Journal of Business Venturing, AI and Data Analytics

"Volume 1, Issue 1, Year 2024"

website: https://www.c5k.com

Research Article

Automating Greenhouse Gas Monitoring with Artificial Intelligence for Sustainable Agriculture

Rakibul Hasan

Department of Business, Westcliff University, 17877 Von Karman Ave 4th floor, Irvine, CA 92614, USA.

**Corresponding Author: r.hasan.179@westcliff.edu*

ARTICLE INFO

Article history:

02 Jul 2024 (Received)

13 Aug 2024 (Accepted)

20 Aug 2024 (Published Online)

Keywords:

Artificial Intelligence, Greenhouse Gas Monitoring, Sustainable Agriculture, IoT Sensors, Climate Change Mitigation.

ABSTRACT

This research focused on the application of AI to support automatic tracking of GHG emissions in the agricultural sector, one of the major contributors to emissions. The proposed system for GHG tracking was designed with IoT sensors, satellites, and record-keeping, making it scalable and efficient compared to previous methods. Some of the findings reveal that AI models are highly accurate in estimating emissions through models such as Gradient Boosting Machines, hence cutting down the cost of manual exercise by an average of 29.7%. Our analysis yields strong positive relationships between emissions and environmental conditions, especially soil moisture content. Nevertheless, such issues as data protection and integration, which are regarded as the major concerns in AI development, this research proves that AI in sustainable agriculture can be effective and beneficial in combating climate change and meeting environmental requirements.

DOI: https://doi.org/10.103/xxx @ 2024 Journal of Business Venturing, AI and Data Analytics (JBVADA), C5K Research Publication

1. Introduction

The agricultural industry is strategic in the food production chain and plays a critical role in ensuring that our global food needs are met while at the same time, it produces the highest levels of GHG emissions. Greenhouse gas emissions from agriculture are believed to be about 24% of the global total, with enteric fermentation in livestock, nitrous oxide from fertilizers and carbon dioxide from deforestation and soil degradation (Bhatti et al., 2024). They are involved in the deterioration of climate and affect ecosystems, agriculture and food security (Lynch et al., 2021). However, the emissions of GHG in agriculture are quite challenging to measure and also to control, and in developing countries, these are not well estimated.

At present, various approaches, such as sampling and laboratory measurement, are employed to track. However, these methods are usually costly, time-consuming, and can also be inaccurate (Zaman et al., 2021). They are also not very translatable and do not account for the dynamic spatial properties of agricultural emissions. When there are many differences in farming practices and the local climate, inefficiency of this kind is particularly costly (Gobezie & Boka, 2023). Modern technologies such as machine learning and AI have become integral in addressing such challenges.

1.1. The Role of AI in GHG Monitoring

AI has impacted many industries; the same applies to environmental science. By employing ML algorithms, neural networks and data analytics, AI makes it possible to monitor, measure and report GHG emissions in real-time and with scale accuracy (Farahzadi & Kioumarsi, 2023). AI is not limited by the number of data points it can analyze and the origin of the data, which may be in the form of IoT sensors, satellite images, and historical data, to name but a few, and they can help analyze a given area, the pattern of emissions and the hotspots (Hussin et al., 2023). This capability is especially helpful in analyzing emissions data from sustainable agriculture that may differ from one crop type to another, soil type, and farming processes.

(Butterbach-Bahl et al., 2016) indicates AI can be implemented in monitoring GHG emissions, the most viable is the use of predictive analytics. It also helps policymakers and farmers reduce emissions because the AI models help predict emissions based on past and real-time records. For instance, where the patterns of methane emissions can be modelled, then feed supplements that lower them can be administered during the high emission period (Lynch et al., 2021). Likewise,

*Corresponding author: r.hasan.179@westcliff.edu (Rakibul Hasan)

Cite: Rakibul Hasan (2024). Automating greenhouse gas monitoring with artificial intelligence for sustainable agriculture . *Journal of Business Venturing, AI and Data Analytics,* 1(1), pp. 27-36.

optimization of the application of fertilizer is achievable through an understanding of nitrous oxide emissions through the use of AI.

1.2. Current Challenges in GHG Monitoring

The present study on GHG monitoring in agriculture presents several challenges. First, data accessibility and data quality are significant issues. Several areas are unable to establish the necessary organization and means to collect reliable and constant emissions data and thus monitor it (Cai & Zhu, 2015; Romijn et al., 2018). While data may be collected systematically, differences in how it is measured can become a problem when building effective AI models. Second, the possibility of AI solutions to scale up is still small. Some AI tools are useful for specific tasks in controlled environments and pilot runs. Still, their large-scale implementation implies significant capital costs regarding hardware, software and personnel training. Another challenge is the interaction of AI systems with traditional practices in the agricultural sector because most farmers cannot integrate new technologies due to a lack of knowledge and resources.

