Preserve: Sensor based Low-cost Emissions Management for Small and Medium Farms

Sachin Mahesh
Castleknock College
Dublin, Ireland
sachin.mahesh@2024.castleknockcollege.ie

Mike Hibbett
Irish Manufacturing Research
Dublin, Ireland
mike.hibbett@imr.ie

***Abstract*— Greenhouse gas (GHG) emissions from the agricultural sector have significant environmental impacts, including contributing to climate change and air pollution. Reducing emissions from the agricultural sector is a crucial step in addressing environmental and economic challenges. While several regulatory measures have been put in place, small and medium farms struggle to control their emissions due to lack of simple, affordable, and user-friendly emission management tools. Finding effective and scalable solutions that can be widely adopted by farmers and producers is a complex task that requires addressing a range of interconnected factors. In this paper, we present Preserve, a low-cost solution that relies on in-situ real-time measurements from farms and provides guidance through predictive models while allowing farmers to monitor via user-friendly cloud interface.**

Keywords*— *Agricultural emissions, GHG emissions, sensor network, predictive models, cloud interface

I. INTRODUCTION

The agriculture, forest, and other land use (AFOLU) sector plays a significant role in greenhouse gases (GHGs) emissions. The AFOLU sector contributes to just under a quarter of global anthropogenic GHG emissions (radiative forcing), producing circa 10-12 gigatons of $CO2$ (GtCO2eq) each year [1]. The sector's GHG emissions is of great concern due to its implications for climate change; several GHGs including Carbon Dioxide (CO2), Methane (CH4), and Nitrous Oxide (N2O), evident within the sector, trap heat in the Earth's atmosphere, increasing the global temperature and its associated environmental impacts. Methane and Nitrous Oxide have a global warming potential (GWP), in a hundred years of 28-36 and 265-298 GWP-100 respectively, significantly more potent than CO2.

Methane emissions in the agriculture sector primarily stem from enteric fermentation in ruminant animals which is produced as a by-product of microbial fermentation occurring in the specimen's rumen. Typical CH4 emissions range from 150.7g/animal/day from cattle to 13.7g/animal/day from goats respectively [2], contributing to the rising temperatures of 1.9*C to 3.9*C predicted by the year 2100. Although enteric fermentation plays a huge role in CH4 emissions in the agriculture sector, optimising the level of manure produced by each type of cattle can also influence emissions. Likewise, emissions from Nitrous Oxide, which primarily include agriculture and soil manure management, contribute to 364.4 million metric tons of Carbon Dioxide (MMTCO2eq) in the US in 2019 with agricultural soil management accounting for 94.56% of total N2O emissions during that year [3]. The total GHG emissions in 2019 in the agriculture sector account for 628.6 million metric tons of carbon equivalents (MMTCE) with CO2 only contributing 7.8 MMTCE to the total whilst having mass of approx. 6.36x the mass of N2O.

Taken together, these statistics outline a huge challenge of keeping the average global temperature rise under 2*C and ideally under 1.5*C as outlined by the Paris Agreement, it is increasingly clear that there is a need for a solution which can reduce the impact of enteric fermentation, soil management, and manure management, limiting the effects of both CH4 and N2O in contributing to GHG emissions.

The United Nations Framework Convention on Climate Change (UNFCCC) recognised the agricultural sector for its significant mitigation potential in the global efforts to stabilise GHG concentrations in the atmosphere. Moreover, the commitments and responsibilities agreed by the UNFCCC Kyoto Protocol include the development, dissemination and adoption of mitigation technologies that reduce GHG emissions from agriculture (UNFCCC, 2008). Since the 2013 reform of the EU's Common Agricultural Policy (CAP), farmers have had to comply with new environmental requirements — often referred to as greening — to receive the full amount of their subsidies (about 30 % of their direct payments).

Agriculture is well known to be a huge contributor of GHG; however, measurement continues to be a challenge. Despite considerable interest in a better characterization of agricultural sector emissions and GHG mitigation potential, available data on global agricultural sector emissions lags data development on fossil fuel emissions [4]. Despite the mandate from various government regulations and policies, lack of simple, user friendly methods of measurement and monitoring have made GHG emissions management extremely challenging for farmers. We explore the challenges of existing methods in the literature review section and in the following sections present 'Preserve' a low cost, real-time, user-friendly solution to manage GHG emissions in a farm.

