Journal of Information Systems Applied Research

Volume 13, Issue 3
November 2020
ISSN: 1946-1836

In this issue:

iscap
Information Systems & Computing Academic Professionals

The **Journal of Information Systems Applied Research** (JISAR) is a double-blind peer reviewed academic journal published by ISCAP, Information Systems and Computing Academic Professionals. Publishing frequency is three issues a year. The first date of publication was December 1, 2008.

JISAR is published online (http://jisar.org) in connection with CONISAR, the Conference on Information Systems Applied Research, which is also double-blind peer reviewed. Our sister publication, the Proceedings of CONISAR, features all papers, panels, workshops, and presentations from the conference. (http://conisar.org)

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the JISAR journal. Currently the target acceptance rate for the journal is about 40%.

Questions should be addressed to the editor at editor@jisar.org or the publisher at publisher@jisar.org. Special thanks to members of EDSIG who perform the editorial and review processes for JISAR.

Journal of Information Systems Applied Research

Trash to Treasure: Predicting Landfill Gas Flow to Optimize Electricity Generation

Edgar Hassler
hassleree@appstate.edu

Dan Emery
emeryde@appstate.edu

Jason Hoyle
hoylejw@appstate.edu

Joseph Cazier
cazierja@appstate.edu

Center for Analytics Research and Education
Appalachian State University
Boone, NC 28608-2049, U.S.A

Abstract

Data analytics and machine learning have the potential to modify and improve many old school businesses. Among the oldest businesses for the human race is managing the waste we generate. In this paper we show how data science can be applied to help derive increased value from a byproduct of that waste, landfill gas. Gas produced from the decomposition of waste in landfills can be captured and transformed into a resource that benefits the local community, environment, and economy. We use analytics to better understand how weather conditions impact the methane content of landfill gas in ways significant enough to interfere with its use as a source of energy. We model methane concentrations in landfill gas and use machine learning techniques to predict future changes in methane concentration using a database of weather, water composition, and landfill gas collection performance metrics. A multilayer predictive model of methane concentration is developed that will aid in the transformation of day-to-day operations of landfill gas collection to maximize the utilization of gas extracted from the landfill, while minimizing the cost of pollution mitigation. This can help transform the industry while mitigating some environmental concerns.

Keywords: Sustainability, Landfill Gas, Analytics, Machine Learning, Methane Prediction

1. INTRODUCTION

Current global municipal solid waste generation levels are approximately 1.3 billion tonnes per year and are expected to increase to approximately 2.2 billion tonnes per year by 2025 (Hoornweg & Bhada-Tata, 2012). Waste that is not recycled or incinerated is placed into landfills where its decomposition releases methane gas and carbon dioxide into the atmosphere accounting for an estimated 17.6% of human-related methane gas emissions in the United States (EPA, 2017).

In addition to climate change impacts, methane and other gases created in a landfill pose risks to air quality as well as human health and safety. To limit these impacts, emissions from landfills are controlled by installing gas collection systems and either burning the landfill gas (LFG) in a flare or utilizing it as an energy source to produce electricity, steam, or even vehicle fuel. Using LFG as an energy source transforms post-closure management of landfills from a cost center into a profit center and has the potential to transform foundational thinking in the waste management industry from viewing trash as waste into viewing trash as a resource—after all, it's only waste if it's not used.

In order to make the gaseous byproducts of our waste more useful, there are several analytics techniques that have the potential to help. These include

- **Modeling –** In order to apply analytic techniques to help convert this waste to useful energy, we must first model and understand the drivers of gaseous levels and concentration. This requires discovery and weighting of key drivers.

- **Prediction –** Next, as the relationships are better understood though discovery, we can apply that understanding by building predictive models based on key variables to forecast what is likely to happen next. This is critical to building a useful solution.

- **Automation –** Finally, as the relationships are understood and can be reliably predicted, these can be incorporated into an automatic solution using information and robotics technologies.

These stages of data analytics to model a current relationship, predict a future one and automate a solution are critical to the use of data science for building usable solutions with real impact.

For example, regulations requiring the monitor and control of LFG emissions specify a minimum frequency of monthly visits (Monitoring of Operations, 2000), and these requirements mean a technician must physically visit each of hundreds of gas wells to take measurements using handheld instruments and make manual adjustments for decades after the landfill closes. This infrequent, labor-intensive, and time-consuming manual process is not only an imprecise, potentially arbitrary, approach to managing LFG collection for the control of emissions, but it also produces an LFG supply that is highly variable in both gas flow and methane concentration that limits the value of using this energy-rich gas as fuel. **Automation** can add real value here, but first it must be modeled and predicted reliably.

