Project check-in # 1: Preliminary findings Report

1. Preliminary Analysis

Data Description: The dataset is of the chemical characterization of 88 food waste substrates with specific importance on the principal nutrient components of the substrate i.e. carbohydrates, proteins, and lipids (fats). A snapshot of the dataset is shown below (figure 1). Each food waste item, hereafter referred to as 'substrate', is assigned to a group based on rough commonalities of the substrate's characteristics, or what would be generally considered a food group in modern society. For example, products containing milk are all assigned to the group 'Dairy Products' while products containing animal flesh (non-fish) are assigned the group 'Meat Products'. There are 15 groups across the 88 food substrates.

In the proposed analysis we will utilize the MATLAB-based Anaerobic Digestion Model #1 (ADM1) to estimate the amount of biogas (methane) that can be produced from each food substrate. The model requires that the mass of the principal nutrients be provided, which can be determined from the characteristics listed in the table using a process described in the Data Preparation section.

```
## Rows: 88
## Columns: 17
## $ num                 <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 1…
## $ G.num               <chr> "DP.1", "DP.2", "DP.3", "DP.4", "DP.5", "DP.6", "FOG…
## $ Food.Wastes         <chr> "Cheese", "Milk", "Baby milk", "Yogurt/yogurt drink"…
## $ Food.Wastes.Clean   <chr> "Cheese", "Milk", "Baby_milk", "Yogurt_yogurt_drink"…
## $ Group               <chr> "DP", "DP", "DP", "DP", "DP", "DP", "FOG", "FOG", "F…
## $ pH                  <dbl> 5.93, 6.76, 7.15, 4.30, 6.37, 7.23, 3.50, 3.20, 3.01…
## $ TS.Perc             <dbl> 49.86, 10.96, 11.38, 30.90, 39.56, 88.58, 75.40, 99.…
## $ VS.Perc             <dbl> 42.63, 10.56, 10.89, 30.30, 37.27, 88.22, 71.63, 98.…
## $ TOC                 <dbl> 29.80, 5.39, 4.41, 12.80, 51.00, 74.90, 86.10, 74.63…
## $ TKN                 <dbl> 2.90, 1.66, 1.95, 0.91, 1.54, 0.51, 0.35, 0.22, 0.69…
## $ Fat.Perc            <dbl> 23.20, 15.60, 23.80, 5.10, 19.50, 83.40, 100.00, 100…
## $ Protein.Perc        <dbl> 18.50, 33.40, 10.35, 14.30, 4.50, 3.27, 0.00, 0.00, …
## $ Carbohydrate.Perc   <dbl> 58.30, 51.00, 65.85, 80.60, 76.00, 13.33, 0.00, 0.00…
## $ TP                  <dbl> 1.14, 1.70, 1.06, 0.70, 1.09, 2.30, 0.01, 0.00, 0.00…
## $ TK                  <dbl> 0.17, 1.10, 1.69, 0.14, 0.60, 1.50, 0.00, 0.01, 0.00…
## $ C.N.Mixture         <dbl> 17.38, 3.32, 17.30, 14.06, 17.31, 17.38, 48.50, 26.9…
## $ BMP                 <dbl> 561.0, 231.0, 315.0, 450.0, 591.0, 660.0, 586.0, 648…
```

Figure 1. 'glimpse' of dataset structure, variables, and values

Initial Analysis:

Carbohydrate, Proteins, Fats Percentages

To better understand the variability of each of the principal nutrients in the dataset, we developed histograms based on their percentages (figure 2). We see that carbohydrates generally dominate the overall mass of the substrates with the majority of foods being >40% carbohydrates by mass of Volatile Solids (VS). Proteins in general occupy a smaller percentage of the VS mass with the majority of foods having <20% proteins. In comparison, more than half the food substrates have 0-10% fats marking them as the lowest represented nutrient by mass, with notable exceptions that are fat dominant such as oils, grease, and fatty dairy products.