II. Literature Review

A. *Estimating GHG emissions from farms*

Recent studies have employed ML (Machine Learning) algorithms for estimating GHG emissions from farm data. [5] Researchers have included a variety of farm types, including dairy, non-dairy, crop, minimal grazing, and mixed farms in their GHG modelling [5]. These models typically include multitude of attributes, such as type of farm, size of land, number of livestock, types of fertilizers applied, quantities applied, liquid and solid wastes, energy consumption etc. Some researchers employed face to face questionnaires for data collection [6], others used datapoints such as animal count, food, and production data from dairy farms [7]. Sensor based methods have also been deployed and studied [8] alongside simulation models. However, simulation models are far from reality since real farms are quite complex to simulate [9].

Data analytics approaches like machine learning can fill the gaps of prior methods such as simulation. Various statistical and machine learning approaches are usually adopted for event detection [10]. Modelling solutions use data and algorithms to estimate GHG emissions based on factors such as the type of crops being grown, the methods being used to produce them, and the size of the operation. Satellite-based measurement methods also provide unique opportunities to improve estimations of GHG emissions [11]. Satellite-based solutions use satellite imagery and remote sensing data to detect and monitor GHG emissions in the agriculture sector. Direct measurement solutions involve collecting data on GHG emissions directly from farms or other agricultural operations. This can be done using sensors, monitoring equipment, and other tools, however, this the most difficult method, as there are unique challenges to developing agricultural data over large spatial scales [12]

B. *Challenges in GHG measurement from Agriculture*

Despite several novel ways to measure from self-reported surveys to satellite based models, several drawbacks exist which impact the adoption and usage of these methods by small and mid-size farms. Table 1 provides a review of the drawbacks of several tools currently available for GHG emissions measurement from farms.

TABLE I. Review of Existing Tools and Their Drawbacks

Tool	***Drawbacks***	***Type***
Global GreenHouse Gas Watch	No real-time tracking, uses monthly measurements and approximations	Satellite based
Carbon AG	One-off measurements only require manually moving the device across the farm	Direct Measurement
AGRISAR	Data gathered twice weekly, not suitable for small farms	Satellite based
Guardian NG	Monitors one type of emission, no cloud storage, farmer friendly UI	Direct Measurement
GrASTech	Relies heavily on animal-based measurement alone	Direct Measurement
GHGsat	Not suitable for small farms	Satellite based
IpCC Software	Requires extensive training to use	Calculation software
ALU Software	Approximate based on national averages	Calculation software
Carbon Navigator	Self-reporting based on auditing the farm, prone to manual errors.	Auditing tool
CoolfarmTool	Online calculator, relies on farmer's self-reporting	Modelling software

CometFarm	Reliant on soil data from the United States, no in-situ measurement	Modelling solution
AgriCalc	Self-reporting based on auditing the farm, prone to manual errors.	Modelling software
Normative	Expensive and unaffordable for small farms	Modelling software

Given these drawbacks and the continuing challenges for small scale farmers to measure GHG cases in a simple cost-efficient manner, we developed Preserve. In the next section, we describe the system design and the prediction model.

III. System Design

A. *Sensor Hardware Prototype Development*

Phase 1 - Low-cost Component Selection: We began by evaluating and reviewing all available components necessary to create a low-cost, cloud-integrated solution. Sensors and supporting electronic circuitry are an integral part of the hardware setup. Existing sensors like the BME688 which has AI (Artificial Intelligence) built-in support were a natural choice, however, a closer look reveals a few challenges. The features provided by BME688 sensor, while useful, are very expensive and in recent times hard to procure due to significant supply chain shortages.

TABLE II. Sensitivity Testing and Calibration

Measurement Variable	***Sensor***	***Control (Mean)***	***Actual (Mean)***	***Mean Difference***
Temperature	10KThermistor	20.00*C	19.87*C	0.13*C
Temperature	DHT11	20.02*C	21*C	0.98*C
Humidity	DHT11	30%	25%	5%
Soil Moisture	MoistureSensor	61%	59%	2%
Gas Emission (Methane)	MQ9	300 ppm	306 ppm	6 ppm

As a low cost but effective alternative a WIFI-based Arduino microcontroller is a perfect solution and allows custom cloud-integrated solutions. After similar rounds of review of various sensors and other electronic components, the following materials were shortlisted.