In terms of **modeling** there is a considerable body of research on the production and diffusion of methane gas in a landfill (Rachor et al. 2013, Alexander 1971, Spokas et al. 2006, Farquhar & Rovers 1973, Bade Shrestha & Narayanan 2008, Toerien & Hattingh 1969, among others); an expanding literature on approaches and strategies to maximize LFG generation, extraction, and use as fuel (Reinhart 2002, Reinhart 1996, Townsend 2018, Ozkaya 2007, Warith 2003, Buivid 1981, Kinman 1987); and a long line of research into the internal dynamics of landfills that influence methane generation emissions, migration, and extraction such as pressure, cover permeability, depth, moisture content, waste composition, temperature, etc. (Arigala et al. 1995, Chen et al. 2003, Xi & Xiong 2013, Hashemi et al. 2002, Sanchez et al. 2006).
On example of a useful model using the fuzzy logic algorithm was derived by Garg et al. (2006) where they built a useful model based on average climate and waste site characteristics.

While these studies and others of a related nature contribute to increased knowledge of methane-related processes inside a landfill, the few that even address time do so on scales of months, years, or decades.

We need **prediction** however, when using LFG as fuel, particularly in a continuous process such as generation of electricity or steam, where changes in the performance of the collection system over periods of months or years are significantly less important than changes in collection system performance over *time intervals of minutes or hours*. Significant changes in methane concentration and, to a lesser extent, changes in gas flow rate will adversely impact the operation, efficiency, and output of LFG-fueled energy conversion devices (e.g. boilers or electricity generators) and at times even cause shutdown events. There exists a limited body of research examining the causes and nature of variability in LFG emissions or gas collection system performance over short time intervals (e.g. hourly or sub-hourly) and few, if any, efforts that attempt to forecast the near-term performance of a landfill gas collection system on an hourly or sub-hourly basis as we have done here.

Absent external influences, landfill gas pressure will stabilize at a landfill-specific level, and once steady-state conditions are attained then gas emission or migration out of the landfill occurs at the rate of gas generation. Lu and Kunz (1981) demonstrated that extracting LFG lowers the gas pressure during the time the vacuum is applied, but when the vacuum is removed the internal gas pressure quickly returns to its steady-state level, implying that gas extraction functions as a substitute for gas emissions and that factors impacting gas emissions will also impact gas extraction.

Young (1990) described the inverse relationship between changes atmospheric pressure and the volume of LFG emissions concluding that the amount by which emissions change over a sub-hourly time interval is proportional to rate of change in atmospheric pressure rather than the actual value of atmospheric pressure. In a further study, Young (1992) explained several factors that cause carbon dioxide and methane – the two largest constituents of LFG – to respond differently to changes in atmospheric pressure. In particular, he demonstrates, in stepped time intervals of less than one-half hour, methane concentrations rising from 50% to 55% over a period of hours and then rapidly and consistently falling to about 40% over the next day or two in response to changing atmospheric pressure.

Aghdam, Scheutz, and Kjeldsen (2019) summarize the conflicting state of conclusions reached in some of the limited number of empirical studies examining the influence of meteorological factors on LFG emissions or collection system performance. Among those factors, absolute level of barometric pressure, rate of change in barometric pressure, solar radiation, wind speed, soil moisture, air temperature, and soil temperature have been found to both have a significant influence and also not to have a significant influence on gas volume and composition.

The inconsistency of results among empirical studies coupled with the complexity resulting from the multitude of interrelated factors that may influence LFG flow and composition over short time intervals provides an opportunity to apply modern data science techniques to improve the performance of LFG energy systems. By carefully analyzing the application and data characteristics, researchers and practitioners can then adopt or develop the appropriate analytical techniques to derive the intended impact (Chen et al. 2012).

Machine learning and applied analytics have the potential to transform the way in which this one-time pollutant is transformed into both a valuable energy source benefiting local economies and a new source of value for the waste management industry, while also mitigating its environmental impact. Analytics also has the potential to automate the current costly process of collection system tuning, identify equipment failures and needed repairs much more rapidly, and greatly improve safety at and around the landfill site by remotely identifying the presence of hazardous gases such as methane and hydrogen sulfide at the landfill surface.

As the economic and environmental benefits of LFG use are increasingly recognized, data analytics and machine learning can aid in the optimization of LFG collection and spur a transformation in the waste management industry. By maximizing the amount of LFG collected and selling the recaptured gas, energy, or carbon credit, a landfill can literally turn trash into treasure.

Research Objective

As gas is extracted from the landfill, methane concentration of the gas fluctuates drastically. If methane concentration drops too low, the gas becomes unusable as an energy source, induces stress on the generating engine and can lead to damage or failure of the engine.

By using data analysis techniques, this project specifically aims to model methane concentration in near-real-time and thus enable landfill operators to predict future changes in methane concentration and gas flow to preemptively respond to methane concentration changes by taking the necessary actions to protect the engine, keep LFG composition optimal, and maximize profit while minimizing the negative effects on the environment.

The remainder of this paper is organized as follows. Section 2 provides background on the business case, environmental impacts and foundational knowledge of LFG production. Section 3 describes the methodology of the study including the dataset development and model construction. In Section 4, the results of the modeling process are presented followed by discussion in Section 5. Section 6 concludes the paper and notes possible future directions for research.