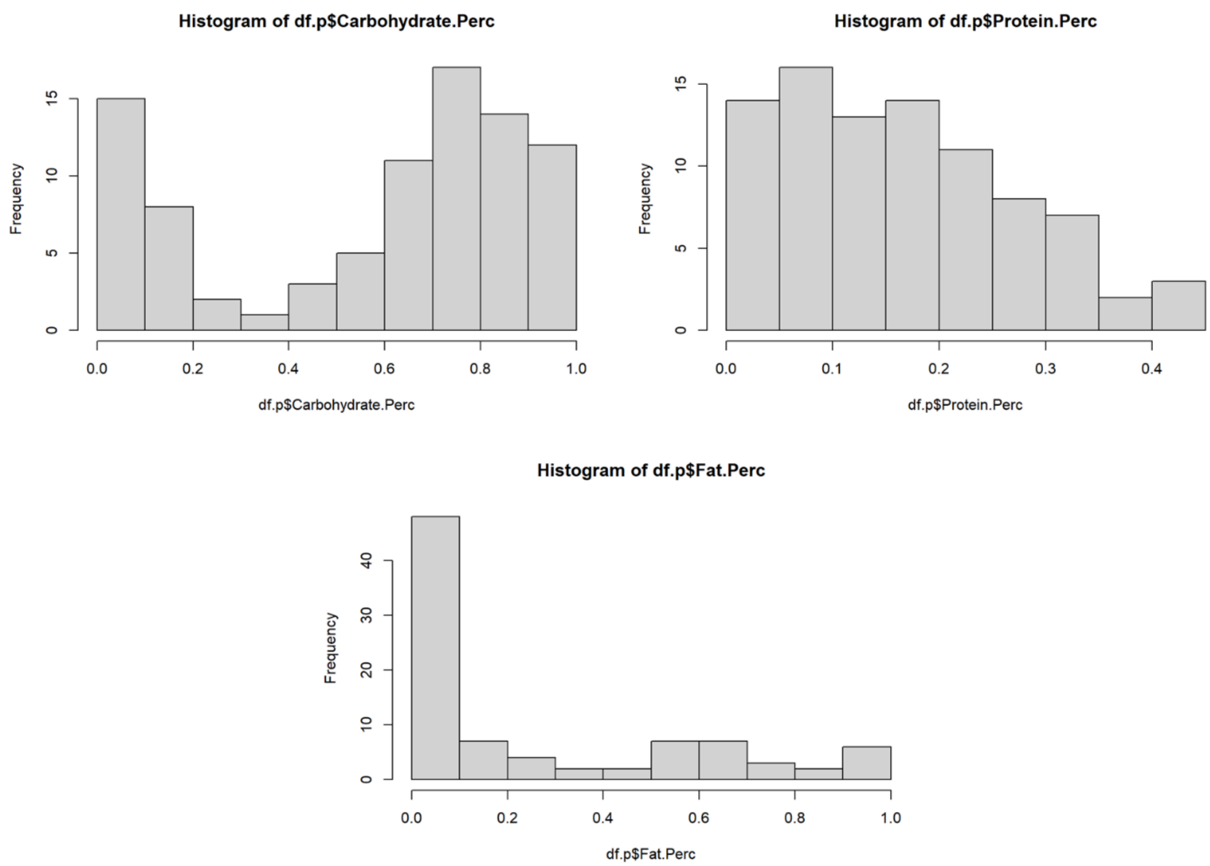

Figure 2. Histograms of the principal nutrient (carbohydrate, proteins, and fats) fractions

To obtain more resolution on the overall fractionation of each of the principal nutrients we also plotted them in 3D space with each axis representing a percentage of the three nutrients as seen below (figure 3). Due to the limitations of the program's color palette it is difficult to infer distinctive groupings of the foods. However, we can see that the greatest percentage of food substrates is in the bottom left quadrant (blue circle, non-statistically based). These substrates possess high carbohydrates, mid proteins, and low fats, which corresponds to the observations of the histograms. A secondary cluster in the upper right (red circle, non-statistically based) describing low carbohydrates, mid proteins, and high fats is also in the spread.

