QUARTERLY PROGRESS REPORT

December 2023 – February 2024

PROJECT TITLE: Evaluating and Optimizing the Value of Anaerobic Digestion of Food Waste using Sensitivity Analysis and Machine Learning

PRINCIPAL INVESTIGATOR(S):

PI: Ana Martin-Ryals, Assistant Professor, admartin@ufl.edu, 352-294-6708

Co-PI: Nikolay Bliznyuk, Associate Professor of Statistics, nbliznyuk@ufl.edu, 352-294-6734

Affiliation: Agricultural & Biological Engineering (ABE), University of Florida, Gainesville, FL

PROJECT WEBSITE: Hinkley Project - UF/IFAS Agricultural and Biological Engineering

WORK ACCOMPLISHED DURING THIS REPORTING PERIOD:

Work accomplished during this reporting period included continuing work on Objectives 1 and 2 as described below.

Objective 1 is to establish a dataset linking feedstock characteristics, anaerobic digestion operating parameters, and methane production. During the previous reporting period, the team carried out an initial review of published experimental studies investigating anaerobic digestion of post-consumer food waste, and extracted data on feedstock characteristics (total solids, volatile solids, carbon to nitrogen ratio, and macronutrient content), operating parameters (volume, retention time, organic loading rate, temperature) and performance metrics (methane production). That review yielded a final data set, Dataset 1, consisting of 100 records, which is summarized in Table 1 and visualized as Figure 1. The greatest challenge faced during the initial literature review was identifying references that had detailed feedstock characterization data. During the current reporting period, further literature review was carried out, with a focus on identifying records with detailed feedstock characteristics, specifically macromolecule content. An additional 202 records were identified during the second round of literature review, Dataset 2, which are summarized in Table 1 and visualized as Figure 2.

One notable difference between the two datasets is that Dataset 2 consisted of more biomethane potential (BMP) experimental results compared to Dataset 1. BMPs are usually carried out using a small volume of material 0.25 – 2 L, to determine the expected methane production potential of a given feedstock. The literature reviewed for Dataset 2 consisted of many BMP tests that had been carried out to specifically evaluate the impact of variable feedstock macromolecule content on methane production potential. Therefore, most of the data in Dataset 2 (179 out of 202 records) represent experimental work that was carried using a volume of 2 L or less, compared to 16 out of 100 records in Dataset 1. Digester volumes report in Dataset 1 ranged from 0.2-900,000 L, while digester volumes in Dataset 2 ranged from 0.5-35 L.

The effect of scale may explain some differences observed between the two datasets in terms of correlation between predictor variables (feedstock characteristics and operating parameters) and methane production. Simple linear regressions for the various predictor variables and experimental methane production are shown in Figures 1, 2, and 3 for Dataset 1, Dataset 2, and the Combined Dataset, respectively. In Dataset 1, protein content and retention time were observed to have a strong positive correlation with methane production. However, this relationship was not observed in Dataset 2 or the Combined Dataset. Therefore, the impact of scale (i.e., digester volume) on methane production and process predictability will be interrogated with more scrutiny during the next reporting period.

In general, the combined dataset highlights the range and potential variability of methane production that may be achieved during anaerobic digestion of post-consumer food waste. Excluding outliers, reported values for methane production ranged from 45-801 mL/g VS added. This translates to a potential energy recovery of 0.38-6.8 MMBtu/ton of wet food waste (summarized in Table 1 and Figure 4). The economic implications of this variability will be considered in the economic and sensitivity analysis, to begin in the next reporting period.

Table 1. Summary of data sets obtained from the first and second quarter literature reviews consisting of feedstock, operational and performance parameters for anaerobic digestion of post-consumer food waste.