2. BACKGROUND

Business Case
The high concentration of methane gas in LFG incentivizes landfill operators to capture the methane gas diffusing out of the landfill for reuse or sale. A study based on 8,196,000 tonnes municipal solid waste generated in Peninsular Malaysia in 2010 estimated that the collection and sale of landfill gas generated 1.9 billion kWh of electricity per year, worth over US$190 million (Johari et al., 2012). Another study found that capturing the methane emissions in Oman from 2016–2030 will attract a total revenue of approximately US$333 million from carbon reduction, and approximately US$291 million from electricity generation (Abushammala et al. 2016). With such money at stake, the importance of understanding more about LFG production and composition in a landfill is heightened.

Currently, the Environmental Protection Agency's Landfill Methane Outreach Program has identified 634 operational landfill gas collection projects in the United States. Combined, these projects generate 291 mmscfd (million standard cubic feet per day)—enough energy to supply the state of Maine's annual natural gas consumption. If sold at $3.50/1000 scf, it would generate US$1,018,500 in revenue every day solely from the sale of the collected gas.

Landfill gas is also a cheap source of energy. Local businesses realize cost savings associated with using LFG as a replacement for more expensive fossil fuels. Some companies could save millions of dollars over the life of their LFG energy projects (LMOP 2017). General Motors has leveraged this cheap energy source into massive savings for the company. A statement from General Motors Green on 13 March 2016 reads, "A decades-long approach to sourcing renewable energy has produced lessons learned that helped GM further reduce its environmental footprint and save $80 million along the way." This testimonial serves as a pilot-study of sorts as to how much the utilization of landfill gas as an energy source can benefit a company.

While the energy savings and environmental benefits of using LFG are potentially substantial, maximizing the value of these benefits requires effective management of the LFG collection system as even short-duration system shutdowns can result in significant financial losses. For example, a landfill selling 2 mmscfd at a price of $3.50/1000 scf loses $7,000 in revenue for every day the system is not operating.

Environmental Impacts
As stated previously, landfill gas is a byproduct of society. It is important to note this paper is not suggesting that collecting and combusting LFG eliminates the need to decrease waste production. It is simply pointing out that capturing and combusting LFG will always be more environmentally friendly than allowing the gas to escape into the atmosphere.

Methane gas (CH_4), has a molar mass of 16.04 g/mol. Carbon Dioxide (CO_2), has a molar mass of 44.01 g/mol. The combustion of methane gas converts 1 mole of methane gas to 1 mole of carbon dioxide gas. Thus, combusting 1 ton of methane gas leads to the release of 2.74 tonnes of carbon dioxide gas.

Previous studies have shown that methane gas has a 20-year global warming potential of 86 when considering carbon-climate feedback (Myhre et al. 2013). Carbon dioxide is used as a baseline for global warming potential values, giving it a value of 1.

Using these global warming potentials, we assert that each ton of methane gas combusted reduces the 20-year global warming potential of the gas by 83.26. Essentially, 1 ton of methane gas released to the atmosphere carries a global warming potential of 86. By converting 1 ton of methane gas to 2.74 tonnes of carbon dioxide, the resultant gas carries a global warming potential of 2.74, a ~97% reduction in global warming potential.

Landfill Gas Collection Systems
Landfill gas collection systems are remarkably complex systems, featuring anywhere from dozens to hundreds of interconnected gas wells located throughout the landfill, all applying a vacuum to extract the gas produced by myriad chemical and biological processes underway within the landfill.

These components all interact with external factors, such as weather, and the resulting gas extracted at a given moment varies significantly in volume and composition. Yet, despite this complexity, LFG collection systems are typically managed by a technician who takes measurements using handheld instruments and adjusts the collection system by turning valves on a largely ad hoc basis.

Using applied analytics techniques to understand and predict the underlying biochemical processes and physical factors that impact LFG quality will enable a far more efficient, cost-effective, and profitable use of LFG as an energy source, while simultaneously increasing environmental and operational benefits.

Processes Within the Landfill

When municipal solid waste is first deposited in a landfill, it undergoes an aerobic (with oxygen) decomposition stage where little methane is generated. Then, typically within less than one year, anaerobic (without oxygen) conditions are established and methane-producing bacteria begin to decompose the waste and generate methane. Thus, landfill gas production can be broken down into two major stages, the non-methanogenic stage and the methanogenic stage (See Figure 1).

The optimum temperature range for aerobic decomposition is 54 to 71ºC, while the optimum temperature range for anaerobic bacteria is 30 to 41ºC. A dramatic drop in activity of anaerobic bacteria has been noted at temperatures below 10ºC (USACE 2013).