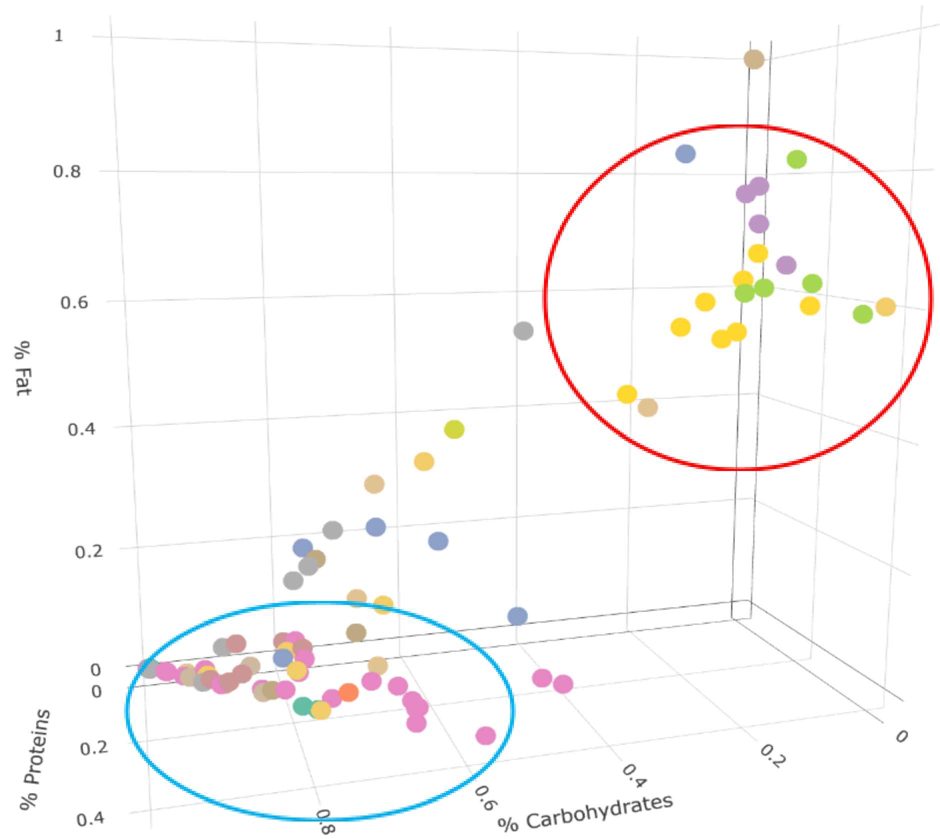

Figure 3. 3D scatterplot of percentages of carbohydrates (x), proteins (y), and fats (z). Points are colored by group (group color not shown to avoid confusion)

Mono-feed Biogas Output

The primary output that will address our research questions is the production of methane biogas; and we are interested in both the rate and the overall quantity produced. Using the default parameters of the ADM1 model we ran an initial test under a batch feed (sequential i.e. a feed every n days) and a continuous feed regime. The model outputs the biogas as a rate (liters of biogas/day) over the reactor time (measured in days). The output for the two studies is below grouped by food category (table below) (table 1).

Table 1. Food Classification used for grouping like food substrates

Food Classification	Symbol Used
Dairy Product	DP
Fats, oils, and grease	FOG
Ice Cream	IC
Fruit and Vegetable	FAV
Confectionary (canned good)	CCG
Cereals and Cereals Products	CP
Bakery Wares	BW
Meat and Meat Products	MP
Fish and Fish Products	FP
Eggs and Egg Products	EP
Sweeteners and Sweet Goods	SSG
Sauces, Spices, and Soups	SS
Beverages	BEV
Ready to eat food or restaurant waste	REWE
Other Expired Food	OT

Batch Feeding

The batch feeding consisted of feeds of 1000 grams of total mass for all substrates. The results are output as a rate (q_gas Liters/day) each day of the 100 day reactor. The output of the batch feeding (occurring day 0, 10, 30, 50, and 70) shows a consistent pattern across all food substrates but the magnitude varies both within and across food groups (figure 4). The most consistently highly productive food groups are cereal products (CP) and bakery wares (BW) likely owing to their high percentage of carbohydrates. Meat products (MP), sweeteners and sweet goods (SSG), and ready to eat foods (REWE) are also highly productive but have greater variability across the foods within that category. Future analysis will investigate the chemical differences between foods within each group to identify possible associations explaining higher or lower production. An important aspect to remember for this analysis is that it utilized a base total mass (includes water and inerts) for each food item, which while more realistic in that food waste is typically presented as a bulk without consideration to the volatile solids, this does present great variability in the relative amounts of available material. For example, 1000 grams of bread has little water or inerts thus the overall VS amount is higher than butter which contains a larger amount of moisture. Normalizing the feed inputs will be another aspect of consideration for future analysis.