	Dataset 1 *(n=100)*			**Dataset 2** *(n=202)*			**Combined Dataset** *(n=302)*		
	Avg.	Median	Range	Avg.	Median	Range	Avg.	Median	Range
Feedstock Characteristics									
Total Solids (%)	16.9	12.3	1.2-40	36.6	27.4	1.34-99	30.1	24	1.2-99
Volatile Solids (%)	14.8	11.0	0.88-35	32.2	24	0.96-99	26.5	20.0	0.88-99
VS/TS (%)	88.2	88.2	75-95	87.8	89.2	12-99.9	87.9	88.2	12-99.9
Carbs (% TS)	36.6	42.9	0-73	45.7	42	0-99	42.7	42	0-99
Protein (% TS)	25.7	19	5-87	22.5	17.9	0-75	23.5	18	0-87
Lipid (% TS)	17.4	19.5	0.3-37	24.3	16.8	0-100	22.0	17.6	0-100
Carbon:Nitrogen	32.8	15.7	9-223	18.8	17.9	2-49	23.1	17.6	2-223
Operating Parameters									
Digester Volume (L)	11,799	85	0.2-900,000	3.75	1	0.5-35	3,909	1	0.2-900,000
Retention Time (days)	73.5	31.6	1-431	40.6	30	15-150	55.9	30	1-431
Organic Loading Rate (g VS/L-d)	3.71	2.52	0.03-13.5	1.0	0.625	0.1-2.8	2.5	1.5	0.03-13.5
Temperature (°C)	42.7	40	20-55	41.6	40	35-55	41.9	40	20-55
Performance									
Methane Yield (L/kg-VS added)	366.8	405.5	0-711	441.7	434	30-1476	416.9	424.9	0-1476
Energy Recovery (MMBtu/US ton)	12.5	13.8	0-24	15.1	14.8	1.0-50	3.5	3.6	0-12.5