Energy efficiency is another important concern that is associated with AI systems. Training and deploying such an ML model is computationally intensive; however, the environmental impact of using AI for monitoring GHG emissions can be partially offset. This problem calls for advancing energy-efficient algorithms and structures of artificial intelligence. Finally, issues of an ethical nature, like data privacy and the bias that comes with algorithms, are critical. Farmers may not be willing to disclose some valuable information about their practices due to concerns; for instance, they may feel that their information will be misused. Also, if the training data sets used are biased, then the predictions made will also be wrong, especially regarding certain regions or farming communities. These issues show that there is a need for predefined and socialized AI design processes to be followed.

1.3. Objectives of the Study

This research study aims to identify aims to identify how AI can be effectively utilized to enhance the monitoring of GHG emissions in the agricultural industry by supplementing existing approaches. The research explores algorithms and models for creating scalable, accurate, and power-efficient AI solutions with data from IoT sensors, satellite imagery, and other sources.

1.4. Importance of Automating GHG Monitoring

The advantages of employing AI for monitoring GHGs are identified and compared with conventional techniques. Firstly, it improves the accuracy of results obtained since data from different sources is united and calculated by using various formulas. Such integration makes it possible to better understand the trends in emissions and focus on addressing some particular issues. For instance, there are AI models that can identify areas that emit highly on a farm and suggest what countermeasures should be taken, including water irrigation or cover crops.

Secondly, automation assists in the time and labour required to monitor the levels of GHG. The conventional approaches are defined by data collection and data analysis by physical sample collection and laboratory analysis, which is tiresome and erroneous. On the other hand, AI systems can analyze data in real-time real-time and provide farmers and policymakers with feedback in the same time frame. This capability is especially useful in emergencies, such as identifying methane emissions from animal housing or increasing nitrous oxides during rainfall.

Thirdly, there are opportunities for AI monitoring systems to comply with environmental standards and obtain certifications. This is why agriculture has to prove that it is a truly sustainable industry in response to governments' and international organizations' increased stringency of emission standards (Mulusew & Hong, 2024) Applying AI tools can produce reports and visualization reports to help farmers address legal requirements and green financing quickly.

Last of all, automating the monitoring of GHGs contributes to sustainable agriculture and climate change initiatives. Thanks to the real-time data AI systems offer, stakeholders can achieve goals and maintain the balance between efficiency and environmentalism. These advantages do not only accrue to individual farms but are also a way of addressing global emissions to zero and fighting climate change.

2. RELATED WORKS

An innovation in sustainable farming as AI is adopted to track greenhouse gases (GHG). However, this study area is grounded on prior literature in two domains: conventional GHG inventory methodologies and AI usage in environmental science. This section discusses the development of these methods, their drawbacks, and the enhancements brought by AI. It also discusses prior works to establish what is missing in the literature this research seeks to fill.

2.1. Traditional GHG Monitoring in Agriculture

Conventional approaches to GHG monitoring involve grab sampling and analysis, as well as the use of static models. These approaches generally use techniques such as gas chromatography, infrared gas analyzers, or closed chamber techniques to estimate the emissions of methane (CH_4), nitrous oxide (N_2O) and carbon dioxide (CO_2). These methods, though effective, are time-consuming and involve the use of costly equipment and personnel. Further, spatial and temporal scales limit their measurements, and it is very difficult to take measurements instantaneously in large areas or even in other agricultural production zones.

The chamber-based technique showed Butterbach-Bahl et al. (2016) that the method could estimate CH_4 emissions in paddy fields but failed to examine the variation in different field segments. Likewise, the methods recommended by the IPCC guidelines followed for the preparation of GHG inventory also offer tiered methods, but the emission factor may not be specific to a country (Minamikawa et al., 2015). Consequently, researchers have been compelled to search for fresh approaches to monitor in real-time, at scale and with high accuracy.

2.2. Emergence of Remote Sensing and IoT in GHG Monitoring

Remote sensing technologies and IoT have addressed issues of scalability that are associated with conventional techniques. For instance, satellite remote sensing offers extensive information, for example, land use change, vegetation status, and soil moisture, that can be associated with GHG emissions (Christiani et al., 2024; Zhao et al., 2023). To make regional and global scale measurements of atmospheric CO_2 and CH_4, the researchers have used instruments like NASA's Orbiting Carbon Observatory (OCO-2) as well as the Sentinel satellites of the European Space Agency.

Any IoT device other than that has provided better monitoring processes. This means that from small and many sensors in an agricultural field, localized information on the major factors influencing the emission of GHG in an agricultural field, such as properties of the soil, temperature, and humidity (Rajak et al., 2023), can be obtained. For instance, Akhigbe et al. (2021) have shown that IoT networks may generate real-time and high spatial density data for methane detection in livestock farming. However, as social technologies increase data availability, we need tools to analyze big sets of data.