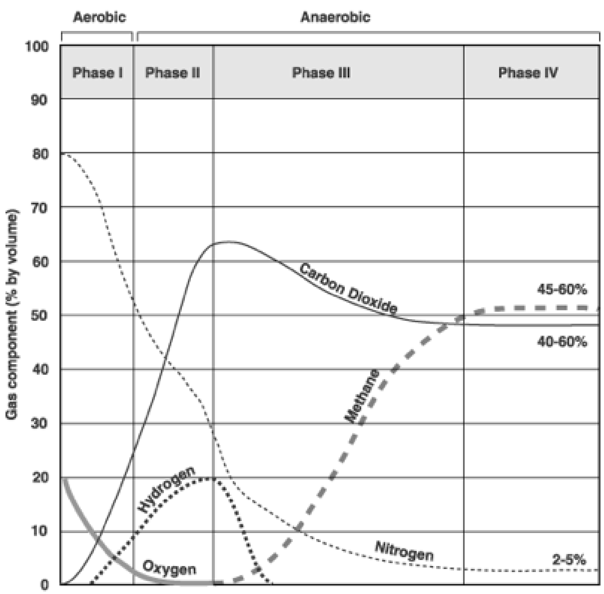

Figure 1. Landfill gas composition over time (EPA, 1997). Phase duration time varies with landfill conditions.

Precipitation dramatically affects the LFG generation process by supplying water to the process and by carrying dissolved oxygen into the waste with the water. High rates of precipitation may also flood sections of the landfill, which will obstruct LFG flow. The amount of precipitation that reaches the waste is highly dependent on the type of landfill cover system (USACE 2013).

Molecular diffusion occurs in a system when a concentration difference exists between two different locations. The concentration of a volatile constituent in the LFG will almost always be higher than that of the surrounding atmosphere, so the constituent will tend to migrate to the atmosphere. Wind often serves to keep the surface concentration at or near zero, which renews the concentration gradient between the surface and the interior of the landfill, thus promoting the migration of vapors to the surface. Geomembranes in landfill covers significantly reduce diffusion as the geomembrane prevents LFG from diffusing to the atmosphere (USACE 2013).

Advective flow occurs where a pressure gradient exists. The rate of LFG movement is generally orders of magnitude faster for advection than for diffusion. LFG will flow from higher pressure to lower pressure regions. In a landfill, advective forces result from the production of vapors from biodegradation processes, chemical reactions, compaction, or an active LFG extraction system. Variations in water table elevations can create small pressure gradients that either push gases out (rising tide) or draw gases in (falling tide). Changes in barometric pressure at the surface can also have an impact on the advective flow of LFG (USACE 2013).

3. METHODS

Data for this project was collected from a midsize rural landfill in the Southeastern United States.

Landfill Monitoring System

Landfill gas flow and methane content are continuously recorded by an Eurotherm 6180A Paperless Chart Recorder. A Thermal Instrument Model 9500 flow meter was used to correct for temperature and pressure at 60 degrees Fahrenheit at 1 atmosphere of pressure. Gas was analyzed using the Landtec FAU-TDL gas analyzer.

Field check calibration for all equipment was performed. The project was verified to the Climate Action Reserve's Landfill Project Protocol Version 3.0. All data was collected prior to the beginning of this research.

Having these systems in place should allow for an automated system to be implemented with little changes to the current setup, which should ease the deployment of the model.

Dataset Description

Observations were recorded every 20 minutes from midnight 14 June 2012 to 5:00 pm 31 August 2013. The initial dataset contained 27,962 observations and 38 variables. Of the 38 variables, 12 describe weather conditions. The weather data was collected from a nearby rural airport weather station.

The 12 weather variables measure six different weather metrics. Each observation recorded the six metrics at the local airport and provided a regional analysis. The dataset also included 25 landfill gas collection system performance metrics. A timestamp of each observation was also included.

There are periods of missing data, half of which occur within a 30-day period in 2012 and during a couple weeks in Feb 2013. The missing time periods for the landfill gas collection system performance metrics are as follows:

- 7/10/2012 - 8/18/2012
- 10/09/2012: 6 hours
- 10/10/2012: 6 hours
- 10/22/2012: 7 hours
- 02/13/2013 - 02/19/2013
- 07/09/2013 - 07/17/2013

There were periods of missing values for the weather data as well. Exploratory analysis of these missing values showed that if one variable was missing, then all values were missing. The time periods never lasted longer than two hours. The short timeframe of missing values for the weather data allowed for the values to be imputed using the following equation:

$$y_i = \frac{(y_{i-1}^* + y_{i+1}^*)}{2} \qquad (2)$$

where y_i is the missing value, y_{i-1}^* is the previous recorded instance of that variable, and y_{i+1}^* is the next recorded instance of that variable.

Feature Construction

Precipitation was not included in the initial weather dataset. Knowing that moisture plays a large role in landfill gas production, precipitation information was acquired through Climate Data Online. Daily precipitation values were joined to the existing dataset by the date of the observation.

Water quality information was obtained from a dataset made available by the United States Geological Survey for a river approximately 50 miles southeast of the county landfill. While not an optimal indicator of conditions on the site, it is included to provide, at a minimum, a regional indication of water quality due to lack of instrumented reading at the facility. Observations were recorded every 15 minutes. The primary variables of interest in this dataset were water elevation, dissolved oxygen in the water, and pH of the water. Prior to joining the datasets, the water quality data was grouped by date and hour of day. Following this step, average elevation, dissolved oxygen, and pH was calculated. The two datasets were then joined by date and hour of the day.