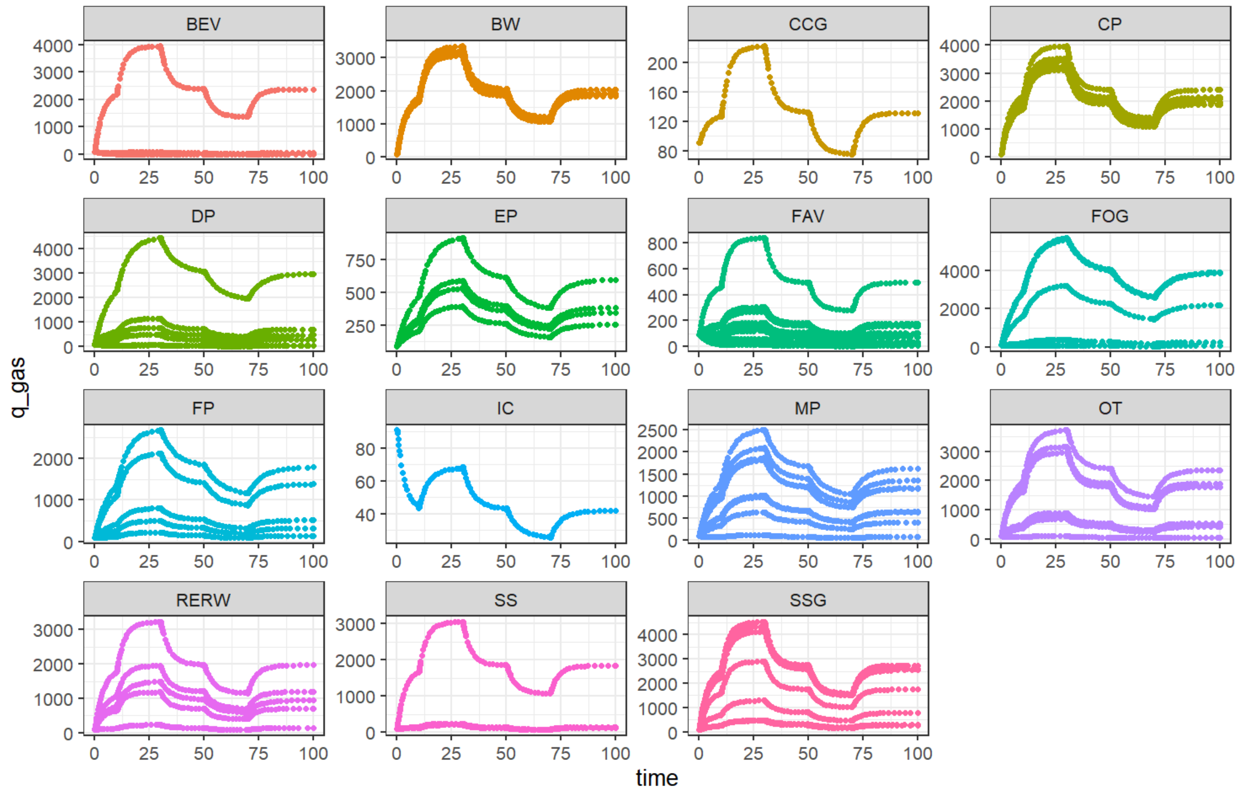

Figure 4: ADM1 Model biogas production output for batch experiment measured in Liters/Day over 100 days. Plots are divided by food classification group

Continuous Feed

For the continuous feed analysis, the mass of feed was normalized to be a consistent total volatile solid amount of 300 grams and presented to the model as a constant daily rate of 300 grams of VS per Liter per day. This was done over the 100-day reactor timeline. The results are output as a rate (q_gas Liters/day) each day of the 100-day reactor (figure 5). Unlike the batch feed the reactors output an exponential growth pattern with a plateau. This can be thought of as the maximum rate of production for each food substrate under the given conditions. There appears to be lower variability between and across food groups than was observed in the batch reactor, which may be due to the fact that the substrate is not limited as it is in the batch reactors. The general range for production rate for all food substrates is between 700-1250 Liters of biogas/day which may be lower than the batch feed but does incorporate a lower amount of substrate and shows stability which could be more desirable overall in a real world scenario.

