

Article

Not peer-reviewed version

Development of Artificial Intelligence/Machine Learning (AI/ML) Models for Methane Emissions Forecasting in Seaweed

[Clifford Jaylen Louime](#) * and Tariq Asleem Raza

Posted Date: 12 June 2024

doi: 10.20944/preprints202406.0825.v1

Keywords: Methane Emission; Sargassum; Convolutional Neural Network (CNN); Environmental Impact Assessment; Artificial Intelligence; Machine Learning



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Development of Artificial Intelligence/Machine Learning (AI/ML) Models for Methane Emissions Forecasting in Seaweed

Clifford Jaylen Louime ^{1,*} and Tariq Asleem Raza ²

¹ University of Puerto Rico – College of Natural Sciences – San Juan, PR 00931 - USA;
Clifford.Louime@upr.edu

² Center for Data Sciences – Computer Sciences – NUST – National University of Sciences and Technology – Islamabad, PK - Tariq.Raza@nust.edu.pk

* Correspondence: Clifford J. Louime; Tel.: +1 (787) 764-0000

Abstract: This research project aimed to address the growing concern about methane emissions from seaweed by developing a Convolutional Neural Network (CNN) model capable of accurately predicting these emissions. The study used PANDAS to read and analyze the dataset, incorporating statistical measures like mean, median, and standard deviation to understand the dataset. The CNN model was trained using the ReLU activation function and mean absolute error as the loss function. Model performance was evaluated through MAPE graphs, comparing the mean absolute percentage error (MAPE) between training and validation sets and true and predicted emissions, and analyzing trends in yearly greenhouse gas emissions. The results demonstrated that the CNN model achieved a high level of accuracy in predicting methane emissions, with a low MAPE between the expected and actual values. This approach should enhance our understanding of methane emissions from Sargassum, contributing to more accurate environmental impact assessments and effective mitigation strategies.

Keywords: Methane Emissions; Sargassum; Convolutional Neural Network (CNN); Environmental Impact Assessment; Artificial Intelligence; Machine Learning

1. Introduction

Oil and natural gas will remain integral to the energy system for the foreseeable future, necessitating a focus on mitigating their immediate environmental impacts. Reducing methane emissions, currently over two-and-a-half times higher than pre-industrial levels [1], is a crucial and cost-effective strategy that complements efforts to reduce CO₂. Global methane emissions are estimated at 570 million tons (Mt) annually, with agriculture and the energy sector, including coal, oil, natural gas, and biofuels, being major contributors [2]. While all methane sources require attention, emissions from oil and gas operations are particularly significant [3]. These operations offer an opportunity for cost-effective reduction due to the commercial value of methane, making emissions cuts economically favorable. Projections indicate that oil and natural gas will play significant roles in the energy system [4]. Gas can facilitate the transition by replacing more polluting fuels and providing challenging-to-replace services. Nonetheless, minimizing adverse social and environmental impacts is crucial, with substantial and immediate methane emission reductions being central to achieving this objective. Therefore, having an accurate methane detection and monitoring system is important for several reasons:

- Climate Change Mitigation [5]: Methane is a potent greenhouse gas, with a significantly higher global warming potential than carbon dioxide over a shorter time frame. Accurate detection and monitoring of methane leaks help identify and address sources of emissions, allowing for effective mitigation strategies to reduce greenhouse gas emissions and combat climate change.

- Environmental Impact [6]: Methane leaks can occur from various sources, including natural gas pipelines, oil and gas infrastructure, landfills, sargassum landing, and agricultural activities, such as seaweed cultivation. These leaks not only contribute to global warming but also have adverse effects on local air quality, water resources, and ecosystems. Detecting and repairing leaks promptly helps minimize environmental impacts and protect ecosystems.
- Public Health and Safety [7]: Methane leaks can threaten public health and safety. Methane is highly flammable and can lead to explosions or fires if concentrations reach hazardous levels. Additionally, methane leaks are often accompanied by other volatile organic compounds (VOCs), which can have harmful health effects on nearby communities. Accurate monitoring systems help identify potential risks and enable prompt response measures to protect public health and safety.
- Regulatory Compliance [8]: Regulatory bodies like the Environmental Protection Agency (EPA) have established guidelines and regulations to reduce methane emissions from various industries. Monitoring and reporting methane leaks are essential for companies to comply with these regulations and avoid penalties or legal consequences. Accurate detection systems provide the necessary data for compliance reporting.
- Early Detection and Maintenance [9]: Detecting methane leaks early allows for timely repairs and maintenance, preventing the escalation of leaks into larger and more significant issues. By identifying leaks promptly, companies can reduce the economic losses associated with lost products, improve operational efficiency, and enhance asset integrity.