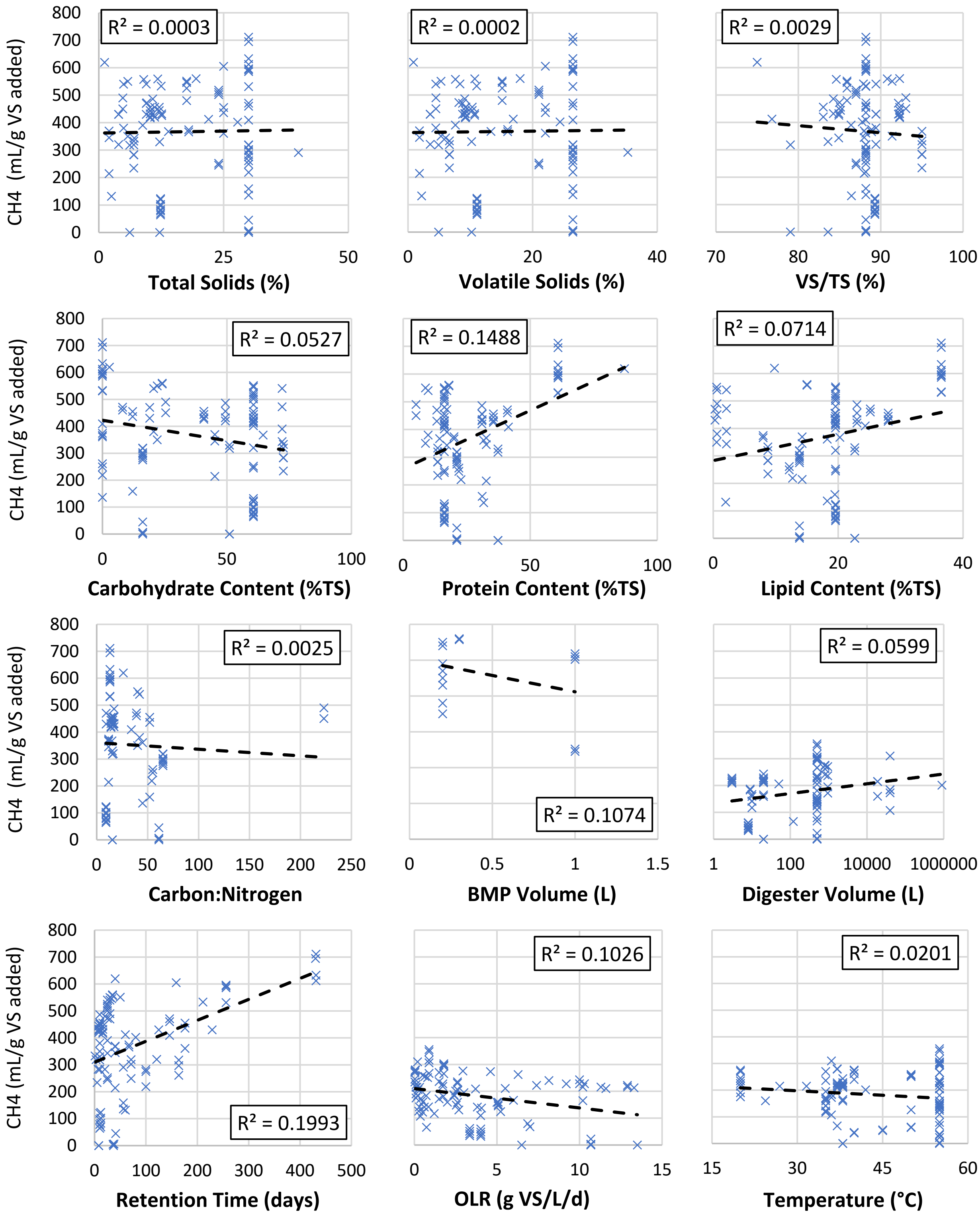

Figure 1. Overview of **Dataset 1**, which was collected during the first quarter literature review. Reported feedstock characteristics and process parameters are shown on the x-axis and correlated to reported methane production on the y-axis.

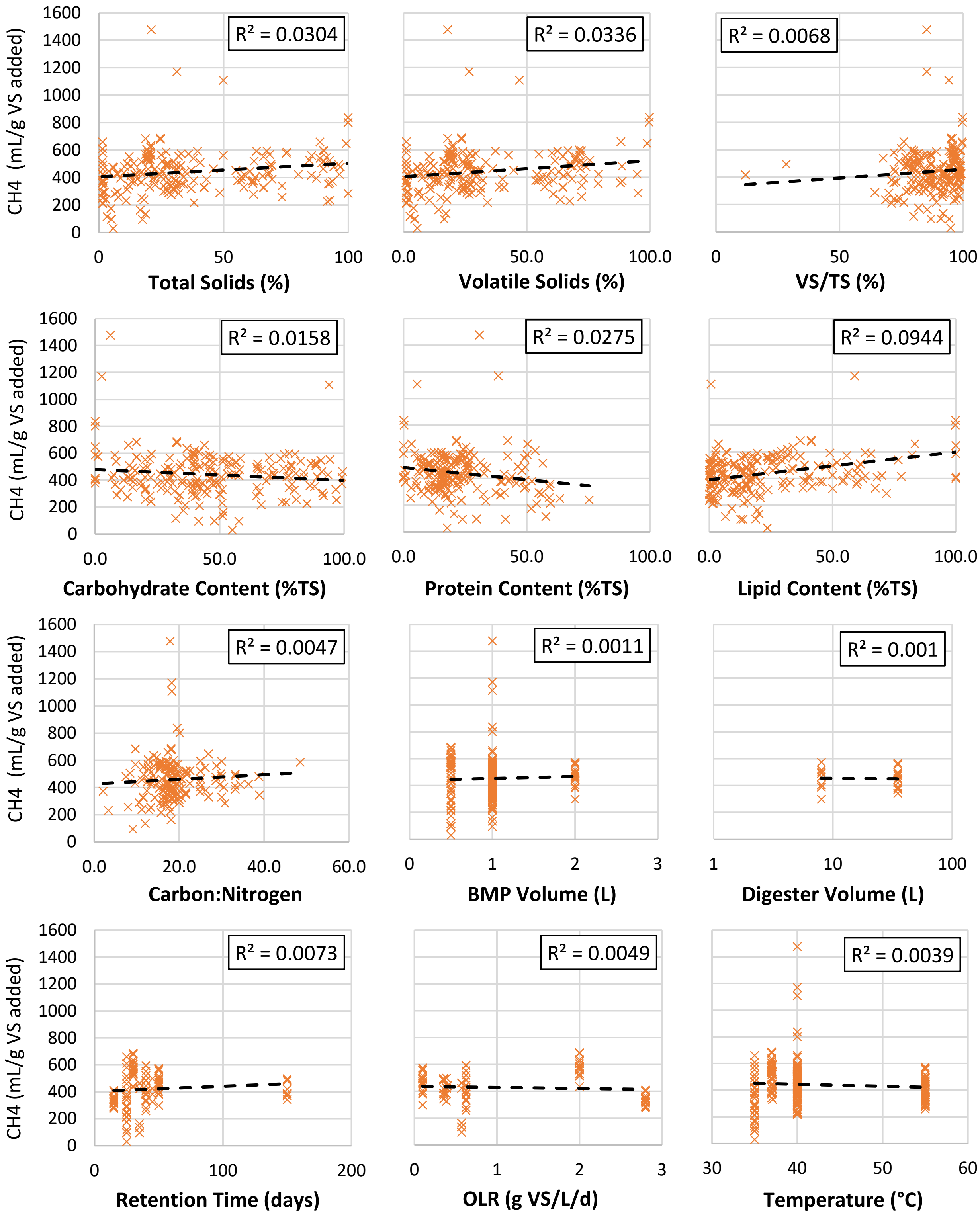

Figure 2. Overview of **Dataset 2**, which was collected during the second quarter literature review. Reported feedstock characteristics and process parameters are shown on the x-axis and correlated to reported methane production on the y-axis.