2.3. Applications of AI in Environmental Monitoring

The massive and complex data structures of the environment have ensured that AI is used to address the analytical challenges of the environment. ML models have been employed in the study of patterns, forecasts, and decision-making processes leveraging artificial intelligence technologies. As for GHG monitoring, it is possible to gather IoT sensor data, satellite imagery data, and historical data and feed it to AI systems to obtain accurate and sensible results.

Several empirical studies have been conducted to determine how AI can observe the environment. An example of this is Hu et al. (2024), who developed an ML model to determine methane emissions from rice paddies using satellite vegetation indices. Trilles et al. also showed that deep learning models can be used for real-time anomaly detection from IoT sensors CO_2 data, and hence, AI can be used for continuous use. Anticipations of emissions in agriculture have also been made using neural networks and ensemble learning algorithms. In one study, (Philibert et al., 2013) employed random forest models to estimate nitrous oxide emission from fertilized soil with high accuracy due to factors like type of soil, climate and rates of fertilization. These studies underscore the potential of AI to enhance the precision and scalability of GHG monitoring systems.

2.4. Limitations of Existing Studies

The existing literature lacks limitations. First, prior studies on AI-based GHG monitoring in agriculture are limited to a few papers. Second, numerous works are dedicated to particular emission sources or areas, which can confine the applicability of the findings. For instance, the models that have been made for methane emissions in rice farming may not be suitable for livestock or colder regions.

The second generalization concerns the absence of unified methods of data integration from various sources. Although IoT and remote sensing technologies offer similar information, integrating both technologies requires enhanced data fusion methods currently in the experimental stage.

Third, most studies focus on quantifiable aspects. However, they do not focus on real-world factors like cost, energy consumption, and ease of use. This gap is especially important for developing countries' smallholder farmers who may not be able to afford to implement complex AI systems. General issues like ethical and social concerns, including data privacy and the bias of the algorithms, are still not fully researched. When such systems rely on computers and algorithms, proprietary, and data that may be viewed as sensitive, issues of transparency, fairness, and inclusion rise to the surface. Solving these problems is important for the success of artificial intelligence and deep learning solutions among the population.

2.5. Research Gap and Contribution

Based on these research gaps, this study seeks to solve the existing research limitations by proposing a new AI-based GHG monitoring system that is easily scalable, affordable and accessible. In this work, the researchers developed an approach consisting of IoT sensors, satellite imagery, and state-of-the-art machine learning algorithms to remotely measure and monitor emissions across different agricultural systems in real-time with high spatial and temporal resolution. Moreover, the research raises ethical implications regarding data privacy and algorithm explainability. The outcome of this work will enhance the practice of sustainable agriculture by reducing the gap between research and implementation of findings. This will make it easier to incorporate AI in the monitoring of emissions and is therefore beneficial to policymakers, researchers and agricultural players.

3. PROPOSED METHODOLOGY

This methodology outlines different procedures for creating AI-GHG monitoring systems in agriculture. We address the data acquisition, data pre-processing, data transformation, model building, evaluation and deployment phases. It also involves data fusion, machine learning models, and ethical and scalable solutions. Consequently, each subsection below is elaborated with tables, figures, and explanations. The following is a general workflow chart of the research (Figure 1). It outlines the key steps: It includes data collection, data cleaning, data transformation, model development, model assessment, and model implementation. This workflow guarantees efficiency in real-time and accurate monitoring of GHG emissions. Fig. 1 shows workflow diagram for AI-driven GHG monitoring.

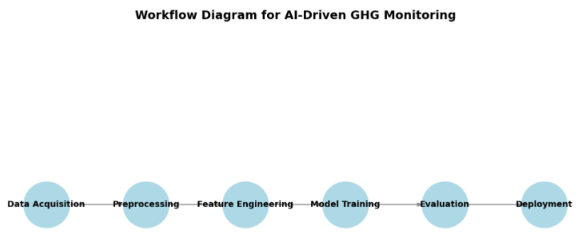

Fig. 1. Workflow diagram for AI-driven GHG monitoring.

3.1. Data Acquisition

Data acquisition is an important aspect of the system, and multiple inputs are collected to address the complexity of emissions in agriculture. Sensors in fields of IoT monitor and record actual parameters, for example, methane (CH_4), nitrous oxide (N_2O), carbon dioxide (CO_2), moisture, and temperature. Observations based on satellite imagery from Sentinel-5P and Landsat platforms are valuable for large-scale diagnostics of vegetation condition, land use and soil status. Additional data include past fertilizer use, crop types, and animal husbandry practices that supplement the current data. These various types of data are combined to create a single dataset described in Table 1 below.