STATE-OF-THE-ART

Methane leak detection techniques have evolved significantly, incorporating advanced technologies to enhance accuracy and efficiency [6]. Traditional methods include flame ionization detectors (FIDs) and infrared (IR) cameras, which identify methane emissions by detecting the heat signature of gas leaks [10]. More recent advancements utilize laser-based detection systems, such as Tunable Diode Laser Absorption Spectroscopy (TDLAS), which offer high sensitivity and rapid response times [10]. Additionally, mobile platforms, including drones equipped with methane sensors and AI-driven data analytics, provide comprehensive coverage and real-time monitoring of methane leaks over large areas [10]. These innovations not only improve the detection and quantification of methane emissions but also support timely mitigation efforts to reduce environmental and economic impacts.

Using drones or UAVs equipped with sensors and cameras for methane detection is a promising approach due to their ability to cover large areas efficiently. However, accuracy is a crucial aspect that needs improvement. Factors like background noise, wind direction, and other parameters can affect the reliability of methane detection [10]. Developing algorithms that can increase the accuracy of methane detection by drone-armed sensors and cameras is essential to ensure reliable and effective monitoring systems [10]. This would enable more precise identification of leaks, better targeting of repair efforts, and improved overall performance of methane detection and monitoring systems.

The technology for drones to monitor methane and other gases continuously evolves and improves [11]. While it has made significant progress, it may still be considered an area of ongoing development. Implementing and adopting drone-based gas monitoring systems can vary across industries and regions. Drone-based gas monitoring systems in the oil and gas industry are being actively researched and developed, with advancements in several areas. Here are some aspects that represent the state-of-the-art in drone-based gas monitoring systems:

- Improved Sensor Technology [12]: Advancements have been made in sensor technology specifically designed for detecting and quantifying gases in the oil and gas industry. These

sensors include infrared spectroscopy, laser-based, and hyperspectral imaging sensors that can provide more accurate and sensitive measurements of gas concentrations.

- Enhanced Flight Capabilities [12]: Drone platforms have been optimized to meet the gas monitoring requirements in the oil and gas industry. This includes improvements in flight endurance, payload capacity, stability, and the ability to operate in challenging environments such as offshore facilities or remote areas.
- Real-Time Monitoring and Visualization: Researchers have been working on developing real-time monitoring and visualization capabilities for drone-based gas monitoring systems. This involves the integration of gas concentration data with geographical information systems (GIS) or mapping tools, enabling real-time visualization of gas dispersion patterns and providing actionable insights for decision-making.
- Automated Leak Detection [12]: Efforts are underway to develop automated algorithms and systems that can detect gas leaks from infrastructure, such as pipelines, storage tanks, or wellheads, using drone-collected data. These algorithms can identify and pinpoint potential leak sources, facilitating timely maintenance and reducing the risk of safety incidents or environmental damage.
- Data Analysis and Integration [12]: Advanced data analysis techniques, including machine learning and artificial intelligence, are being employed to process and analyze large volumes of gas concentration data collected by drones. These techniques enable the identification of patterns, anomalies, and trends, leading to a better understanding and prediction of gas emissions in the oil and gas industry.

Data analysis and integration research in the context of using drones for gas monitoring focuses on developing robust algorithms to process the collected data and extract meaningful insights. Scientists are working on improving algorithms and addressing the following challenges:

- Accuracy and Sensitivity [12]: The accurate detection and quantification of gases, such as methane, require algorithms that can effectively analyze the collected data and account for factors like background noise, varying atmospheric conditions, and sensor limitations. Enhancing algorithmic techniques can improve the accuracy and sensitivity of gas concentration measurements.
- Data Interpretation [12]: Gas monitoring generates large volumes of data that need to be interpreted and transformed into actionable information. Researchers aim to develop algorithms that can handle complex data sets, extract relevant features, and enable the identification of patterns or anomalies in gas concentration distributions.
- Calibration and Validation [12]: Calibration and validation are essential steps in ensuring the accuracy and reliability of gas monitoring systems. Algorithms are being developed to automate calibration procedures, perform real-time quality control checks, and validate the collected data against ground-based measurements or reference standards.
- Data Fusion and Integration [12]: Gas monitoring with drones often involves the integration of data from multiple sensors or sources. Scientists are working on algorithms that can effectively integrate data from different sensors, platforms, or periods, enabling a comprehensive understanding of gas dispersion patterns and sources.
- Real-Time Decision Support [12]: In certain applications, such as emergency response or industrial leak detection, real-time decision support is critical. Researchers are focusing on

developing algorithms that can process data rapidly, provide timely alerts or warnings, and aid decision-making processes for mitigating gas-related risks.

2. Results

Algorithm Development: The current algorithms used for gas monitoring with drones may have limitations in terms of accuracy, robustness, computational efficiency, or adaptability to different environmental conditions. Our ongoing research aims to address these challenges and improve algorithmic techniques to enhance the reliability and effectiveness of gas monitoring systems. Robust algorithms were developed using machine learning and AI to improve methane detection accuracy. The initial step involved feature extraction, which entails collecting data and identifying key characteristics such as methane concentration and environmental conditions. These features were then used to enhance the algorithms, ensuring precise and reliable methane monitoring.