One common thread between all landfill gas models is the importance of landfill moisture. Access to weather data will allow for more information on moisture. The three and five-day cumulative precipitation values were calculated from the daily precipitation values.

$$3\ day\ precipitation\ for\ row\ i\ = \sum_{j=1}^{3} P_{i-j} \qquad (3)$$

$$5\ day\ precipitation\ for\ row\ i\ = \sum_{j=1}^{4} P_{i-j} \qquad (4)$$

Oxygen content in the landfill is yet another obvious driver of methane gas production. As mentioned previously, the primary methane-producing reaction in a landfill is anaerobic. The dissolved oxygen content of a nearby river serves as a proxy for dissolved oxygen in the water present in the landfill. The problem with this variable is that dissolved oxygen can be high, but if no precipitation occurs while it is high, no new oxygen is introduced to the system. To account for this scenario, two variables were created.

First, dissolved oxygen of the river water was multiplied by daily precipitation. The justification of this variable is that it provides a means of measuring total oxygen introduced into the system. Next, a three-day cumulative value of this variable was created using the following equation:

$$\boldsymbol{3\ day\ summed\ feature\ for\ row\ i} = \sum_{j=1}^{3} \boldsymbol{F_{i-j}} \qquad (5)$$

Preliminary modeling shows that wind direction is relatively important, but that wind speed may not be significant. Intuitively, we know that these two variables are related and that the combination of the two variables gives us more information than each variable separately, e.g. a 10 mph wind from the west may affect methane gas concentration more than a 1 mph wind from the west. These variables were combined using the following equation:

$$\begin{aligned}&\boldsymbol{Wind\ feature}\\&= \boldsymbol{Wind\ Speed}\\&* \mathbf{sin}\left(\frac{\mathbf{2}\boldsymbol{\pi}[\boldsymbol{Wind\ Direction} - \mathbf{55}]}{\mathbf{360}}\right)\end{aligned} \tag{6}$$

Equation 1 shows that the changes in gaseous concentration over time depends on gas velocity, gas dispersivity, and the molecular gas diffusion coefficient in the soil gas phase. Solving this equation does not fit within the scope of a data-driven approach to methane gas modeling; however, it does provide insights into useful variables. The molecular gas diffusion coefficient in the soil gas phase is proportional to T^1.5/p, where T represents temperature of the system and p represents pressure of the gaseous components of the system. This value was computed in the dataset by dividing recorded temperature, raised to the 3/2 by barometric pressure.

The final variables added to the dataset are lagged performance variables. Preliminary modeling attempts showed that it is not sufficient to simply use weather and water composition variables to predict methane concentration. By adding in lagged variables, we gain more insight into the activite components of the system. For this study, the closest lagged variables were four hours.

Initial Analysis
This project has two major components:

- Develop a regression model that accurately predicts the methane concentration of the landfill gas.
- Develop a classification model to predict if methane concentration will fall below 40% (1 = methane concentration <= 40%, 0 = methane concentration > 40%)

To develop a regression model for methane concentration, we must account for the drastic impact that extracting gas from the landfill has on methane concentration. Figure 2 illustrates this relationship. When the header valve is open, meaning that LFG is being extracted, we observe methane concentration of the LFG fluctuates drastically over the course of a day. When the header valve is closed, we observe methane concentration generally increases with time with much smaller fluctuations in the general path. Although this presents some challenges with the regression model, it opens another path of study.

By creating two regression models, one for each header valve scenario, we can develop a further understanding of the system. The regression model focused on the case when the header valve is closed has the potential to establish a baseline methane concentration growth rate. This baseline prediction may also have relevance in the regression model, focusing on the scenario in which the header valve is open and gas is being extracted.

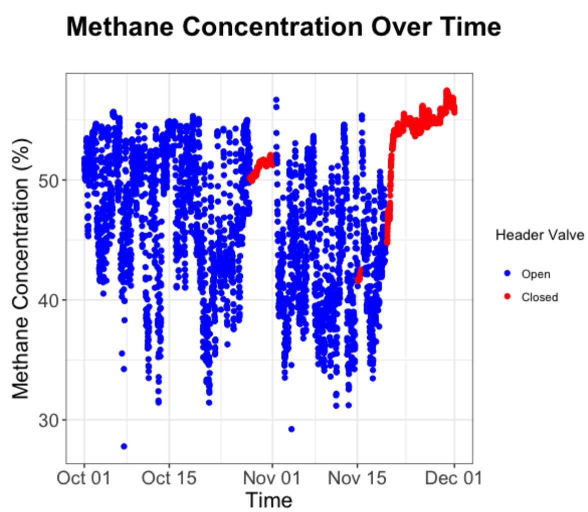

Figure 2. Displays LFG methane concentration at Rockingham County Landfill from 1 Oct 2012 to 31 Nov 2012.

