), . Academic Press, Oxford, pp. 722–730. doi:10.1016/B978-0-12-809633-8.20360-9.

Antonopoulou, G., Gavala, H.N., Skiadas, I.V., Lyberatos, G., 2012. Modeling of fermentative hydrogen production from sweet sorghum extract based on modified ADM1. Int. J. Hydrogen Energy 37, 191–208. https://doi.org/10.1016/j.ijhydene.2011.09.081.

Bai, J., Liu, H., Yin, B., Ma, H., Chen, X., 2017. Modified ADM1 for modeling free ammonia inhibition in anaerobic acidogenic fermentation with high-solid sludge. J. Environ. Sci. 52, 58–65. https://doi.org/10.1016/j.jes.2016.03.004.

Batstone, D.J., Keller, J., 2003. Industrial applications of the IWA anaerobic digestion model No. 1 (ADM1). Water Sci. Technol.

Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S.V., Pavlostathis, S.G., Rozzi, A., Sanders, W.T.M., Siegrist, H., Vavilin, V.A., 2002. The IWA Anaerobic Digestion Model No 1 (ADM1). Water Sci. Technol. 45, 65–73.

Batstone, D.J., Keller, J., Steyer, J.P., 2006. A review of ADM1 extensions, applications, and analysis: 2002–2005. Water Sci. Technol. 54, 1–10.

Biernacki, P., Steinigeweg, S., Borchert, A., Uhlenhut, F., 2013. Application of Anaerobic Digestion Model No. 1 for describing anaerobic digestion of grass, maize, green weed silage, and industrial glycerine. Bioresour. Technol. 127, 188–194. https://doi.org/10.1016/j.biortech.2012.09.128.

Bohutskyi, P., Phan, D., Kopachevsky, A.M., Chow, S., Bouwer, E.J., Betenbaugh, M.J., 2018. Synergistic co-digestion of wastewater grown algae-bacteria polyculture biomass and cellulose to optimize carbon-to-nitrogen ratio and application of kinetic models to predict anaerobic digestion energy balance. Bioresour. Technol. 269, 210–220. https://doi.org/10.1016/j.biortech.2018.08.085.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Chen, T., Guestrin, C., 2016. XGBoost, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16. ACM Press, New York, New York, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785

Climent, F., Momparler, A., Carmona, P., 2018. Anticipating bank distress in the Eurozone: an extreme gradient boosting approach. J. Bus. Res.

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2018. Predicting methane yield by linear regression models: a validation study for grassland biomass. Bioresour. Technol. 265, 372–379. https://doi.org/10.1016/j.biortech.2018.06.030.

Derbal, K., Bencheikh-lehocine, M., Cecchi, F., Meniai, A.-H., Pavan, P., 2009. Application of the IWA ADM1 model to simulate anaerobic co-digestion of organic waste with waste activated sludge in mesophilic condition. Bioresour. Technol. 100, 1539–1543. https://doi.org/10.1016/j.biortech.2008.07.064.

Donoso-Bravo, A., Mailier, J., Martin, C., Rodríguez, J., Aceves-Lara, C.A., Wouwer, A. Vande, 2011. Model selection, identification and validation in anaerobic digestion: a review. Water Res. 45, 5347–5364. https://doi.org/10.1016/j.watres.2011.08.059.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.

Girault, R., Bridoux, G., Nauleau, F., Poullain, C., Buffet, J., Steyer, J.-P., Sadowski, A.G., Béline, F., 2012. A waste characterisation procedure for ADM1 implementation based on degradation kinetics. Water Res. 46, 4099–4110. https://doi.org/10.1016/j.watres.2012.04.028.

Goldstein, A., Kapelner, A., Bleich, J., Pitkin, E., 2015. Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. J. Comput. Graph. Stat. 24, 44–65. https://doi.org/10.1080/10618600.2014.907095.

Gunaseelan, V.N., 2009. Predicting ultimate methane yields of Jatropha curcus and Morus indica from their chemical composition. Bioresour. Technol. 100, 3426–3429. https://doi.org/10.1016/j.biortech.2009.02.005.

Hagos, K., Zong, J., Li, D., Liu, C., Lu, X., 2017. Anaerobic co-digestion process for biogas production: progress, challenges and perspectives. Renew. Sustain. Energy Rev. 76, 1485–1496. https://doi.org/10.1016/j.rser.2016.11.184.

Hamed, M.M., Khalafallah, M.G., Hassanien, E.A., 2004. Prediction of wastewater treatment plant performance using artificial neural networks. Environ. Model. Softw. 19, 919–928. https://doi.org/10.1016/j.envsoft.2003.10.005.