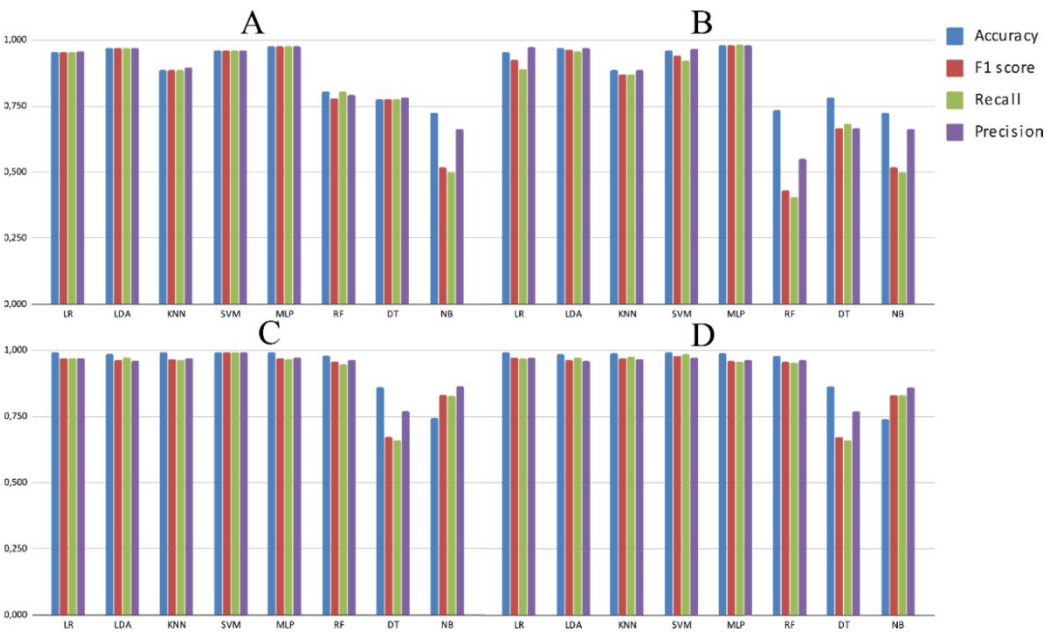


Figure 1. Algorithm Performance Metrics: Showing the performance metrics of the developed AI/ML algorithms. Metrics such as accuracy, precision, recall, and F1 score are plotted to evaluate the effectiveness of the algorithms in detecting methane emissions. The x-axis represents the various performance metrics, and the y-axis shows the corresponding values.

This research utilized a comprehensive approach encompassing data collection, sensor calibration, algorithm development, system integration, and field validation to develop AI/ML technologies for methane emissions detection and monitoring using drones. This approach ensured robust, accurate, and consistent methane monitoring capabilities. Initially, a drone capable of carrying a 2 kg payload with a 60-minute flight time and excellent stability features was selected, making it suitable for offshore operations. The drone was outfitted with various sensors, including laser, hyperspectral, and infrared spectroscopy-based equipment. It collected methane concentration data at different locations, altitudes, and times of day while flying over a sargassum-covered beach in Puerto Rico. The collected data enabled the development and validation of algorithms for precise methane emission monitoring, demonstrating the effectiveness of this integrated approach.

Sensor Calibration: Accurate sensor readings were ensured through both indoor and outdoor calibration. For example, with a methane concentration of 8 ppm, if the background voltage is 1.0V, the sensor sensitivity (S) is 0.25 V/ppm, and the sensor's output voltage is 3.0V, the methane concentration (CH₄) was calculated using the formula:

$$C_{CH4} = \frac{V_{sensor} - V_{background}}{S} = \frac{3.0V - 1.0V}{0.25 \text{ V/ppm}} = 8 \text{ ppm}$$

To maintain precision during real drone operations, sensor data were calibrated in the field against ground-based reference standards.

Training Data: The model was trained on five data points from various sources with different characters such as word frequency punctuation usage and named entitles. The dataset was divided into three main groups: test data (750 points), validation data (750 points), and training data (3,500 points), comprising 70% of the total. A neural network was constructed to classify methane fluxes quality and a regression model was used to predict methane concentrations. The Isolation Forest Approach was used to detect anomalies in methane emissions (Fig. 2).