- NodeMCU Board - Serves as a WIFI integrated Microcontroller
- MQ9 Gas sensor - To monitor Methane and other gas emissions
- 10K Ohm NTC Thermistor - To read temperature to 2.d.p
- Soil Moisture Sensor - To get water content in soil
- DHT11 - To read humidity and verify measured temperature

B. *Sensitivity Testing*

To assess the accuracy of the sensors, several measurements were made alongside a control that was already calibrated. Further care was taken to mimic the real-world conditions in a lab setting. For temperature we used a calibrated radiator using an industry grade thermostat controller. We took humidity measurements by placing a humidifier inside a box measuring 1 meter cubed. Soil moisture measurements were made by pouring 30ml of water in a pot with dried fertiliser absorbent pellets. Gas Emissions were tested using a specific concentration of fresh cow manure.

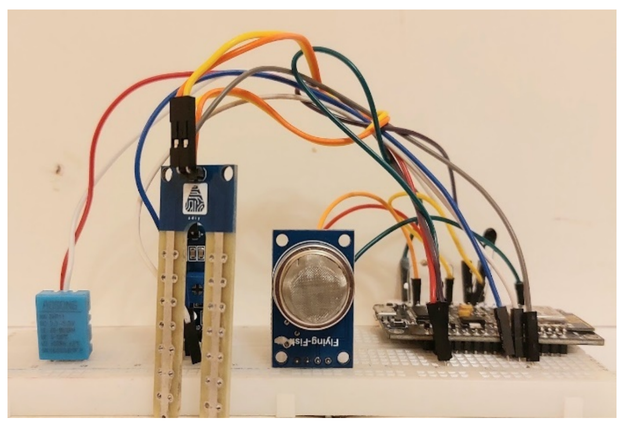

Fig. 1. Simple Prototype of the Circuit Hardware

C. *Spatial density of deployments*

This is a crucial factor for Preserve and a tradeoff between cost and accuracy. We want to keep Preserve an affordable solution for small farms by keeping our costs low. The number of sensor units deployed has a direct impact on the cost of ownership. After talking to small farms and visually inspecting a few farms, we minimized the number of sensor units by prioritizing deployment in problem-prone points and frequent maintenance checks by using a gyrometer attached that sent an alert when a sensor moved out of position. There is an opportunity to capture wind speed and look for correlation with emissions.

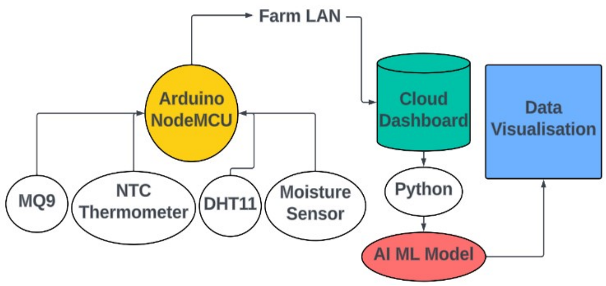

Fig. 2. Block diagram of various Preserve modules

During the deployment phase, spanning over a one-two-month period problem-prone areas will be identified through sampling various locations. This enables Preserve to identify hotspots in emissions while generalizing over other measured locations without further re-calibration. The size of the farm has a direct correlation with the number of sensors required, however, using calibration during the deployment phase, we are able to effectively identify hotspots for sensor placement.

IV. Predicting Emissions

A. *Dataset Development*

The dataset was developed through continuous real-time collection of sensor data. N_2O has high Global Warming Potential which is 298 times higher than CO_2. When soil temperature rises and high moisture conditions prevail during rainy months N_2O release is known to increase. The dataset from Preserve sensor unit was used to measure Volumetric Water content and Temperature at different distances (30 - 150cm, with an interval of 30cm from the tip of the sensor to the surface of the soil) and local emissions Nitrous Oxide from subsoil).

B. *Model Development*

We use multiple linear regression (MLR), a statistical technique that predicts the future values of a dependent variable by using the values of two or more independent variables. MLR assumes that there is a linear relationship between predictor and target values, the residuals are normally distributed, and the independent variables are not correlated with each other. We checked these assumptions by calculating the correlation coefficients, and then normality tested using the KS test and the variance inflation factor to test for multicollinearity.