According to the design specifications of the collection system engine used by the landfill, the engine is designed to utilize gas composed of 50% methane by volume. As concentration decreases, more stress is placed on the engine. Once methane concentration falls below 40%, the engine possesses a high risk of failure. Although the engine needs 40% methane concentration, the flare can likely operate with lower concentrations. Figure 3 shows the counts of each scenario for the binary variable created for the classification model.

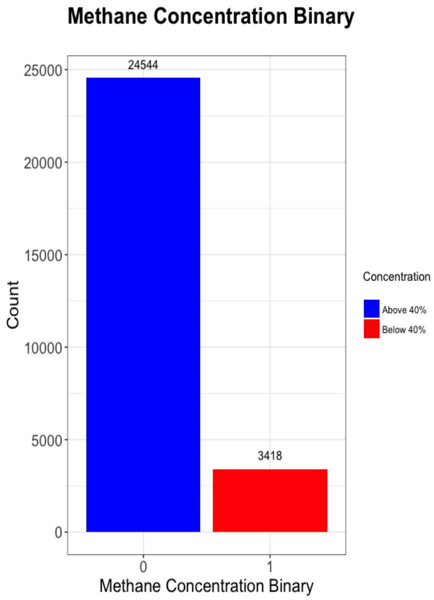

Figure 3. Displays the count of each scenario regarding LFG methane concentration.

We were unable to find any statistically significant relationship between the weather variables and methane concentration.

Model Construction

Prior to using any machine learning algorithm to develop a model, a baseline prediction was made. For regression models, this involved simply assuming the average methane concentration for every point in time. For classification, this involved assuming the methane concentration was always greater than or equal to 40%. This assumption was made due to the imbalanced structure of the dataset. This analysis was done in order to assess the effectiveness of each machine learning algorithm.

Multilinear regression and logistic regression were then used to provide a first attempt at the model. Both models utilize the same underlying math; however, logistic regression is used to predict the probability of an event occurring (methane concentration falling below 40%) while multilinear regression predicts a value (methane concentration). This is represented by Equation 7.

$$\boldsymbol{y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \; for \; i = 1, 2, \ldots n} \tag{7}$$

The final model utilized in the study is known as a multilayer machine learning algorithm. This technique utilizes a variety of algorithms and passes predictions from those algorithms to one or multiple higher-level algorithms. If multiple higher-level algorithms are used, the predictions from each algorithm are typically blended together to create an ensemble.

All models were developed using Python 3.6. On each model, a 4-fold cross validation was used.

4. RESULTS

Model Overview

To assess each classification model, sensitivity was used as the evaluation metric. Sensitivity is a measure of the model's ability to predict the positive case. This metric was selected because we want the model to detect an event which does not happen frequently. The regression model was assessed using mean absolute error (Equation 8). Mean absolute error was chosen because it shows, on average, how far off the estimate is from the true value. Table 1 shows the mean absolute error and sensitivity for each model.

$$Mean\ Absolute\ Error = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{8}$$

	Model 0 Simple Average	Model 1 Multilinear/ Logistic Regression	Model 2 Multilayer Machine Learning
Mean Absolute Error	5.6%	3.1%	1.3%
Sensitivity	0%	0%	81.3%

Table 1. Mean absolute error and classification sensitivity for each of the three models used in this study.

Model 2 - Multilayered Machine Learning

Model 2, which was a multilayered machine learning model, clearly outperformed all other models used in this study. The basic structure of this model was as follows:

- The base level contained a combination of Multilinear/Logistic regression, Random Forests, Extremely Randomized Trees, Gradient Boosted Trees, Extreme Gradient Boosted Trees, K-Nearest Neighbors, and Neural Networks
- The intermediate level utilized two Gradient Boosted Tree structures from the LightGBM and XGBoost libraries.
- The final level was a weighted average between the two intermediate levels.

The mean absolute error of the predicted methane concentration and the actual methane concentration was 1.30% with a sensitivity of 81.3% (see Table 1).

The model had a classification accuracy of 95.0% (see Table 2). R^2=0.9032 when comparing the predicted methane concentration to actual methane concentration (see Figure 4).

	Predict Above 40%	Predict Below 40%
Actual Above 40%	23,935	808
Actual Below 40%	601	2,606

Table 2. Confusion matrix for classification model 2.

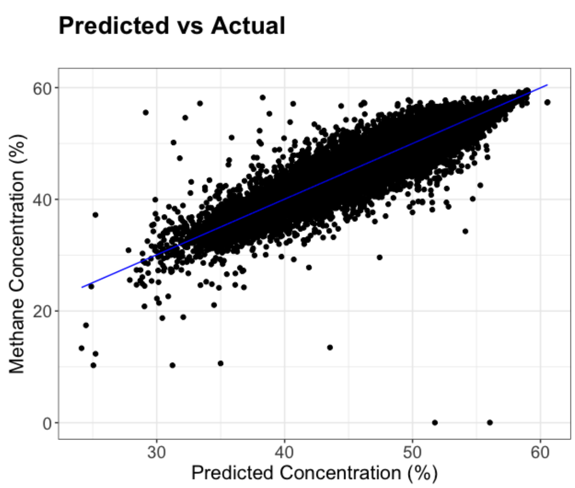

Figure 4. Predicted methane concentration of the LFG vs the actual methane concentration. R^2=0.9032.