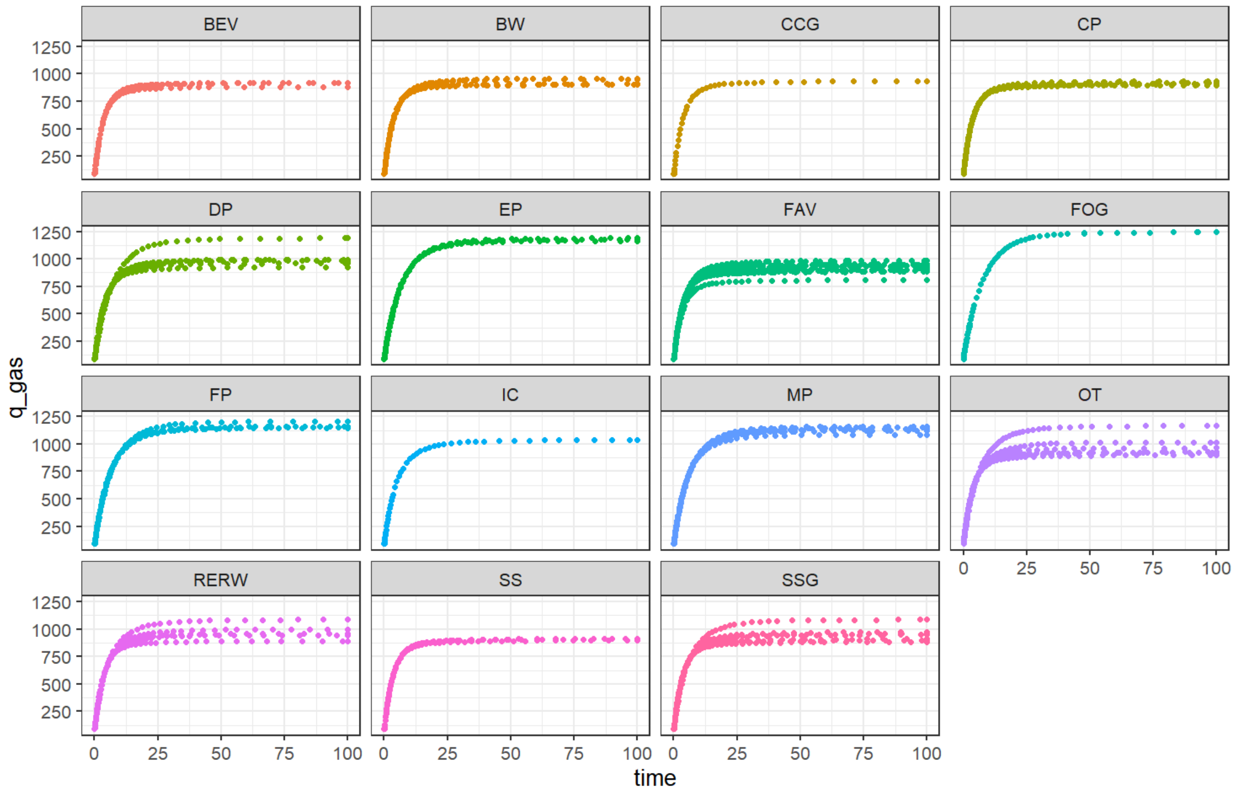

Figure 5: ADM1 Model biogas production output for continuous feed experiment measured in Liters/Day over 100 days. Plots are divided by food classification group

Data Issues

We currently have no issues with the data itself, our primary considerations that will require rectification are in how we treat the data, specifically how we normalize it or scale to allow us to make direct comparisons. For example, in our preliminary batch and continuous feed analysis we utilized a bulk feed based on total mass for the batch analysis and a feed rate based on total VS in the continuous analysis. Neither method is necessarily incorrect; they simply represent different scenarios and address different questions. However for the future of the analysis it would be useful to be able to compare the batch and continuous feed scenarios directly, requiring a common unit.

One way we may address this is to work from a total system mass, which is to say that over the 100 day period the same total mass of VS is introduced into the system whether it is done in pulses in the batch or over time in the continuous feed.

2. Approach

Research Questions

What is the correlation between the principal nutrients of a given substrate and its biogas production capacity?

Are there consistent factors (i.e. ratios of the principal nutrients) that amount to high biogas production?

What combinations of substrates achieve the highest production of biogas and how does this compare to individual substrates?

How does the biogas production of bulk feeds compare to the biogas production of continuous feeds for equivalent masses of food waste?

Data Preparation

Due to some differences in formatting between the initial data set and the ADM1 feed inputs, some intermediate data preparation was required to prepare the food waste data for ADM1. As we are using the most simplified version of ADM1 (ADM1-R4), which only considers the principal nutrients, these are the only values from the initial data set that need to be formatted for input into ADM1. The amount of substrate material used as input is determined starting with the total solid percentage (TS.Perc). This is the mass of material that is not water. From the total solid percentage we then take the volatile solids percentage (VS.Perc) which is the fraction that can engage in anaerobic digestion. From the volatile solids we can then fraction out the principal nutrients (carbohydrates, proteins, and fats) by their listed percentages. The final mass of each of these are formatted into the input file and fed into ADM1 runs.

ADM1 accepts feed inputs as a single csv file, representing a schedule of inputs. Each row of the feed inputs file represents a single feeding event at some point during the 100 day reaction, with the 'time' value of each row encoding the feed day. The inputs can also be multiplied by an arbitrary factor represented by 'q_in'. This allows multiple differently-sized inputs of the same food item to differ only in their input 'time' and 'q_in'. To generate these files, we use pandas to iterate over the rows of the data, extracting the principal nutrient percentage values. For each food item, we multiply each principal nutrient percentage by the total input VS mass to get the input mass of each principal nutrient. For schedules with combinations of food items, we multiply half the total VS mass by each substrate's principal nutrient percentages and add the results. These values are then formatted to match the feed inputs file, with the appropriate 'time' and 'q_in' values for each feed schedule (batch or continuous).