Results of the multiple linear regression indicated that there was a very strong collective significant effect between the X1, X4, X6, X10, and Y, ($F(4, 194) = 94.71$, $p < .001$, $R2 = 0.66$, $R2adj = 0.65$).

$$\hat{Y} = -16.47 + 7.39\,X1 + 50.67\,X4 + 0.54\,X6 + 0.37\,X10$$

The $\hat{Y}$ represents the amount of future Nitrous Oxide emissions and the variables X1, X4, X6, and X10 represent Volumetric Water Content at 30cm & 150cm and Temperature at 30cm & 120cm and respectively. The individual predictors were examined further and indicated that X4 ($t = 6.553$, $p < .001$) and X6 ($t = 5.197$, $p < .001$) and X10 ($t = 6.546$, $p < .001$) were significant predictors in the model, and X1 ($t = 1.588$, $p = .114$) was non-significant predictor in the model.

Goodness of fit: Overall regression: right-tailed, $F(4,194) = 94.712934$, p-value = 0. Since p-value < α (0.05), we reject the H0. The linear regression model, $Y = b0 + b1X1 + ... + bpXp + \varepsilon$, provides a better fit than the model without the independent variables resulting in, $Y = b0 + \varepsilon$. The Y-intercept (b): two-tailed, $T = -8.162802$, p-value = 4.00791e-14. Hence b is significantly different from zero. There is no multicollinearity concern as VIF values are less than 2.5.

C. *Python Automation*

The model developed in the previous section entailed using a one-off data gathering exercise at a local farm to demonstrate that usefulness of regression modelling in predicting NO_2 emissions. However, an average farmer will not be able to run such an analysis. To solve this, a python module was built which periodically revises the prediction model parameters using datapoints from the most recent measurements gathered from Preserve installed in each farm.

V. Cloud User Interface

To visualise data, predictions, and other insights, Preserve utilises a custom Cloud Dashboard developed through an open source Thingspeak API [12] and IFTTT (If This Then That) [13]. We used the Thinkspeak API to add a custom module to send data from the Preserve Arduino real-time sensor data. Preserve does not collect any personally identifiable data considering user privacy. Data collected in the farms via the Preserve sensors are uploaded often to the cloud and a farm owner can in real-time view the latest metrics from their farm.

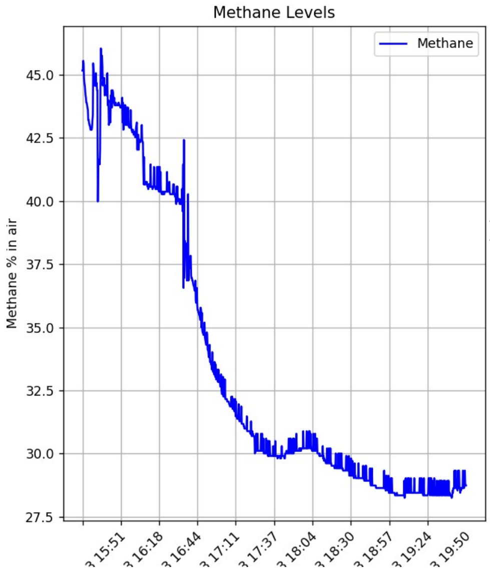

Fig. 3. Cloud UI showing measurements at different times of day

Further a farm owner can set thresholds for emissions and the prediction model starts alerting the farmer before the emissions reach higher levels, giving the farmer an intervention opportunity. To further improve the user experience of a busy farm owner, we use IFTTT to send SMS alerts or email alerts when pre-defined monitoring criteria are met.

Improved monthly view metrics and goal target user interface are currently under development and testing. Figure 4 shows the mock user interface.