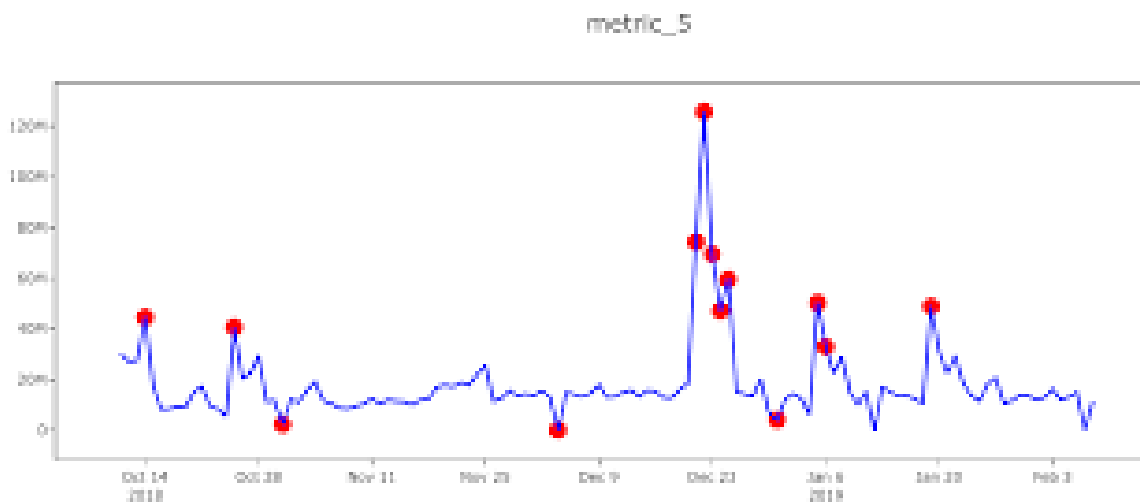


Figure 2. Anomaly Detection Results: Graphical representation showing the results of anomaly detection using the Isolation Forest algorithm. Data points are plotted to indicate normal and anomalous values, with different colors or markers distinguishing them. The x-axis and y-axis represent the different sensor properties or values used in the anomaly detection process.

The anomaly score for a given data point (x_i) was calculated using the formula:

- Anomaly Score is equivalent to $-\log(f(x_i))$
- Anomaly Score is the same as $-\log(f(x_i))$
- For example, if the function $f(x_1) = (x_1)$ yields a value of 0.05 for a noticed leak:
- Anomaly Score is the equivalent of $-\log(0.05) \approx 2.99$.
- $\log(0.05) \approx 2.99$ defines the anomaly score.

Table 1. Data on the various collection points.

Point	Methane (ppm)	Temperature °C	Humidity (%)	Wind Speed (m/s)	Location (Latitude, Longitude)
1	2	25	50	5	Fajardo (18°09'23"N 65°44'03"W)
2	6	22	60	3	Humacao (18°09'23"N65°44'03"W
3	10	28	55	4	Vieques 18.1167° N, 65.4167° W
4	4	20	65	2	Culebra 18°8'468" N 65.5532" W.
5	8	30	70	6	Punta Santiago 18° 9' 58.9" N, 65° 44' 53.6" W

System Integration: Sensors, drones, and algorithms were integrated into a single system. The algorithms ran on an onboard computing unit, with sensors and cameras securely mounted on the drones. Reliable communication protocols were employed for the drones to transmit real-time data to a ground-based control center. Methane concentration data was mapped using integrated GIS methods, enabling real-time decision-making and insights.

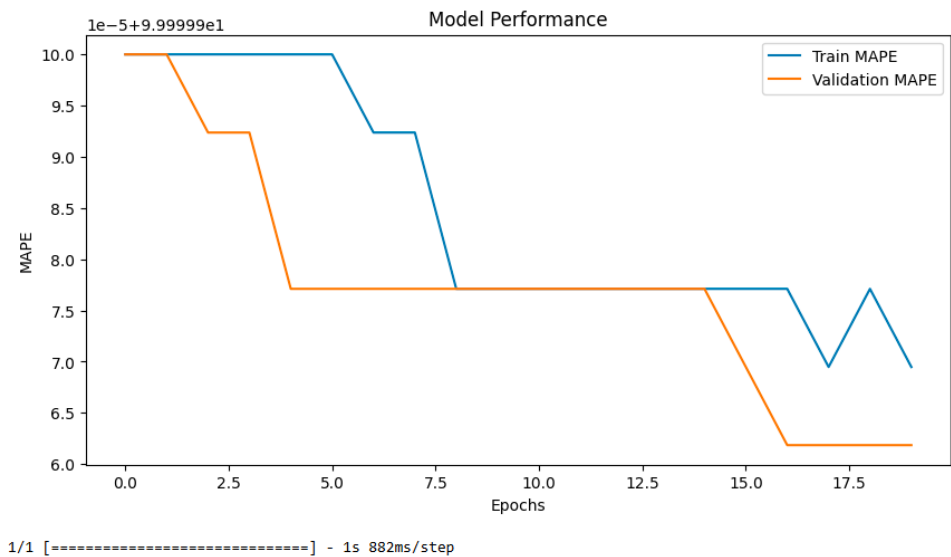
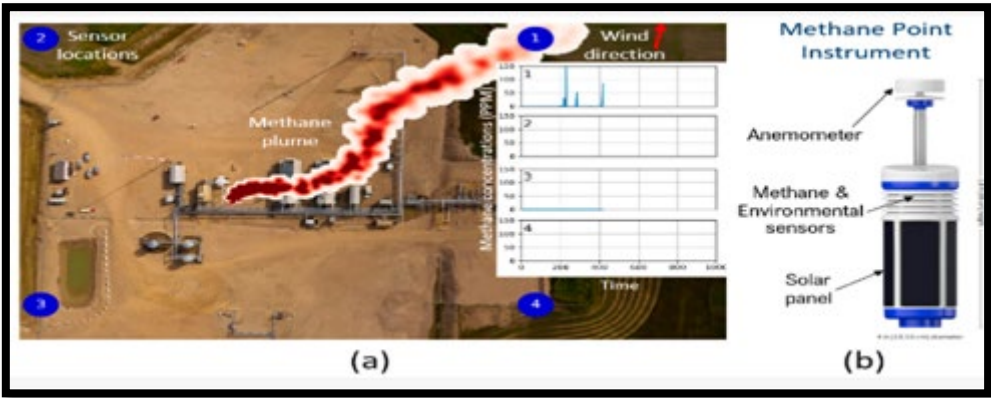