5. DISCUSSION

The purpose of this project is to serve as an aid in the decision-making process for landfill operations. Our results show that through adequate data collection, variable manipulation, and proper predictive modeling techniques, LFG methane concentration can be predicted within a time horizon useful for operations.

By incorporating data science techniques into the day-to-day operations of a landfill gas collection system, the value and quantity of gas extracted from the landfill for generation purposes can be maximized. This predictive capability offers the potential for development of strategies guiding intervention and alteration of collection system operations prior to methane levels dropping below minimal useful thresholds, thus ensuring higher reliability and quality from landfill gas systems.

The classification model laid out in this paper has a sensitivity of 81.3%. If properly implemented, this model could prevent four out of five stress events from occurring. This result is quite impactful, as preventing engine stress and, in turn, preventing engine damage, can improve revenue and decrease the environmental impact of methane gas released by the landfill. This model also provides use to a fully automated system, as it can feed inputs into the system, changing its course of action.

With the ability to accurately predict engine stress events four hours in advance of the event occurring, landfill operators can alter their actions to alleviate the stress event. With the collection system database growing every day, we can improve the decision-making process, and modeling of events, by fully utilizing this data.

6. CONCLUSIONS

Landfill gas is a significant contributor to climate change, but its capture and use both reduces greenhouse gas emissions and provides a low-cost source of local renewable energy that can stimulate local economies. With the potential to benefit the local community, environment, and economy, the decomposition of solid waste should be seen as a potential resource. In fact, it should only be considered waste if it is not used.

This research provides a foundation for a new data-driven LFG model. By expanding data traditionally recorded during operations with secondary data that can serve as an appropriate proxy, it is possible to construct sub-hourly models of LFG flows using analytics techniques.

Additionally, this paper shows clearly how data analytics, combined with domain knowledge, can model a current relationship, predict a future one and eventually automate the process to transform industry.

Future research should investigate additional modeling techniques. Regression techniques were suboptimal in this case; however the Multi-layered machine learning model results suggest that Random Forests, Extreme Randomization, Gradient Boosting, Clustering techniques, and Neural Networks may improve predictions.

By moving forward with this research, we can assist waste management organizations with automation and better utilize data to maximize their resources through the incorporation of predictive analytics, ultimately converting more trash to treasure, while reducing the environmental impact of methane escaping from landfills.

7. REFERENCES

Abushammala, M. F., Qazi, W. A., Azam, M.-H., Mehmood, U. A., Al-Mufragi, G. A., & Alrawahi, N.-A. (2016). Economic and environmental benefits of landfill gas utilisation in Oman. *Waste Management & Research*, *34*(8), 717–723. https://doi.org/10.1177/0734242X16628983

Aghdam, E. F., Scheutz, C., & Kjeldsen, P. (2019). Impact of meteorological parameters on extracted landfill gas composition and flow. Waste Management, 87, 905-914.

Alexander, M. (1971). Microbial ecology. Wiley, New York.

Arigala, S. G., Tsotsis, T. T., Webster, I. A., Yortsos, Y. C., & Kattapuram, J. J. (1995). Gas generation, transport, and extraction in landfills. Journal of environmental engineering, 121(1), 33-44.

Buivid, M. G., Wise, D. L., Blanchet, M. J., Remedios, E. C., Jenkins, B. M., Boyd, W. F., & Pacey, J. G. (1981). Fuel gas enhancement by controlled landfilling of municipal solid waste. Resources and Conservation, 6(1), 3-20.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly: Management Information Systems, 36(4), 1165–1188.

Chen, Y. C., Chen, K. S., & Wu, C. H. (2003). Numerical simulation of gas flow around a passive vent in a sanitary landfill. Journal of Hazardous Materials, 100(1-3), 39-52.

Farquhar, G. J., & Rovers, F. A. (1973). Gas production during refuse decomposition. Water, Air, and Soil Pollution, 2(4), 483–495. https://doi.org/10.1007/BF00585092

Garg, A., Achari, G., & Joshi, R. C. (2006). A model to estimate the methane generation rate constant in sanitary landfills using fuzzy synthetic evaluation. *Waste Management & Research*, *24*(4), 363–375. https://doi.org/10.1177/0734242X06065189

Harrill, D. J. (2014). Analysis of Changes in Landfill Gas OutPut and the Economic Potential for Development of a Landfill Gas Control Prototype (Master of Science, Appalachian State University). Retrieved from https://libres.uncg.edu/ir/asu/f/Harrill,%20David_2014_Thesis.pdf

Hashemi, M., Kavak, H. I., Tsotsis, T. T., & Sahimi, M. (2002). Computer simulation of gas generation and transport in landfills--I: quasi-steady-state condition. Chemical Engineering Science, 57(13), 2475-2501.