Hamidieh, K., 2018. A data-driven statistical model for predicting the critical temperature of a superconductor. Comput. Mater. Sci. 154, 346–354. https://doi.org/10.1016/j.commatsci.2018.07.052.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, 2nd ed. Springer-Verlag, New York, NY, USA. doi:10.1007/978-0-387-84858-7

Ho, T.K., 1995. Random Decision Forests, in: Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1, ICDAR '95. IEEE Computer Society, Washington, DC, USA, p. 278–.

Hu, C., Yan, B., Wang, K., Xiao, X., 2018. Modeling the performance of anaerobic digestion reactor by the anaerobic digestion system model (ADSM). J. Environ. Chem. Eng. 6, 2095–2104. https://doi.org/10.1016/j.jece.2018.03.018.

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. Bioresour. Technol. 214, 386–395. https://doi.org/10.1016/j.biortech.2016.04.068.

Jiang, Y., Dennehy, C., Lawlor, P.G., Hu, Z., Yang, Q., McCarthy, G., Tan, S.P., Zhan, X., Gardiner, G.E., 2018. Inactivation of Salmonella during dry co-digestion of food waste and pig manure. Waste Manag. 82, 231–240. https://doi.org/10.1016/j.wasman.2018.10.037.

Jin, W., Xu, X., Yang, F., Li, C., Zhou, M., 2018. Performance enhancement by rumen cultures in anaerobic co-digestion of corn straw with pig manure. Biomass Bioenergy 115, 120–129. https://doi.org/10.1016/j.biombioe.2018.05.001.

Kalyuzhnyi, S., Veeken, A., Hamelers, B., 2000. Two-particle model of anaerobic solid state fermentation. Water Sci. Technol.

Koch, K., Lübken, M., Gehring, T., Wichern, M., Horn, H., 2010. Biogas from grass silage – Measurements and modeling with ADM1. Bioresour. Technol. 101, 8158–8165. https://doi.org/10.1016/j.biortech.2010.06.009.

Lauwers, J., Appels, L., Thompson, I.P., Degrève, J., Van Impe, J.F., Dewil, R., 2013. Mathematical modelling of anaerobic digestion of biomass and waste: Power and limitations. Prog. Energy Combust. Sci. 39, 383–402. https://doi.org/10.1016/j.pecs.2013.03.003.

Le Hyaric, R., Benbelkacem, H., Bollon, J., Bayard, R., Escudié, R., Buffière, P., 2012. Influence of moisture content on the specific methanogenic activity of dry mesophilic municipal solid waste digestate. J. Chem. Technol. Biotechnol. 87, 1032–1035. https://doi.org/10.1002/jctb.2722.

López González, L.M., Pereda Reyes, I., Romero Romero, O., 2017. Anaerobic co-digestion of sugarcane press mud with Vinasse on methane yield. Waste Manag. 68, 139–145. https://doi.org/10.1016/j.wasman.2017.07.016.

Mahanty, B., Zafar, M., Park, H.-S., 2013. Characterization of co-digestion of industrial sludges for biogas production by artificial neural network and statistical regression models. Environ. Technol. 34, 2145–2153. https://doi.org/10.1080/09593330.2013.819022.

Maharaj, B.C., Mattei, M.R., Frunzo, L., van Hullebusch, E.D., Esposito, G., 2018. ADM1 based mathematical model of trace element precipitation/dissolution in anaerobic digestion processes. Bioresour. Technol. 267, 666–676. https://doi.org/10.1016/j.biortech.2018.06.099.

Manchala, K.R., Sun, Y., Zhang, D., Wang, Z.-W., 2017. Chapter Two - Anaerobic Digestion Modelling, in: Li, Y., Ge, X.B.T.-A. in B. (Eds.), Elsevier, pp. 69–141. doi:10.1016/bs.aibe.2017.01.001

Marques, I.P., Batista, A.P., Coelho, A., Lopes da Silva, T., 2018. Co-digestion of Rhodosporidium toruloides biorefinery wastes for biogas production. Process Biochem. 64, 221–227. https://doi.org/10.1016/j.procbio.2017.09.023.

Martin, D.J., Potts, L.G.A., Heslop, V.A., 2003. Reaction mechanisms in solid-state anaerobic digestion: 1. the reaction front hypothesis. Process Saf. Environ. Prot. 81, 171–179. https://doi.org/10.1205/095758203765639870.

Matuszewska, A., Owczuk, M., Zamojska-Jaroszewicz, A., Jakubiak-Lasocka, J., Lasocki, J., Orliński, P., 2016. Evaluation of the biological methane potential of various feedstock for the production of biogas to supply agricultural tractors. Energy Convers. Manag. 125, 309–319. https://doi.org/10.1016/j.enconman.2016.02.072.

Mendes, C., Esquerre, K., Matos Queiroz, L., 2015. Application of Anaerobic Digestion Model No. 1 for simulating anaerobic mesophilic sludge digestion. Waste Manag. 35, 89–95. https://doi.org/10.1016/j.wasman.2014.10.013.