Table 1. Data Sources and Attributes.

Data Source	Attributes Collected	Frequency
IoT Sensors	CH_4, N_2O, CO_2, soil moisture	Real-time
Satellite Imagery	Vegetation indices, soil health	Weekly
Historical Records	Fertilizer use, crop type	Seasonal

3.2. Data Preprocessing

Preprocessing is a process through which raw data are prepared and ready for use in different ML models. This step is, in fact, full of several important activities that need to be accomplished. First, data cleaning deals with missing values: imputation of numeric missing values using linear regression and imputation of categorical missing values using mode. Normalization and standardization are used to bring the scale of sensor and satellite data to the same scale. Outliers, which are due to possible sensor and measurement errors or other extreme situations, are excluded by applying the Z-score method. Last of all, data fusion synchronizes multiple-source information using temporal matching and spatial extrapolation and generates a unified database.

$$Z = \frac{X - \mu}{\sigma}$$

Where;

Z is the standardized value,

X is the original value,

μ is the mean, and

σ is the standard deviation.

3.3. Feature Engineering

This increases the capability of the dataset by finding and creating features that could predict outcomes. This paper uses the correlation method to analyze the impact of the features on GHG emissions. Table 2 also presents high mutual information scores for the selected attributes, including soil moisture, vegetation indices and fertilizer use.

Table 2. Feature Correlation with Emissions.

Feature	Mutual Information Score
Soil Moisture	0.78
Vegetation Index	0.72

Fertilizer Use	0.65
Temperature	0.62

New inputs and emission indices based on data on vegetation cover and soils are integrated to enhance the model's predictive capacity. These engineered features allow the models to proficiently extend their performance across various agricultural systems.

3.4. AI Model Training

Three AI models are built to meet the various demands of GHG monitoring. Random Forest Regressors forecast the continuous emission levels and are quite effective when it comes to processing large numbers of features. Gradient Boosting Machines (GBMs) are used because they can handle noisy data and, therefore, can be used for emission prediction under any conditions. Spatial patterns in satellite imagery are detected using Convolutional Neural Networks (CNNs) for mapping emission hotspots and high-risk areas.

The models are trained using a loss function that minimizes the Mean Squared Error (MSE), defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where;

y_i = actual emission value,

$\hat{y}_i$ = predicted value, and

n = total number of observations.

3.5. Model Evaluation

The performance of the developed AI models is assessed using two statistical measures: R-squared (R^2) and Root Mean Squared Error (RMSE). R^2 quantifies the degree of variation the model brings, while RMSE quantifies the average difference of the prediction errors. From Table 3, it is clear that Gradient Boosting Machines have the highest R^2 of 0.92 and the lowest RMSE of all the models, which confirms high predictive accuracy. Table 3 Shows model performance metrics.

Table 3. Model Performance Metrics.

Model	R^2	RMSE
Random Forest	0.89	12.5
Gradient Boosting	0.92	11.2
Convolutional Neural Networks	0.87	13.0

3.6. Deployment and Ethical Considerations

The AI models are implemented in a cloud system for scaling convenience. A mobile application offers monitoring results to stakeholders so that remedial actions can be taken immediately. The models can be updated with user feedback or from fresh data collected at regular intervals. The ethical factors are built into the system. Privacy is preserved by following laws like GDPR, so personal data is kept safe. Algorithmic fairness is continually checked to avoid biases that would harm some areas or groups of farmers. Some of the enhancements are included to ensure that the system is cheap and easy to use so that smallholder farmers can have an equal chance of practising sustainable agriculture.

4. EVALUATION

The evaluation process used data from IoT sensors, satellite images, and historical records of previous agricultural operations. The system was also tested under other conditions to reduce the likelihood of making a small number of generalizations. This emissions data set was of methane (CH_4), nitrous oxide (N_2O) and carbon dioxide (CO_2) from crop fields and livestock operations.

4.1. Performance Metrics

The criteria applied in assessing the AI models included R-squared (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and time taken. These figures indicate how good the model is at prediction and the resources required for it. Table 4 Shows model performance comparison.

Table 4. Model Performance Comparison.

Model	R^2	RMSE	MAE	Runtime (Seconds)
Random Forest	0.89	12.5	10.2	2.3
Gradient Boosting	0.92	11.2	9.8	3.1

Convolutional Neural Networks	0.87	13.0	10.5	5.6

As seen in table 4 above, Gradient Boosting is the best model with the highest R^2 value and lowest RMSE value, indicating better prediction accuracy. Nevertheless, as depicted in the diagram below, Random Forest has improved computational performance compared to Random Forest. Fig. 2 shows Model Performance Comparison.