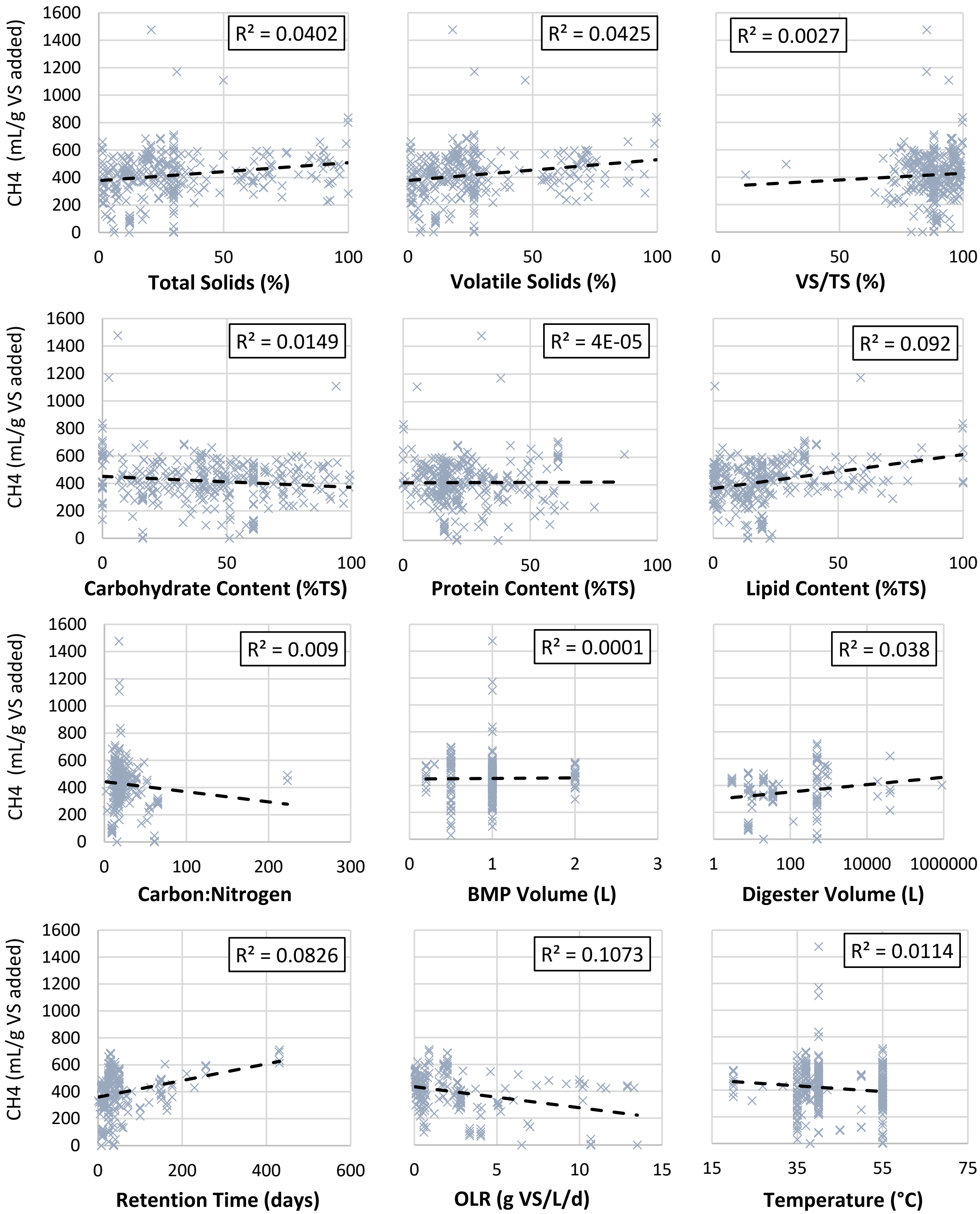

Figure 3. Overview of **Combined Dataset**, consisting of data collected during first and second quarter literature reviews. Reported feedstock characteristics and process parameters are shown on the x-axis and correlated to reported methane production on the y-axis.