Meneses-Reyes, J.C., Hernández-Eugenio, G., Huber, D.H., Balagurusamy, N., Espinosa-Solares, T., 2018. Oil-extracted Chlorella vulgaris biomass and glycerol bioconversion to methane via continuous anaerobic co-digestion with chicken litter. Renew. Energy 128, 223–229. https://doi.org/10.1016/j.renene.2018.05.053.

Mirmasoumi, S., Ebrahimi, S., Saray, R.K., 2018. Enhancement of biogas production from sewage sludge in a wastewater treatment plant: Evaluation of pretreatment techniques and co-digestion under mesophilic and thermophilic conditions. Energy 157, 707–717. https://doi.org/10.1016/j.energy.2018.06.003.

Mokomele, T., da Costa Sousa, L., Balan, V., van Rensburg, E., Dale, B.E., Görgens, J.F., 2019. Incorporating anaerobic co-digestion of steam exploded or ammonia fiber expansion pretreated sugarcane residues with manure into a sugarcane-based bioenergy-livestock nexus. Bioresour. Technol. 272, 326–336. https://doi.org/10.1016/j.biortech.2018.10.049.

Motte, J.-C., Escudié, R., Bernet, N., Delgenes, J.-P., Steyer, J.-P., Dumas, C., 2013. Dynamic effect of total solid content, low substrate/inoculum ratio and particle size on solid-state anaerobic digestion. Bioresour. Technol. 144, 141–148. https://doi.org/10.1016/j.biortech.2013.06.057.

Nordlander, E., Thorin, E., Yan, J., 2017. Investigating the possibility of applying an ADM1 based model to a full-scale co-digestion plant. Biochem. Eng. J. 120, 73–83. https://doi.org/10.1016/j.bej.2016.12.014.

Papadopoulos, S., Kontokosta, C.E., 2019. Grading buildings on energy performance using city benchmarking data. Appl. Energy 233–234, 244–253. https://doi.org/10.1016/j.apenergy.2018.10.053.

Pastor-Poquet, V., Papirio, S., Trably, E., Rintala, J., Escudié, R., Esposito, G., 2018. Semi-continuous mono-digestion of OFMSW and Co-digestion of OFMSW with beech sawdust: Assessment of the maximum operational total solid content. J. Environ. Manage.

Pinto, M.P.M., Mudhoo, A., de Alencar Neves, T., Berni, M.D., Forster-Carneiro, T., 2018. Co–digestion of coffee residues and sugarcane vinasse for biohythane generation. J. Environ. Chem. Eng. 6, 146–155. https://doi.org/10.1016/j.jece.2017.11.064.

Poggio, D., Walker, M., Nimmo, W., Ma, L., Pourkashanian, M., 2016. Modelling the anaerobic digestion of solid organic waste – Substrate characterisation method for ADM1 using a combined biochemical and kinetic parameter estimation approach. Waste Manag. 53, 40–54. https://doi.org/10.1016/j.wasman.2016.04.024.

Rahman, M.A., Møller, H.B., Saha, C.K., Alam, M.M., Wahid, R., Feng, L., 2018. Anaerobic co-digestion of poultry droppings and briquetted wheat straw at mesophilic and thermophilic conditions: Influence of alkali pretreatment. Renew. Energy 128, 241–249. https://doi.org/10.1016/j.renene.2018.05.076.

Raposo, F., De la Rubia, M.A., Fernández-Cegrí, V., Borja, R., 2012. Anaerobic digestion of solid organic substrates in batch mode: An overview relating to methane yields and experimental procedures. Renew. Sustain. Energy Rev. 16, 861–877. https://doi.org/10.1016/j.rser.2011.09.008.

Sinha, S., Bose, P., Jawed, M., John, S., Tare, V., 2002. Application of neural network for simulation of upflow anaerobic sludge blanket (UASB) reactor performance. Biotechnol. Bioeng. 77, 806–814. https://doi.org/10.1002/bit.10168.

Spyridonidis, A., Skamagkis, T., Lambropoulos, L., Stamatelatou, K., 2018. Modeling of anaerobic digestion of slaughterhouse wastes after thermal treatment using ADM1. J. Environ. Manage. 224, 49–57. https://doi.org/10.1016/j.jenvman.2018.07.001.

Stojić, A., Stanić, N., Vuković, G., Stanišić, S., Perišić, M., Šoštarić, A., Lazić, L., 2019. Explainable extreme gradient boosting tree-based prediction of toluene, ethylbenzene and xylene wet deposition. Sci. Total Environ. 653, 140–147. https://doi.org/10.1016/j.scitotenv.2018.10.368.