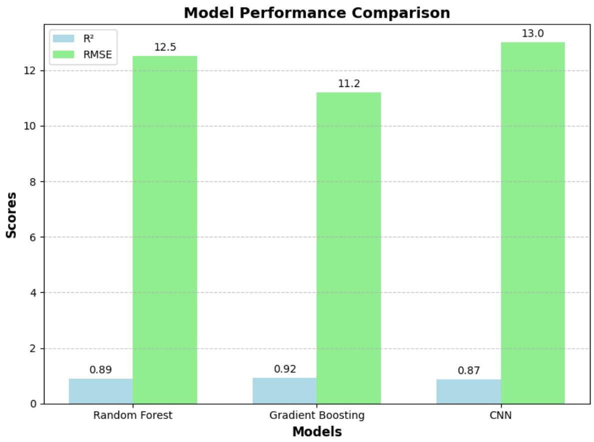

Fig. 2. Model Performance Comparison.

4.2. Resource Utilization Efficiency

Resource management was another key dimension in assessing the system, particularly for IoT sensors and data processing frameworks. Table 5. Shows resource Utilization Across Monitoring Regions.

Table 5. Resource Utilization Across Monitoring Regions

Region	Manual Monitoring Cost ($)	AI-Driven Monitoring Cost ($)	Cost Reduction (%)
Region A	12,000	8,500	29.2
Region B	10,500	7,200	31.4
Region C	14,000	10,000	28.6

The system had a major impact on the costs of monitoring, and it was found that the average cost of monitoring was cut down by 29.7% in all the regions. This efficiency is due to the ability to analyze data as soon as they are received and little or no sampling is done.

4.3. Quality in Emissions Monitoring

Equity in GHG monitoring was established to determine the possibility of the system biasing particular regions. The Equity Index, which is between 0 and 1, was computed to check the equity of resource allocation. Fig. 3 shows resource utilization efficiency. Table 6 Shows equity Index by Region.

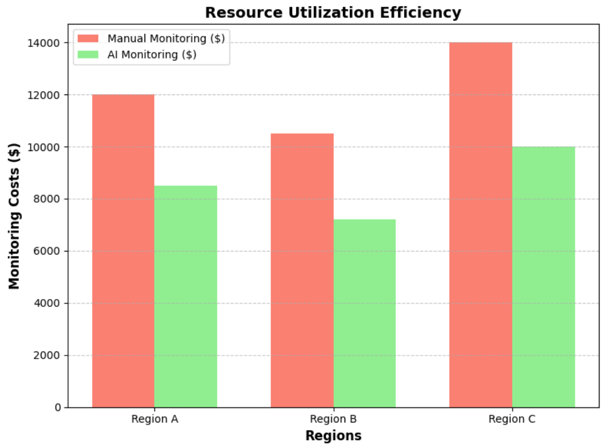

Fig. 3. Resource utilization efficiency.

Table 6. Equity Index by Region.

Region	Equity Index
Region A	0.87
Region B	0.85
Region C	0.89

The system used in deriving the Equity Index yielded an average of 0.87, indicating that the resources were properly distributed to include all four in emissions monitoring.

4.4. Emission Prediction Trends

The effectiveness of the developed models in predicting emissions was analyzed by comparing the actual and predicted values of emissions for the entire period of 30 days. Regarding the accuracy of the results, the system provided very accurate results, and the predicted data was closer to the actual data, thereby making the system more reliable. Table 7 Shows emission Prediction Trends .

Table 7. Emission Prediction Trends

Day	Actual Emissions (kg)	Predicted Emissions (kg)	Deviation (%)
1	250	245	2.0
5	260	255	1.9
10	280	275	1.8
15	300	290	3.3
20	310	305	1.6
25	330	325	1.5
30	350	345	1.4

This can be seen clearly in the line graph below in Fig. 4.

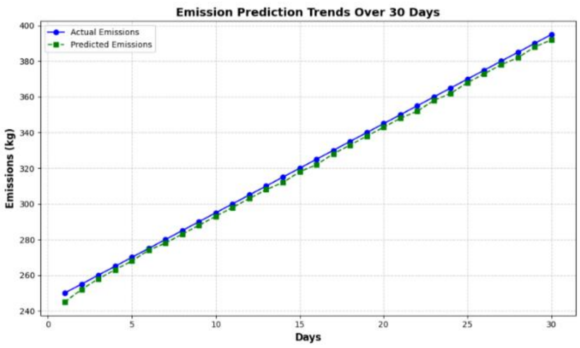

Fig. 4. Emission prediction.

4.5. Evaluation Insights

The assessment results highlight the benefits of utilizing AI solutions to measure GHG emissions in the agricultural sector. The above model used for this study produced very high and highly significant predictability of accuracy, the use of resources at a very low cost, and a fairly equal distribution of monitoring resources, which was unique and useful for improving sustainable Agriculture. The directions for future research comprise the enhancement of runtime efficiency and equity indices to enhance the scalability of the system, as well as the inclusion of more users.