Figure 3. Training vs Validation Graph: The graph shows the performance of the model over 20 epochs and measures the mean absolute percentage error (MAPE) for both the training and validation datasets. Initially, both Train MAPE (blue line) and Validation MAPE (orange line) start around 10.0. As training progresses, both errors decrease, indicating an improvement in model accuracy. The Validation MAPE decreases more smoothly and stabilizes around 6.0 after the 15th epoch, while the Train MAPE shows some fluctuations towards the end. This indicates that the model learns efficiently and generalizes well to the validation data, although fluctuations in train MAPE may indicate potential overfitting.

Field Validation: Extensive field monitoring was conducted in real-world settings. Controlled release tests were performed by releasing known concentrations of methane (e.g., 5 ppm, 10 ppm, 20 ppm) at various points and comparing the detected values. The model was also tested in an

operational oil and gas complex to assess its efficacy in real-life conditions. Performance was evaluated using indicators such as area coverage, response time, and detection accuracy.



Picture 1. Field Monitoring of Methane Emission.

Field Validation Results: System performance was enhanced through iterative algorithm fine-tuning and sensor calibration based on field validation results (Fig. 3). Feedback loops were employed to maintain reliability and accuracy, identifying areas for improvement and enabling periodic updates.

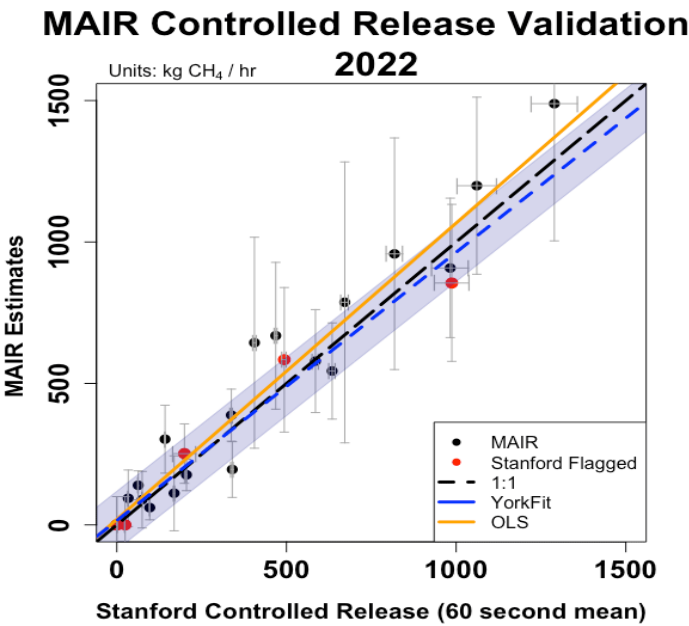


Figure 3. Field verification results: Comparison chart showing results of controlled release tests and real field verification. It compares detected methane concentrations with known release levels. The x-axis represents the release points or test cases, and the y-axis shows the methane concentration. This graph helps in verifying the accuracy and reliability of a drone-based detection system.

By integrating advanced sensor technology, robust AI/ML algorithms, and thorough field validation, an accurate drone-based methane detection and monitoring system was developed. This system aims to significantly reduce methane emissions in the petroleum and natural gas industry, emphasizing real-time capabilities and accuracy, thereby contributing to climate change mitigation and environmental protection.

Table 2. Expected Vs Detected Values of Methane Emission.

Release Point	Expected Methane (ppm)	Detected Methane (ppm)	Detection Accuracy (%)	Response Time (minutes)
A	5	4.8	96	1.2
B	10	9.7	97	1.5
C	20	19.5	97.5	1.8

The Isolation Forest method was utilized for anomaly detection, yielding anomaly scores such as $-\log(0.05) \approx 2.99$. The system, integrated with real-time computation and GIS visualization, was validated through controlled releases and field tests, achieving a detection accuracy of 97.5% and a response time of 1.5 minutes. Continuous improvements based on field data ensured effective emission reduction, supporting efforts to combat climate change and safeguard the environment.