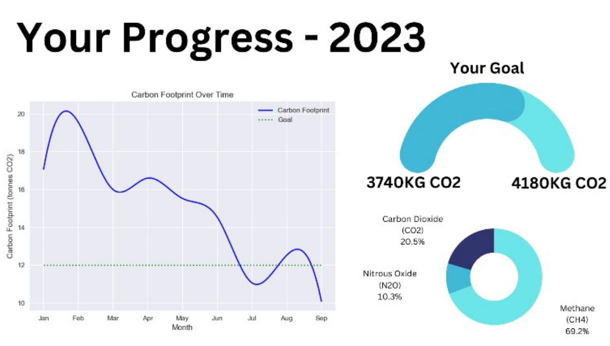

Fig. 4. Mocks from Improved Cloud User Interface

VI. Discussion

A. Implications

Overall, Preserve met the various goals we set out to accomplish, enabling an average farmer owning a small to medium-sized farm to manage their GHG emissions. Table 3 summarizes the feature set of Preserve and the impacted goals.

TABLE III. Preserve Design Considerations

Goal	*Preserve Design Consideration*
Low-cost	Use of affordable sensors that have been sensitivity tested and pre-calibrated in lieu of expensive sensors that offer precision but at unaffordable costs. Use of open source APIs affordable sensors that have been sensitivity tested and pre-calibrated instead like thinkspeak API have also contributed ot keeping costs low at around €50.
Real-time data	Preserve uses a live sensor tracking through continuous in-situ measurements at various points within the farm.
Predictive to allow prevention	Using MLR methods along with repeated Python based model testing and re-build, Preserve offers prediction support to farms that wish to keep tabs of rising emission levels, intervene and take corrective action before they exceed allowable limits
Access	Cloud based dashboard allows a farm owner to access the sensor data even remotely.
User Friendly	The intuitive UI offers a quick and straightforward way to understand the emissions status at a given point and also look at historical trends and future projections Farm owners can also setup alerts that can be triggered through custom threshold configurations based on individual farm needs.

B. Limitations

In this first version of Preserve, we focused on leveraging the relationship between soil moisture and temperature on GHG emissions. Several researchers [14],[15],[16] have demonstrated that there is a linear relationship between soil moisture, temperature and several GHG emission rates. This was the primary reason for us to use linear regression. There is of course a debate if linear regression models are the best approach, in fact, researchers also argue that simple regression models linking the main driving factors (temperature, moisture, soil type) with outputs (e.g., greenhouse gas emissions) by means of simple linear relationships with empirical coefficients have low predictive power. This further validates why our in-situ measurements only had an R^2 of 0.7 to 0.8. GHG emissions are also influenced by several factors beyond the soil moisture and temperature, as we add more variables into the mix, it is likely that a linear regression model may not be appropriate. We plan to use methods such as factor analysis and principal component analysis to determine which predictors can best explain the variance and then determine the regression model we might need to build. To do so, a larger, robust dataset is required. We are currently working with small farms to implement longer term data collection measures. This will be part of our future research.

We will also need different models for several types of farms, for example, farms with livestock vs farms that do not have livestock. Given a farm is a dynamic environment, we plan to address this limitation by building a prediction model that uses reinforcement learning techniques and use human-in-the-loop models with feedback from the farmer on model prediction and performance.

Another limitation of Preserve is its reliance on Wi-Fi communication, covering less than 1km (ESP8266). Using communication protocols like LoRaWAN or NBIOT has been proven to be effective over long distances typically upwards of 15km and 10km in rural farm environments respectively while using significantly less power. In the future, we seek to implement these communication protocols, preferably LoRaWAN (Arduino MKR WAN-1310) due to built-in TTN (The Things Network) support, though twice the initial cost.

C. *Future Research Directions*

In our opinion, Preserve addresses the drawbacks of several existing methods (including satellite and modelling solutions), however, there is an opportunity for improvement which provides us with the ability for future research and development.

Firstly, there are several challenges pertaining to data collection, for example, different farm structures (e.g., plant-based farms, cattle-only farms, mixed farms) will require the implementation of different models. We are studying volunteer farms and note the adjustments needed based on the nature of the farms.

Secondly, the sensor data generated from Preserve while is readily available for sharing, further design considerations are needed for interoperability. The key challenge of integrating different livestock agricultural systems is how to deal with the heterogeneity of multiple information resources. There is tremendous opportunity to create an international standard around emissions data coming from IoT and Sensor based networks that allow for comparison, portability, and auditability of datasets, though there are existing standards like Asset Administration Shell (AAS) or OPC/UA that face challenges due to a lack of official/published standards.