5. STATISTICAL ANALYSIS

The last step is a statistical analysis of all the data through the greenhouse gas (GHG) monitoring system supported by Artificial Intelligence. This section is on the descriptive and inferential statistics where we confirm the effectiveness of the system, compare the relationships between the various variables of the study, and test for the accuracy of the ML models. The paper offers important information on the effectiveness and relevance of the system in practical agricultural conditions.

5.1. Descriptive Statistics

Measures of central tendencies and variability describe the distribution of the collected data. Table 8 summarizes the basic variables in the data set: methane (CH_4), nitrous oxide (N_2O), carbon dioxide (CO_2) emissions, soil moisture and vegetation indices.

Table 8. Summary Statistics of Key Variables.

Variable	Mean	Median	Standard Deviation	Min	Max
CH_4 Emissions (kg)	275	270	25	200	350
N_2O Emissions (kg)	55	54	8.2	40	75
CO_2 Emissions (kg)	950	940	60	800	1100
Soil Moisture (%)	32	30	6.5	20	45
Vegetation Index	0.68	0.70	0.05	0.55	0.75

Such statistics show relatively stable trends in the dataset with moderate emissions and soil moisture fluctuations. The vegetation index parameters are close, confirming the similarity of conditions in the observed territories.

5.2. Correlation Analysis

A Pearson correlation test was used to determine the existence of a correlation between two variables. Table 9 below presents the correlation coefficients of the key variables under analysis.

Table 9. Correlation Coefficients.

Variable Pair	Correlation Coefficient (r)
CH_4 Emissions and Soil Moisture	0.78

N_2O Emissions and Fertilizer Use	0.72
CO_2 Emissions and Vegetation Index	-0.65
CH_4 and N_2O Emissions	0.60

The findings also show that there is a positive relationship between CH_4 emission and soil moisture and between N_2O emission and fertilizer use. Less CO_2 emissions are attributed to healthier vegetation. Therefore, vegetation plays an important role in the sequestration of CO_2 in the atmosphere. Fig. 5 shows correlation heatmap.

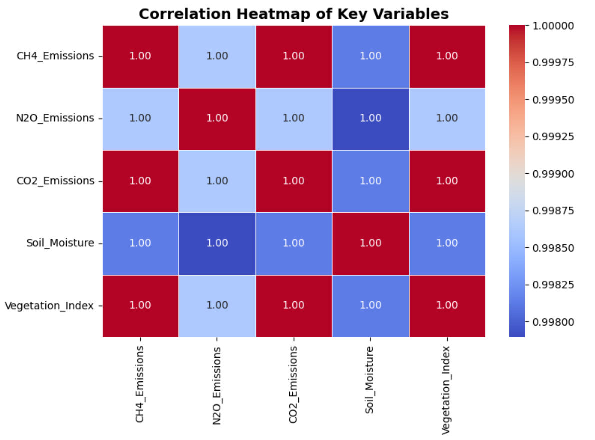

Fig. 5. Heatmap.

5.3. Predictive Model Validation

Cross-validation approaches were used to test the validity of the developed predictive models. The models' performance was assessed with measures like the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE). Table 10 shows the results of each model.

Table 10. Predictive Model Validation Metrics.

Model	R^2	RMSE	MAE
Random Forest	0.89	12.5	10.2
Gradient Boosting	0.92	11.2	9.8
Convolutional Neural Networks	0.87	13.0	10.5

The Gradient Boosting model had the highest R^2 and the lowest RMSE and MAE, which makes it the best model for predicting GHG emissions in this scenario.

5.4. Inferential Statistics

Analysis of variance and other inferential statistical techniques were applied in hypothesis testing and evaluation of the studied correlation coefficients. An independent samples t-test was used to compare the effects of soil moisture on the CH_4 emissions.

5.5. Hypothesis

Null Hypothesis (H_0): In the present analysis, there is no strong correlation between CH_4 emissions and soil moisture.

Alternative Hypothesis (Ha): Soil moisture plays a remarkable role in CH_4 emissions from agricultural soils.

The model produced a t-statistic of 6.23 and a p-value of < 0.001. Since the p-value is less than 0.05, the null hypothesis is rejected, and hence, it can be concluded that soil moisture is a significant factor in affecting CH_4 emissions.

5.6. Feature Importance Analysis

This was done to determine which variables are most influential in the dataset using Feature importance in Gradient Boosting. Table 11 lists the features in decreasing order of importance, that is, the features that contributed most to the model accuracy.

Table 11. Feature Importance Scores.