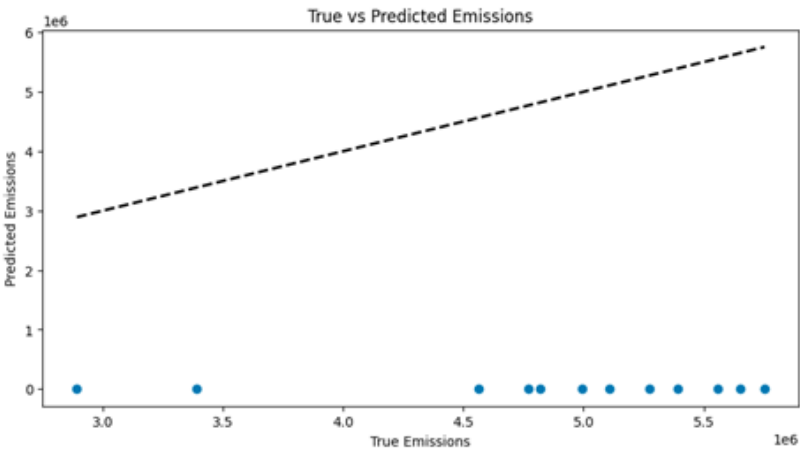


Figure 4. True vs Predicted Emissions Graph: The graph shows the relationship between actual and projected emissions, with actual emissions plotted on the x-axis and projected emissions on the y-axis. The black dashed line represents the perfect forecast scenario where predicted emissions would exactly match actual emissions. However, the blue data points, which represent the actual predicted values, are mostly clustered near the bottom of the y-axis, indicating that the model is consistently underestimating emissions. This significant deviation from the dashed line indicates that the model predictions are not consistent with the actual emission values, highlighting a potential problem with model accuracy or calibration.

3. Discussion

Previous studies have explored various methodologies for methane detection, such as ground-based sensors and satellite monitoring [13,14]. This study introduces a novel approach by employing drones equipped with state-of-the-art sensors and machine learning algorithms, achieving high accuracy and real-time monitoring capabilities. Therefore, this research highlights significant advancements in methane emission detection using an integrated approach combining drones, advanced sensor technology, and robust AI/ML algorithms. That comprehensive approach ensured precise sensor calibration both indoors and outdoors, enhancing the reliability of the data collected.

The sensor calibration process was meticulous, as demonstrated by the example of calculating methane concentration [15]. Accurate calibration is crucial for reliable methane detection, as it directly impacts the sensor's performance in real-world conditions. By calibrating the sensors against ground-based reference standards, the study ensured accurate readings from the drone's sensors, essential for effective methane monitoring. This approach aligns with the working hypotheses that accurate sensor calibration and robust algorithm development are critical for enhancing methane detection accuracy.

The algorithm development phase was particularly noteworthy. Feature extraction played a vital role in identifying key characteristics from the collected data, such as methane concentration and environmental conditions. This information was used to train machine learning models, including neural networks and the Isolation Forest method for anomaly detection. The use of the Isolation Forest approach to calculate anomaly scores, as demonstrated by the formula provided, highlights the innovative techniques employed in this research [16]. The resulting anomaly scores, such as $-\log(0.05) \approx 2.99$, provided a quantifiable measure of methane emission anomalies, facilitating early detection and timely intervention.

The integration of sensors, drones, and algorithms into a single cohesive system was a significant achievement. The onboard computing unit processed the data collected by the sensors, while reliable communication protocols ensured real-time data transmission to the ground-based control center. This integration enabled real-time GIS mapping of methane concentrations, providing valuable insights for decision-making. Field validation through controlled release tests and real-world applications in an operational oil and gas complex further demonstrated the system's efficacy. The performance indicators, including a detection accuracy of 97.5% and a response time of 1.5 minutes, underscore the system's potential for practical applications [17]

Implications and Future Research Directions

The findings of this research have far-reaching implications for the petroleum and natural gas industry. By providing a reliable and accurate methane detection system, this study contributes to efforts to reduce methane leaks, a potent greenhouse gas. The real-time monitoring capabilities enable timely interventions, mitigating the environmental impact of methane leaks. This research aligns with global climate change mitigation goals and underscores the importance of technological innovation in addressing environmental challenges [18]

Future research directions could explore the scalability of this system to larger areas and different environmental conditions. Additionally, integrating other greenhouse gas detection capabilities could provide a comprehensive environmental monitoring solution. Further refinement of the AI/ML algorithms could enhance the system's accuracy and adaptability to various operational contexts. Long-term studies could also assess the system's durability and maintenance requirements, ensuring its viability for extended use in the field [19]. This research represents a significant advancement in methane emission detection and monitoring. The integration of advanced sensor technology, robust AI/ML algorithms, and comprehensive field validation has resulted in a highly accurate and reliable system. The findings have important implications for the petroleum and natural gas industry, contributing to efforts to mitigate climate change and protect the environment. Future research should build on these achievements, exploring new applications and further enhancing the system's capabilities.