Thirdly, we use an MLR based model for prediction, however, we may encounter farms with emission patterns that may not lend well to linear regression. Sophisticated non-linear prediction models and heuristics-based approaches need to be developed to improve accuracy of prediction models.

Lastly, while the goal is to automatically manage emissions data without human intervention, there is abundance of opportunity to use input from the farm owners at early-stage deployment and testing, including use of farmers feedback on data collected to build a human in the loop system.

VII. Conclusion

In conclusion, Preserve is a low-cost Wi-Fi based, real-time GHG management solution that is designed to help farmers and producers measure and optimise emission rates caused by the agricultural sector. Using predictive analytics and multi-regression models in Python, Preserve helps farmers and producers to take preventative measures when the model predicts an impending emissions spike. Preserve can help farmers and producers to identify opportunities for reducing emissions and improving their operations, helping address the issue of climate change and improve air quality.

References

[1] Intergovernmental panel on climate change. World Meteorological Organization. 2007;52.

[2] J.Broucek, "Production of methane emissions from ruminant husbandry". Journal of Environmental Protection. 2014;5(15):1482.

[3] Environmental Protection Agency GHG report, Ireland's National Inventory Report 2021

[4] FN Tubiello, Salvatore M, Cóndor Golec RD, Ferrara A, Rossi S, Biancalani R, Federici S, Jacobs H, Flammini A. "Agriculture, forestry and other land use emissions by sources and removals by sinks". Rome, Italy. 2014 Mar.

[5] E. Pérez-Miñana, Krause PJ, Thornton J. "Bayesian Networks for the management of greenhouse gas emissions in the British agricultural sector". Environmental Modelling & Software. 2012 ;35:132-48.

[6] HO. Hosseinzadeh-bandbafha, Safarzadeh DA, Ahmadi EB, Nabavi-Pelesaraei AS. "Modelling output energy and greenhouse gas emissions of dairy farms using adaptive neural fuzzy inference system". Agricultural Communications. 2016;4(2):14-23.

[7] M. Zhang, Wang L, Feng H, Zhang L, Zhang X, Li J. Modeling method for cost and carbon emission of sheep transportation based on path optimization. Sustainability. 2020;12(3):835.

[8] T. Norton, Chen C, Larsen ML, Berckmans D. Precision livestock farming: Building 'digital representations' to bring the animals closer to the farmer. Animal. 2019 ;13(12):3009-17.

[9] T. Norton , Berckmans D. Developing precision livestock farming tools for precision dairy farming. Animal Frontiers. 2017;7(1):18-23.

[10] F. Adrion, Brown-Brandl TM, Jones D, Gallmann E. Identifying anomalous decreases in feeding time of grow-finish pigs. In10th International Livestock Environment Symposium (ILES X) 2018 (p. 1). American Society of Agricultural and Biological Engineers.

[11] T. Matsunaga, Maksyutov S. A Guidebook on the Use of Satellite Greenhouse Gases Observation Data to Evaluate and Improve Greenhouse Gas Emission Inventories, Satellite Observation Center. National Institute for Environmental Studies Japan, 2018

[12] MA. Maureira, Oldenhof D, Teernstra L. "ThingSpeak–an API and Web Service for the Internet of Things". World Wide Web. 2011;25:1-4.

[13] S. Ovadia. "Automate the internet with "if this then that"(IFTTT). Behavioral & social sciences librarian. 2014 ;33(4):208-11.

[14] J. Gao, Yan Y, Hou X, Liu X, Zhang Y, Huang S, Wang P. "Vertical distribution and seasonal variation of soil moisture after drip-irrigation affects greenhouse gas emissions and maize production during the growth season." Science of the Total Environment. 2021

[15] J. van Lent, Hergoualc'h K, Verchot L, Oenema O, van Groenigen JW. Greenhouse gas emissions along a peat swamp forest degradation gradient in the Peruvian Amazon: soil moisture and palm roots effects. Mitigation and Adaptation Strategies for Global Change. 2019 Apr 15;24:625-43.

[16] S. Mangalassery, Sjögersten S, Sparkes DL, Sturrock CJ, Craigon J, Mooney SJ. To what extent can zero tillage lead to a reduction in greenhouse gas emissions from temperate soils?. Scientific reports. 2014