Strömberg, S., Nistor, M., Liu, J., 2015. Early prediction of Biochemical Methane Potential through statistical and kinetic modelling of initial gas production. Bioresour. Technol. 176, 233–241. https://doi.org/10.1016/j.biortech.2014.11.033.

Tan, B., Zhang, J., Wang, L., 2011. Semi-supervised Elastic net for pedestrian counting. Pattern Recognit. 44, 2297–2304. https://doi.org/10.1016/j.patcog.2010.10.002.

Tan, H.M., Gouwanda, D., Poh, P.E., 2018. Adaptive neural-fuzzy inference system vs. anaerobic digestion model No.1 for performance prediction of thermophilic anaerobic digestion of palm oil mill effluent. Process Saf. Environ. Prot. 117, 92–99. https://doi.org/10.1016/j.psep.2018.04.013.

Thomsen, S.T., Spliid, H., Østergård, H., 2014. Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass. Bioresour. Technol. 154, 80–86. https://doi.org/10.1016/j.biortech.2013.12.029.

Thorin, E., Olsson, J., Schwede, S., Nehrenheim, E., 2018. Co-digestion of sewage sludge and microalgae – Biogas production investigations. Appl. Energy 227, 64–72. https://doi.org/10.1016/j.apenergy.2017.08.085.

Tong, X., Smith, L.H., McCarty, P.L., 1990. Methane fermentation of selected lignocellulosic materials. Biomass 21, 239–255. https://doi.org/10.1016/0144-4565(90)90075-U.

Tsapekos, P., Alvarado-Morales, M., Kougias, P.G., Konstantopoulos, K., Angelidaki, I., 2019. Co-digestion of municipal waste biopulp with marine macroalgae focusing on sodium inhibition. Energy Convers. Manag. 180, 931–937. https://doi.org/10.1016/j.enconman.2018.11.048.

Uhlenhut, F., Schlüter, K., Gallert, C., 2018. Wet biowaste digestion: ADM1 model improvement by implementation of known genera and activity of propionate oxidizing bacteria. Water Res. 129, 384–393. https://doi.org/10.1016/j.watres.2017.11.012.

Valenti, F., Porto, S.M.C., Selvaggi, R., Pecorino, B., 2018a. Evaluation of biomethane potential from by-products and agricultural residues co-digestion in southern Italy. J. Environ. Manage. 223, 834–840. https://doi.org/10.1016/j.jenvman.2018.06.098.

Valenti, F., Zhong, Y., Sun, M., Porto, S.M.C., Toscano, A., Dale, B.E., Sibilla, F., Liao, W., 2018b. Anaerobic co-digestion of multiple agricultural residues to enhance biogas production in southern Italy. Waste Manag. 78, 151–157. https://doi.org/10.1016/j.wasman.2018.05.037.

Vavilin, V.A., Rytov, S.V., Lokshina, L.Y., Pavlostathis, S.G., Barlaz, M.A., 2003. Distributed model of solid waste anaerobic digestion: effects of leachate recirculation and pH adjustment. Biotechnol. Bioeng. 81, 66–73. https://doi.org/10.1002/bit.10450.

Vivekanand, V., Mulat, D.G., Eijsink, V.G.H., Horn, S.J., 2018. Synergistic effects of anaerobic co-digestion of whey, manure and fish ensilage. Bioresour. Technol. 249, 35–41. https://doi.org/10.1016/j.biortech.2017.09.169.

Wang, X., Bai, X., Li, Z., Zhou, X., Cheng, S., Sun, J., Liu, T., 2018. Evaluation of artificial neural network models for online monitoring of alkalinity in anaerobic co-digestion system. Biochem. Eng. J. 140, 85–92. https://doi.org/10.1016/j.bej.2018.09.010.

Xu, F., Li, Y., Wang, Z.-W., 2015. Mathematical modeling of solid-state anaerobic digestion. Prog. Energy Combust. Sci. 51, 49–66. https://doi.org/10.1016/j.pecs.2015.09.001.

Xu, F., Wang, Z.-W., Tang, L., Li, Y., 2014. A mass diffusion-based interpretation of the effect of total solids content on solid-state anaerobic digestion of cellulosic biomass. Bioresour. Technol. 167, 178–185. https://doi.org/10.1016/j.biortech.2014.05.114.

Zhang, Y., Piccard, S., Zhou, W., 2015. Improved ADM1 model for anaerobic digestion process considering physico-chemical reactions. Bioresour. Technol. 196, 279–289. https://doi.org/10.1016/j.biortech.2015.07.065.

Zhao, X., Li, L., Wu, D., Xiao, T., Ma, Y., Peng, X., 2019. Modified Anaerobic Digestion Model No. 1 for modeling methane production from food waste in batch and semi-continuous anaerobic digestions. Bioresour. Technol. 271, 109–117. https://doi.org/10.1016/j.biortech.2018.09.091.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B (Statistical Methodol. 67, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.