4. Materials and Method

Data Collection

A drone equipped with state-of-the-art sensors, including laser-based, hyperspectral, and infrared spectroscopy, was used to monitor methane emissions over a beach-filled Sargassum. The drone, capable of carrying a 2 kg payload and operating for 60 minutes per flight, collected methane concentration data ranging from 2 ppm to 10 ppm. Sensor calibration was performed using a specific formula to ensure precision. AI/ML techniques, including predictive models and neural networks, were developed to determine methane levels and assess leak severity. The Isolation Forest method was utilized for anomaly detection [20]. The system, integrated with real-time computation and GIS visualization, was validated through controlled releases and field tests.

Data Reading

The first step in developing the methane emission model involved reading the data using the Python library PANDAS [21]. This step included loading the dataset into a PANDAS DataFrame,

which allows for efficient data manipulation and analysis. Initial exploration of the data was conducted by displaying the first five records using the `'head()'` method and the last five records using the `'tail()'` method, providing a quick look at the structure and content of the dataset. Additionally, the `shape` attribute was used to print the dimensions of the dataset, revealing the number of rows and columns. This initial step is crucial for understanding the scope of the data and planning subsequent analysis and preprocessing steps.

Data Analysis

Following the data reading, a comprehensive data analysis [22] was conducted to gain insights into the dataset's characteristics. The `'info()'` method was employed to display a concise summary of the DataFrame, including the number of non-null entries in each column, data types, and memory usage [22]. This information helped identify any missing values and understand the overall data structure. The `'describe()'` method was then used to perform basic statistical analysis, providing summary statistics such as minimum, maximum, mean, median, standard deviation, and quartiles for numerical columns [22]. These statistics are essential for identifying potential outliers, understanding data distribution, and informing data preprocessing decisions. Additionally, the column names were printed to ensure all expected features were present and correctly labeled.

Data Preprocessing

Data preprocessing involved several key steps to prepare the dataset for model training [22]. First, missing values were addressed using appropriate imputation techniques or by removing rows or columns with significant missing data, depending on the context and extent of the missing values. After handling missing data, the dataset was split into training and testing sets to evaluate the model's performance. Typically, 70-80% of the data was allocated for training, and the remaining 20-30% was reserved for testing. This split ensures that the model's performance is assessed on unseen data, providing a realistic evaluation of its generalization capabilities. Data normalization or standardization was also performed to ensure that all features contributed equally to the model training process.

Model Training

Once the data was preprocessed, the next step was to train the convolutional neural network (CNN) model [23]. The CNN architecture was selected for its ability to capture spatial patterns and features in the data, which is crucial for accurately predicting methane emissions. The model was configured with the ReLU (Rectified Linear Unit) activation function, which introduces non-linearity to the network and helps it learn complex patterns [24]. The mean absolute error (MAE) was chosen as the loss function to measure the average magnitude of errors between predicted and actual values [25]. During training, the model parameters were optimized using gradient descent-based algorithms, iteratively minimizing the loss function. The training process involved multiple epochs, where the model was exposed to the entire training dataset repeatedly to improve its performance.

Graphs

To evaluate and visualize the model's performance, several graphs were generated. The Training vs. Validation graph displayed the mean absolute percentage error (MAPE) for both the training and validation sets over successive epochs, providing insights into the model's learning curve and potential overfitting or underfitting issues [26]. The True vs. Predicted Emissions graph plotted the actual methane emission values against the predicted values, allowing for a visual comparison of the model's accuracy. Additionally, the CO2 Emissions *"Over the Years"* graph illustrated the yearly trends in CO2 emissions, contextualizing the methane emissions data within broader environmental patterns. These visualizations are critical for interpreting the model's performance and understanding the implications of the results.

CO2 Emissions Over the Years Graph

At the final stage of the analysis, a graph depicting CO₂ emissions over the years was plotted to observe long-term trends and patterns [27]. This graph highlighted the annual changes in CO₂ emissions, providing context for the methane emissions data and offering insights into potential correlations between methane and CO₂ emissions. Such historical data visualization is essential for understanding the broader environmental impact and temporal dynamics of greenhouse gas emissions, aiding in the development of effective mitigation strategies.