Hoornweg, D., & Bhada-Tata, P. (2012). What a Waste: A Global Review of Solid Waste Management (Review No. Knowledge papers no. 15). Retrieved from World Bank website: https://openknowledge.worldbank.org/handle/10986/17388

IPCC. (2017). IPCC — Intergovernmental Panel on Climate Change. Retrieved August 19, 2017, from IPCC — Intergovernmental Panel on Climate Change website: https://www.ipcc.ch/

Johari, A., Ahmed, S. I., Hashim, H., Alkali, H., & Ramli, M. (2012). Economic and environmental benefits of landfill gas from municipal solid waste in Malaysia. Renewable and Sustainable Energy Reviews, 16(5), 2907–2912. https://doi.org/10.1016/j.rser.2012.02.005

Kinman, R. N., Nutini, D. L., Walsh, J. J., Vogt, W. G., Stamm, J., & Rickabaugh, J. (1987). Gas enhancement techniques in landfill simulators. Waste management & research, 5(1), 13-25.

Lu, A. H., & Kunz, C. O. (1981). Gas-flow model to determine methane production at sanitary landfills. Environmental science & technology, 15(4), 436-440.

Monitoring of Operations (2000). 40 C.F.R. §60.756

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., … Zhang, H.

(2013). Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Retrieved from https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter08_FINAL.pdf

OECD. (2003). OECD Environmental Indicators: Development, Measurement, and Use [Reference]. Retrieved from Organization for Economic Cooperation Development website: http://www.oecd.org/environment/indicators-modelling-outlooks/24993546.pdf

Ozkaya, B., Demir, A., & Bilgili, M. S. (2007). Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors. Environmental Modelling & Software, 22(6), 815-822.

Rachor, I. M., Gebert, J., Gröngröft, A., & Pfeiffer, E.-M. (2013). Variability of methane emissions from an old landfill over different time-scales. European Journal of Soil Science, 64(1), 16–26. https://doi.org/10.1111/ejss.12004

Reinhart, D. R., & Basel Al-Yousfi, A. (1996). The impact of leachate recirculation on municipal solid waste landfill operating characteristics. Waste Management & Research, 14(4), 337-346.

Reinhart, D. R., McCreanor, P. T., & Townsend, T. (2002). The bioreactor landfill: Its status and future. Waste Management & Research, 20(2), 172-186.

Sanchez, R., Hashemi, M., Tsotsis, T. T., & Sahimi, M. (2006). Computer simulation of gas generation and transport in landfills II: Dynamic conditions. Chemical Engineering Science, 61(14), 4750-4761.

Shrestha, S. O. B., & Narayanan, G. (2008). Landfill gas with hydrogen addition – A fuel for SI engines. Fuel, 87(17), 3616–3626. https://doi.org/10.1016/j.fuel.2008.06.019

Spokas, K., Bogner, J., Chanton, J. P., Morcet, M., Aran, C., Graff, C., … Hebe, I. (2006). Methane mass balance at three landfill sites: What is the efficiency of capture by gas collection systems? Waste Management, 26(5), 516–525. https://doi.org/10.1016/j.wasman.2005.07.021

Toerien, D. F., & Hattingh, W. H. J. (1969). Anaerobic digestion I. The microbiology of anaerobic digestion. Water Research, 3(6), 385–416. https://doi.org/10.1016/0043-1354(69)90002-5

Townsend, T. G. (2018). Landfill bioreactor design & operation. Routledge.

US EPA. (2016). Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2015 (Reports and Assessments No. EPA 430-P-17-001; p. 633). Retrieved from U.S. Environmental Protection Agency website: https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2015

USACE. (2013). Landfill Gas Collection and Treatment Systems (Manual No. EM 200-1-22). Retrieved from US Army Corps of Engineers website: https://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_200-1-22.pdf

Warith, M. A. (2003). Solid waste management: new trends in landfill design. Emirates Journal for Engineering Research, 8(1), 61-70.

World Development Report 1992: Development and Environment, Andrew Steet. 1992. Oxford University Press, New York, NY. 308 pages. ISBN: 0-19-520876-5. (1994). Bulletin of Science, Technology & Society, 14(2), 117–118. https://doi.org/10.1177/027046769401400236

Xi, Y., & Xiong, H. (2013). Numerical simulation of landfill gas pressure distribution in landfills. Waste Management & Research, 31(11), 1140-1147.

Young, A. (1990). Volumetric changes in landfill gas flux in response to variations in atmospheric pressure. Waste Management & Research, 8(5), 379-385.

Young, A. (1992). The effects of fluctuations in atmospheric pressure on landfill gas migration and composition. Water, Air, and Soil Pollution, 64(3-4), 601-616