Once we've generated a .csv file for each feeding schedule for each food item, we're able to provide these files as input to the ADM1 simulation. The results of these simulations contain the rate of biogas production for each day. Because the biogas output is given as a rate of production rather than total volume produced, and because the numerical integrator used takes uneven time steps, we must compute the total biogas produced. This is done by applying the linear trapezoidal rule to the gas production rate time series data to estimate the area under the curve.

Analysis Plan

Once we've determined the capacity of each monofeed and pair-feed substrate to produce biogas, we plan to use a dimensionality reduction technique such as Principal Component Analysis (or some nonlinear method, if the data turns out to not be linearly separable) to identify how much each of the principal nutrients contributes to the biogas production capacity of a substrate. We will also visualize the results by creating a 3D scatterplot such as the one in the Initial Analysis section, with x, y, and z representing the percentages of each principal nutrient in the VS, and encoding the biogas productivity in the color of the point.

3. Next Steps

Additional Data or Methods

We do not anticipate needing additional data beyond the biogas productive capacity data we're in the process of generating.

Scope Adjustments

As of now we have not adjusted the scope, though as we begin to analyze the data and address our research questions it is likely that additional questions will arise and interesting trends may present themselves that warrant further analysis.

4. Literature Review

Citation: De Jonge, N., Davidsson, Å., la Cour Jansen, J., & Nielsen, J. L. (2020). Characterisation of microbial communities for improved management of anaerobic digestion of food waste. *Waste Management*, *117*, 124-135. **Summary:** Unlike other conventional waste streams such as wastewater or manure, food waste streams are physiochemically variable and complex, potentially resulting in a diverse microbial community that is unique to each food waste stream. This would make predicting and modelling anaerobic digestion difficult. This paper researched 18 food waste based reactors under different operational parameters and investigated the microbial communities. They found that indeed, no one microbial community was consistent in food waste however, food waste was more plentiful in

key microbial members that aid in anaerobic digestion such as syntrophic acetate oxidizing microbes.

Relevance: Microbes are the primary driver of anaerobic digestion in a bioreactor and their activity defines the production of biogas. Most models utilize a general collection of microbial processes to estimate biogas production which is applicable to waste streams that are consistent. However, food waste streams are highly variable thus understanding the microbial composition is integral to accurately predicting biogas production and yet understudied. As this study shows, there actually may be highly beneficial (in regards to anaerobic digestion) microbes contained in certain food wastes that could be optimized for better biogas production

Further References:

Lee, J., Kim, E., Han, G., Tongco, J.V., Shin, S.G., Hwang, S., 2018. Microbial communities underpinning mesophilic anaerobic digesters treating food wastewater or sewage sludge: A full-scale study. Bioresour. Technol. 259, 388–397. https://doi.org/10.1016/j.biortech.2018.03.052. Lee, J., Shin, S.G., Han, G., Koo, T., Hwang, S., 2017. Bacteria and archaea communities in full-scale thermophilic and mesophilic anaerobic digesters treating food wastewater: Key process parameters and microbial indicators of process instability. Bioresour. Technol. 245, 689–697. https://doi.org/10.1016/j. biortech.2017.09.015.

Citation:

Dhar, H., Kumar, P., Kumar, S., Mukherjee, S., & Vaidya, A. N. (2016). Effect of organic loading rate during anaerobic digestion of municipal solid waste. *Bioresource Technology*, *217*, 56-61.

Summary: The organic loading rate (rate of VS introduced into a reactor) is highly determinant of the overall reactor function and can be optimized on a per reactor and per substrate bases to achieve optimum biogas production. In this paper different loading rates of the organic fraction of municipal solid waste were trialed in a laboratory-scale digester. They found that increased loading rate resulted in increased biogas production however the increases in rate quickly decreased as loading rate increased demonstrating the system maximum.

Relevance: Loading rate is a key parameter in reactor production and can easily be over or under estimated especially when the substrate is less known. Having empirically defined benchmark values will aid us by providing benchmarked realistic values that can be trailed in the ADM1 model to give us more realistic results

Further References:

Chen, Y., Cheng, J. J., & Creamer, K. S. (2008). Inhibition of anaerobic digestion process: a review. *Bioresource technology*, *99*(10), 4044-4064.

Razaviarani, V., & Buchanan, I. D. (2014). Reactor performance and microbial community dynamics during anaerobic co-digestion of municipal wastewater sludge with restaurant grease waste at steady state and overloading stages. *Bioresource technology*, *172*, 232-240.