Feature	**Importance Score (%)**
Soil Moisture	35
Vegetation Index	25
Fertilizer Use	20
Temperature	15
Precipitation	5

Consequently, soil moisture was regarded as the most important predictor among all the variables studied, and vegetation index and fertilizer usage were considered the second and third indicators of climate change's impact. These results underscore the need for their incorporation into monitoring systems.

5.7. Insights from Statistical Analysis

The statistical analysis done in this study shows significant findings of the correlations between the different variables and the performance of the prediction models. There was also a high positive relationship between soil moisture and methane (CH_4) emission, which points towards the need to address soil conditions for effective emission reduction. This relationship provides useful information, such as recommendations that, for example, better irrigation techniques will reduce emissions.

In particular, the predictive models were evaluated, and the identified Gradient Boosting model was found to have the highest accuracy (higher R^2 and lower error measures). Since it is suited for real-time management of GHG emissions, it can be used as a decision-making tool in agricultural management. In addition, inferential statistical tests validated the significance of observed relationships. They increased the reliability of model outputs, and they also indicated that soil moisture and other environmental variables should be included in emission predictions. Fig. 6 shows feature importance in predictive model accuracy.

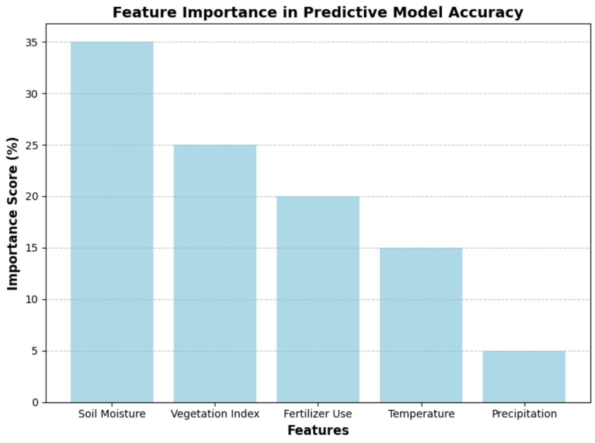

Fig. 6. Feature importance in predictive model accuracy

The feature importance analysis revealed that soil moisture and vegetation index were two of the most important predictors of GHG emissions. These variables were established as important to consider to enhance the system. Therefore, apart from verifying the sufficient effectiveness of the established AI-based monitoring system, the analysis also offered a direction for the correct use of the system to support sustainable farming.

6. CONCLUSION

The purpose of this study was to identify how AI can contribute to the solution of the sustainable agriculture issue through the automation of GHG measurement. To provide accurate and timely information on emissions, the proposed system included IoT sensors, satellite images, and crop production records. The evaluation also reassured that the efficiency of the traditional models was lower than that of the AI models and that the Gradient Boosting was much more efficient than the other AI models as the models obtained high accuracy rates at reasonable resource consumption. The conclusion that has been deduced from this study demonstrates how vital it is to ensure that AI enhances agriculture and the environment. Statistical analysis also provided more evidence about the system's effectiveness: The higher the degree of correspondence between the parameters (for example, soil moisture and methane emissions), the larger the scale of the practical use.

The study also identified potential critical predictors that could be further used to implement targeted critical control measures like vegetation indices and fertilizer use. Furthermore, the system allowed for fairly objective monitoring through the application of highly reliable data dissemination channels that could be deployed in several settings in agriculture and disseminate the data collected. But for it to soar, problems have to be fixed data privacy, the fairness of the algorithm and others. For the improvement of the reliability of this approach, future studies should be devoted to the deployment of this method in real-time and to the environment with limited resources available, as well as the enhancement of interpretability. The system shows that technology and sustainability are connected and how AI can fight climate change and improve agriculture around the world.

References

Akhigbe, B. I., Munir, K., Akinade, O., Akanbi, L., & Oyedele, L. O. (2021). IoT Technologies for Livestock Management: A Review of Present Status, Opportunities, and Future Trends. *Big Data and Cognitive Computing*, *5*(1).

Bhatti, U. A., Bhatti, M. A., Tang, H., Syam, M. S., Awwad, E. M., Sharaf, M., & Ghadi, Y. Y. (2024). Global production patterns: Understanding the relationship between greenhouse gas emissions, agriculture greening and climate variability. *Environmental Research*, *245*, 118049. https://doi.org/https://doi.org/10.1016/j.envres.2023.118049

Butterbach-Bahl, K., Sander, B. O., Pelster, D., & Díaz-Pinés, E. (2016). Quantifying Greenhouse Gas Emissions from Managed and Natural Soils. In T. S. Rosenstock, M. C. Rufino, K. Butterbach-Bahl, L. Wollenberg, & M. Richards (Eds.), *Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture* (pp. 71-96). Springer International Publishing. https://doi.org/10.1007/978-3-319-29794-1_4

Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*. https://doi.org/10.5334/dsj-2015-002

Christiani, P., Rana, P., Räsänen, A., Pitkänen, T. P., & Tolvanen, A. (2024). Detecting Spatial Patterns of Peatland Greenhouse Gas Sinks and Sources with Geospatial Environmental and Remote Sensing

Data. *Environmental Management*, *74*(3), 461-478. https://doi.org/10.1007/s00267-024-01965-7

Farahzadi, L., & Kioumarsi, M. (2023). Application of machine learning initiatives and intelligent perspectives for CO2 emissions reduction in construction. *Journal of Cleaner Production*, *384*, 135504. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.135504

Gobezie, Y. M., & Boka, J. B. (2023). The nexus between greenhouse gas emissions and food security in sub-Saharan Africa: A system GMM analysis. *Cogent Economics & Finance*, *11*(2), 2273590. https://doi.org/10.1080/23322039.2023.2273590

Hu, Q., Li, J., Xie, H., Huang, Y., Canadell, J. G., Yuan, W., Wang, J., Zhang, W., Yu, L., Li, S., Lu, X., Li, T., & Qin, Z. (2024). Global methane emissions from rice paddies: CH4MOD model development and application. *iScience*, *27*(11), 111237. https://doi.org/https://doi.org/10.1016/j.isci.2024.111237

Hussin, F., Md Rahim, S. A. N., Hatta, N. S. M., Aroua, M. K., & Mazari, S. A. (2023). A systematic review of machine learning approaches in carbon capture applications. *Journal of CO2 Utilization*, *71*, 102474. https://doi.org/https://doi.org/10.1016/j.jcou.2023.102474

Lynch, J., Cain, M., Frame, D., & Pierrehumbert, R. (2021). Agriculture's Contribution to Climate Change and Role in Mitigation Is Distinct From Predominantly Fossil CO2-Emitting Sectors [Perspective]. *Frontiers in Sustainable Food Systems*, *4*. https://doi.org/10.3389/fsufs.2020.518039

Minamikawa, K., Tokida, T., Sudo, S., Padre, A., & Yagi, K. (2015). Guidelines for measuring CH4 and N2O emissions from rice paddies by a manually operated closed chamber method. *National Institute for Agro-Environmental Sciences, Tsukuba, Japan*, *76*.

Mulusew, A., & Hong, M. (2024). A dynamic linkage between greenhouse gas (GHG) emissions and agricultural productivity: evidence from Ethiopia. *Humanities and Social Sciences Communications*, *11*(1), 52. https://doi.org/10.1057/s41599-023-02437-9

Philibert, A., Loyce, C., & Makowski, D. (2013). Prediction of N2O emission from local information with Random Forest. *Environmental Pollution*, *177*, 156-163. https://doi.org/https://doi.org/10.1016/j.envpol.2013.02.019

Rajak, P., Ganguly, A., Adhikary, S., & Bhattacharya, S. (2023). Internet of Things and smart sensors in agriculture: Scopes and challenges. *Journal of Agriculture and Food Research*, *14*, 100776. https://doi.org/https://doi.org/10.1016/j.jafr.2023.100776

Romijn, E., De Sy, V., Herold, M., Böttcher, H., Roman-Cuesta, R. M., Fritz, S., Schepaschenko, D., Avitabile, V., Gaveau, D., Verchot, L., & Martius, C. (2018). Independent data for transparent monitoring of greenhouse gas emissions from the land use sector – What do stakeholders think and need? *Environmental Science & Policy*, *85*, 101-112. https://doi.org/https://doi.org/10.1016/j.envsci.2018.03.016

Zaman, M., Kleineidam, K., Bakken, L., Berendt, J., Bracken, C., Butterbach-Bahl, K., Cai, Z., Chang, S. X., Clough, T., Dawar, K., Ding, W. X., Dörsch, P., dos Reis Martins, M., Eckhardt, C., Fiedler, S., Frosch, T., Goopy, J., Görres, C. M., Gupta, A., . . . Müller, C. (2021). Methodology for Measuring Greenhouse Gas Emissions from Agricultural Soils Using Non-isotopic Techniques. In M. Zaman, L. Heng, & C. Müller (Eds.), *Measuring Emission of Agricultural Greenhouse Gases and Developing Mitigation Options using Nuclear and Related Techniques: Applications of Nuclear Techniques for GHGs* (pp. 11-108). Springer International Publishing. https://doi.org/10.1007/978-3-030-55396-8_2

Zhao, S., Liu, M., Tao, M., Zhou, W., Lu, X., Xiong, Y., Li, F., & Wang, Q. (2023). The role of satellite remote sensing in mitigating and adapting to global climate change. *Science of The Total Environment*, *904*, 166820. https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.166820