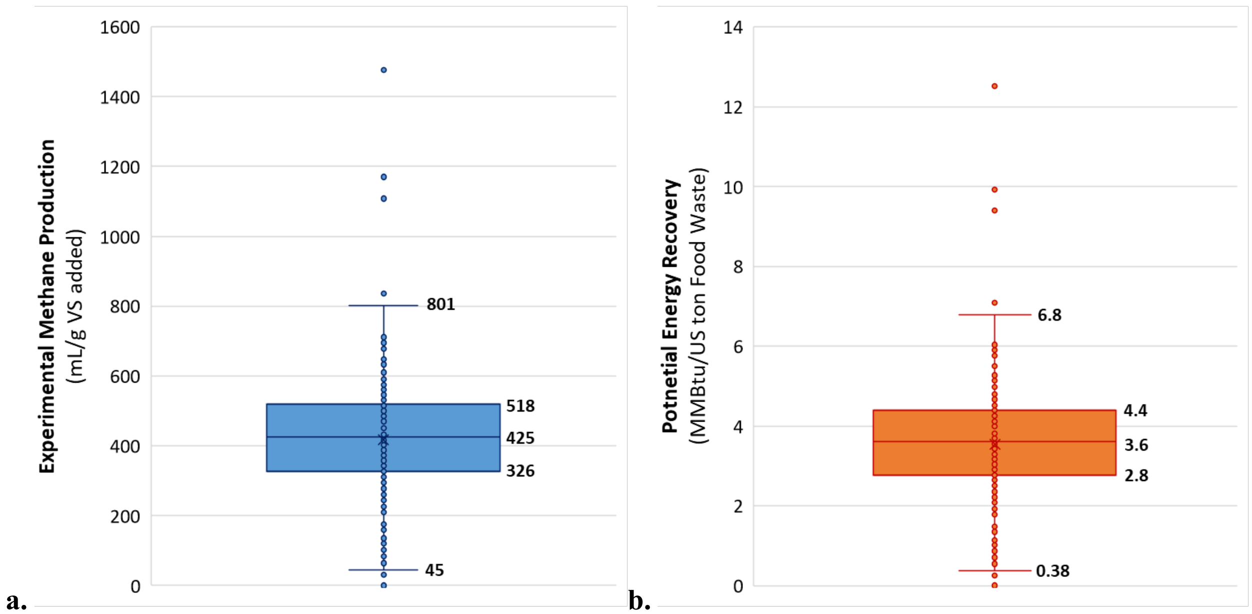

Figure 4. Overview of ranges for (a) reported experimental methane production as mL methane per gram of volatile solids added, and (b) calculated potential energy recovery as MMBtu per ton of wet food waste treated.

Objective 2 is to develop a data-driven, machine-learning based model that can predict anaerobic digestion performance as a function of feedstock characteristics and operating conditions.

During the current reporting period, Dataset 1 was used to carry out the initial machine-learning model development. 10-fold cross validation was integrated with random forest regression trees to build a preliminary means of predicting methane yield from feedstock characteristics and digester operating parameters. The complete dataset was split into training and testing datasets for pre-screening by random forest. The RandomForestRegressor method available in sk.learn (*4*, *5*) was fit to all four datasets generated from imputation and encoding for feature importance analysis. This initial analysis identified organic loading rate as the most important predictor variable, followed by feedstock protein content. This was consistent with was observed through simple linear regression of Dataset 1 as shown in Figure 1. Volatile solids content and carbon-nitrogen ratio had similar importance, with a gradual decline in variable importance thereafter. Surprisingly, no other digester operating conditions apart from organic loading rate were ranked in the top six contributors to the RF model's prediction accuracy (Figure 5).

The initial prediction accuracy was relatively low with a Mean Squared Error (MSE) greater than 7000. During the next reporting period, we will continue to improve the machine-learning model utilizing the expanded data set and additional algorithms beyond random forest. We will also investigate the impact of scale and co-variants on performance predictability. Finally, we will also continue reviewing the literature to expand our dataset for anaerobic digestion of post-consumer food waste and to include co-digestion of food waste and other relevant feed stocks (wastewater biosolids, and fats, oils, and grease), for further machine-learning model development and validation.