Citation:

Agency USEP (2014) Food Waste Management in the United States. Office of Resource Conservation and Recovery. Available at https://www.epa.gov/sites/production/files/2016-12/documents/food_waste_management_2014_12082016_508.pdf. Accessed 21st Jan 2021

Summary: This report is a summary of the food waste management practices in the United States circa 2014. It outlines the mechanisms of food waste reduction from donation to composting by state and by year from 2004-2014. It also touches upon policies that have impacted food waste management in the decade of analysis.

Relevance: While a decade old it still touches on the fundamentals of the purpose of this research which is to reduce the amount of terminal food waste (transported to landfill without additional purpose). Thinking about this research and analysis in a larger context is valuable in understanding the impact that alternative food waste protocols can have in the overall waste stream and can also identify specific states or locals where outcomes of this research may be most valuable.

Further References:

U.S. EPA Region 1. 2013. Anaerobic Digestion of Food Waste in New England Summer 2013 Report. Revised 2-9-2015.
http://www.ct.gov/deep/lib/deep/compost/compost_pdf/ad_of_food_waste_in_new_england.pdf

Platt, B.; Goldstein, N. 2014. State of Composting in the U.S. BioCycle 55(6): 19. http://www.biocycle.net/2014/07/16/state-of-composting-in-the-u-s/. Accessed March 2016

Citation:

Meegoda, J. N., Li, B., Patel, K., & Wang, L. B. (2018). A review of the processes, parameters, and optimization of anaerobic digestion. *International journal of environmental research and public health*, *15*(10), 2224.

Summary: There are a multitude of processes, parameters, and design considerations that must be taken into account to achieve optimal production of biogas for any anaerobic digestion bioreactor. Some example parameters are temperature, pH, loading rate, total solids, and carbon/nitrogen ratio. This paper walks through each of the typical parameters and assesses the impact of the changes one could make and methods for choosing an optimal setting.

Relevance: The ADM1 Model utilized in our analysis is an effort to reduce the number of parameters that have to be optimized in order to allow the user to achieve realistic estimations of biogas production quickly. However, it is important to understand what each parameter contributes to the production rate and how changes will impact said rate. Establishing a baseline rate and then adjusting parameters to understand the limit is a fundamental aspect of reactor design that, time willing, will be implemented in future analysis.

Further References:

Wang, X.; Yang, G.; Feng, Y.; Ren, G.; Han, X. Optimizing feeding composition and carbon–nitrogen ratios for improved methane yield during anaerobic co-digestion of dairy, chicken manure and wheat straw. Bioresour. Technol. 2012, 120, 78–83

Ferguson, R.M.W.; Coulon, F.; Villa, R. Organic loading rate: A promising microbial management tool in anaerobic digestion. Water Res. 2016, 100, 348–356.

Citation:

Menzel, T., Neubauer, P., & Junne, S. (2020). Role of microbial hydrolysis in anaerobic digestion. *Energies*, *13*(21), 5555. **Summary:** Hydrolysis is the first rate-limiting step for any anaerobic digestion reactor as it breaks down the principal nutrients into the more degradable products. The microbes responsible for hydrolysis will therefore determine the rate of downstream processes depending on their ability to interact with the substrate. This paper investigates improving microbial hydrolysis efficiency through parameter and process manipulation, staging, substrate pretreatment, and reactor design adjustments. **Relevance:** For complex substrates like food waste the process of microbial hydrolysis will be variable corresponding to the diversity of foods in the waste stream. Understanding mechanisms to improve overall hydrolysis will be important in understanding rate-limiting parameters to achieving higher biogas production **Further References:** Shrestha, S.; Fonoll, X.; Khanal, S.K.; Raskin, L. Biological Strategies for Enhanced Hydrolysis of Lignocellulosic Biomass during Anaerobic Digestion: Current Status and Future Perspectives. Bioresour. Technol. 2017, 245, 1245–1257 Gottardo, M.; Micolucci, F.; Bolzonella, D.; Uellendahl, H.; Pavan, P. Pilot Scale Fermentation Coupled with Anaerobic Digestion of Food Waste—Effect of Dynamic Digestate Recirculation. Renew. Energy 2017, 114, 455–463.
Citation: Donoso-Bravo, A., Mailier, J., Ruiz-Filippi, G. et al. (2013). Identification in an anaerobic batch system: global sensitivity analysis, multi-start strategy and optimization criterion selection. *Bioprocess Biosyst Eng* 36, 35–43 **Summary:** This paper investigates questions of structural and parametric identifiability for mathematical models of anaerobic digestion. It presents a procedure for estimating parameters in a simplified AD model, including an application of global sensitivity analysis. **Relevance:** Sensitivity analysis is of particular interest to our investigation of ADM1 because of the complexity of the model and our data. Understanding how sensitivity analysis has been conducted in related areas will help our understanding of our results. **Further References:** Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C., & Gujer, W. (2011). Global sensitivity analysis in wastewater treatment plant model applications: prioritizing sources of uncertainty. *Water research*, 45(2), 639-651. Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3), 271-280.
Citation: Silva, M.I., De Bortoli, A.L. (2020). Sensitivity Analysis for Verification of an Anaerobic Digestion Model. *Int. J. Appl. Comput. Math* 6, 38 **Summary:** This article demonstrates the calculation of first-order sensitivity coefficients for a system of nonlinear equations from an anaerobic digestion model.

Their results show the importance of each chemical reaction for each species involved in this model of anaerobic digestion.
Relevance: As stated previously, understanding how other researchers in the field have conducted sensitivity analysis is instrumental in developing an understanding of how sensitivity analysis could be conducted in our own use of such complex models.
Further References:
Valko, P., & Vajda, S. (1984). An extended ODE solver for sensitivity calculations. *Computers & Chemistry*, 8(4), 255-271.
Shen, J. (1999). A direct method of calculating sensitivity coefficients of chemical kinetics. *The Journal of chemical physics*, 111(16), 7209-7214.

Citation:
Jeong, HS., Suh, CW., Lim, JL. et al. (2005). Analysis and application of ADM1 for anaerobic methane production. *Bioprocess Biosyst Eng* 27, 81–89
Summary: An anaerobic model based on ADM1 was developed for the sealed reactor or "serum bottle" test and used to perform sensitivity analysis on the yield parameters. They determined that the most sensitive component was the concentration of methane gas, and were able to verify many of their simulated findings with experimental results.
Relevance: Understanding the sensitivity of ADM1 and related models to its parameters is important for validating the results of our own simulated experiments in anaerobic digestion.
Further References:
Wang, Q. J. (1997). Using genetic algorithms to optimise model parameters. *Environmental Modelling & Software*, 12(1), 27-34.
Mussati, M., Gernaey, K., Gani, R., & Jørgensen, S. (2002). Computer aided model analysis and dynamic simulation of a wastewater treatment plant. *Clean technologies and environmental policy*, 4, 100-114.

Citation:
J.B. Holm-Nielsen, T. Al Seadi, P. Oleskowicz-Popiel. (2009). The future of anaerobic digestion and biogas utilization. *Bioresource Technology* 100(22), 5478-5484
Summary: The EU has set renewable energy production goals that could be met, in part, by biogas production from anaerobic digestion of animal manure. This paper takes a high level look at the value and impact of anaerobic digestion in the efficient disposal and recycling of biological waste. The reclamation of methane from anaerobic digestion for use as fuel serves the dual purpose of reducing methane's release into the atmosphere where it acts as a greenhouse gas, while simultaneously providing a new source of energy in the form of biogas.
Relevance: This paper articulates the motivations for conducting our own research. Though the cited paper looks primarily at animal manure and organic wastes, the motivating principles spelled out in the paper apply equally to our application with food waste streams.
Further References:
Hjorth, M., Nielsen, A. M., Nyord, T., Hansen, M. N., Nissen, P., & Sommer, S. G. (2009). Nutrient value, odour emission and energy production of manure as

influenced by anaerobic digestion and separation. *Agronomy for sustainable development*, 29, 329-338. Braun, R., & Wellinger, A. (2003). Potential of co-digestion. In *IEA Bioenergy, task* (Vol. 37).
Citation: Caldeira, C., De Laurentiis, V., Ghose, A., Corrado, S., Sala, S. (2021). Grown and thrown: Exploring approaches to estimate food waste in EU countries. *Resources, Conservation and Recycling* 168, 105426 **Summary:** National studies on food waste produce varied results due to differences in approach. This paper aims to develop a harmonized system for modelling food waste generation in order to assess amounts reported by EU member countries. Three EU countries are used to illustrate the approaches and compare results. **Relevance:** This paper contains systematically determined estimates for food waste generated by Denmark, Germany, and Italy, broken down by each food group. This data could potentially be integrated into our study to estimate how differences in the type of food waste produced by different national diets & supply chains affect biogas productive capacity of various food waste streams. **Further References:** Caldeira, C., De Laurentiis, V., Corrado, S., van Holsteijn, F., & Sala, S. (2019). Quantification of food waste per product group along the food supply chain in the European Union: a mass flow analysis. Resources, Conservation and Recycling, 149, 479-488. Edjabou, M. E., Jensen, M. B., Götze, R., Pivnenko, K., Petersen, C., Scheutz, C., & Astrup, T. F. (2015). Municipal solid waste composition: Sampling methodology, statistical analyses, and case study evaluation. Waste Management, 36, 12-23.