We plan to hold a second TAG meeting in the next month to review project progress thus far, and to discuss the next phases of the project which will include economic and sensitivity analysis. During the next reporting period, we will establish the economic and sensitivity analysis framework which will be used to evaluate the impact of food waste compositional variability on carbon intensity and economic viability of a full-scale digester process.

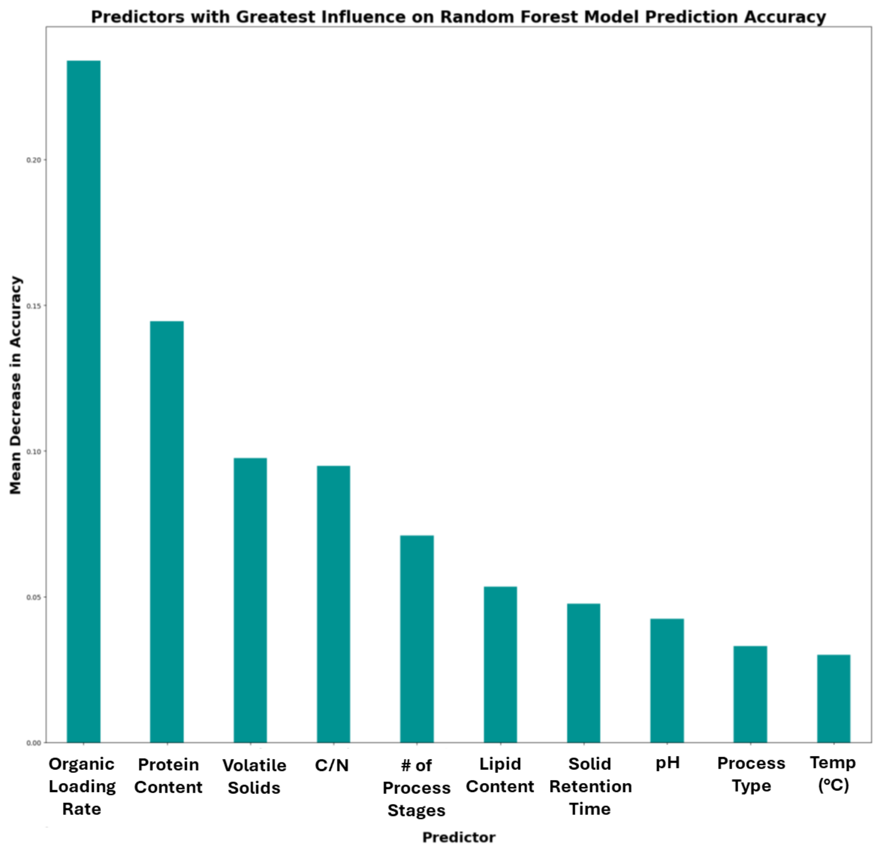

Figure 5. Overview of most important feature contributors to Random Forest model prediction power.

TAG MEETINGS:

No TAG Meetings were held during this reporting period. PI Martin-Ryals did meet with one TAG member, Soren Jorgensen, Senior VP of Global Business Development of Bigadan A/S on January 16, 2024, to discuss the possibility of getting full-scale digester data for use in model.

METRICS REPORTING:

1. Summary of input provided by the TAG during this period.

PI Martin-Ryals met with Soren Jorgensen of Bigadan and one of his colleagues to discuss the possibility of getting full-scale digester data for use in developing the predictive model. Bigadan is in the process of organizing their operational data, and it was indicated that they will likely be able to provide PI Martin-Ryals with a data set when it's available. This data set will be used to test, validate, and update the predictive model.

2. Publications resulting from **THIS** Hinkley Center project.

    None

3. Research presentations resulting from (or about) **THIS** Hinkley Center project.

- Martin-Ryals, Ana. "Advancing carbon, water and nutrient recovery from anaerobic digestion of food waste". Manure Lunch Seminar Series. Alachua, Florida, ***Rescheduled*** from Dec 1, 2023, to March 1, 2024.

4. List who has referenced or cited your publications from this project.

    None

5. How have the research results from **THIS** Hinkley Center project been leveraged to secure additional research funding? What additional sources of funding are you seeking or have you sought?

    No change – no additional funding has been secured at this time.

6. What new collaborations were initiated based on **THIS** Hinkley Center project?

    No change

7. How have the results from **THIS** Hinkley Center funded project been used (not will be used) by the FDEP or other stakeholders?

    None

PICTURES:

None at this time.