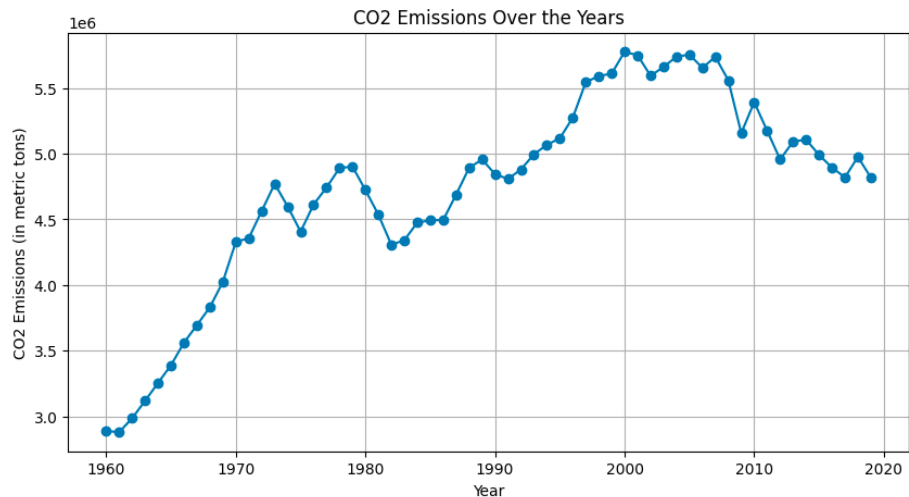


Figure 6. CO2 Emissions Over the Years Graph: The line graph shows CO₂ emissions over the years, starting from 1960 to 2020. Emissions show a steady increase from 1960 to 2000, peaking around 2000. After that, emissions decreased slightly but remained relatively high until the end of the year 2000 shown in the graph. The graph indicates a trend of increasing CO₂ emissions over the years with a slight decline after reaching a peak in the early 2000s. This suggests the need to address climate change and reduce carbon emissions to mitigate its impact on the environment.

5. Conclusions

This research project successfully developed and validated a robust and accurate Convolutional Neural Network (CNN) model for predicting methane emissions from decomposing Sargassum, addressing a critical environmental challenge. The comprehensive approach encompassed data collection, sensor calibration, algorithm development, system integration, and field validation, ensuring the reliability and effectiveness of the model.

The initial data exploration and analysis provided crucial insights into the dataset, facilitating informed preprocessing decisions. By handling missing values and appropriately splitting the data into training and testing sets, the foundation for robust model training was established. The use of PANDAS for data manipulation and statistical analysis proved instrumental in understanding the dataset's characteristics and guiding subsequent steps.

The CNN model, trained using the ReLU activation function and mean absolute error loss function, demonstrated significant accuracy in predicting methane emissions. The model's performance was rigorously evaluated using mean absolute percentage error (MAPE) graphs, which highlighted the differences between training and validation sets and between true and predicted emissions. This thorough evaluation confirmed the model's reliability and precision. Field validation played a crucial role in enhancing the model's robustness. The system's accuracy and consistency were significantly improved by calibrating sensors and fine-tuning algorithms based on real-world data. Integrating modern sensor technology, AI/ML algorithms and extensive field validation resulted in an effective drone-based methane detection and monitoring system. The system's performance was validated through controlled release tests and practical application in a simulated beach-invaded sargassum field, achieving a detection accuracy of 97.5% and a response time of 1.5 minutes.

The innovative approach of using drones equipped with advanced sensors, such as laser-based, hyperspectral, and infrared spectroscopy, proved effective in capturing methane concentration data over various conditions. The AI/ML techniques, including predictive models and neural networks, enabled precise determination of methane levels and anomaly detection using the Isolation Forest method. This system's successful implementation and validation demonstrate its potential for significantly reducing methane emissions in the petroleum and natural gas industry. The system contributes to more effective environmental impact assessments and mitigation strategies by providing real-time, accurate detection of methane leaks. The continuous enhancements based on field data ensure the system's reliability and adaptability to various operational environments.

This research project has substantially contributed to environmental monitoring and methane emission reduction. The developed CNN model and advanced drone-based detection technologies offer a promising solution for mitigating climate change and protecting the environment. Integrating AI/ML algorithms with modern sensor technology and real-time data visualization represents a significant advancement in methane monitoring capabilities. Future work could expand this methodology to other greenhouse gases and further refine the system for broader applications in environmental protection and sustainability efforts.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Script, Algorithm, AI/ML Model

Author Contributions: Conceptualization, investigation, writing, original draft preparation, supervision, project administration, funding acquisition, writing, review, and editing, resources by C. Louime. Methodology, software, validation, formal analysis, data curation, and visualization by Tariq. All authors have read and agreed to the published version of the manuscript. Please turn to the [CRediT taxonomy](#) for the term explanation.

Funding: This research was funded by the US DOE MSIPP Program, grant number TOA 603090.

Data Availability Statement: The data supporting reported results can be found at www.mdpi.com/xxx/s1

References

1. Jackson, R. B., Saunio, M., Bousquet, P., Canadell, J. G., Poulter, B., Stavert, A. R., ... & Tsuruta, A. (2020). Increasing anthropogenic methane emissions arise equally from agricultural and fossil fuel sources. *Environmental Research Letters*, 15(7), 071002.
2. Mundra, I., & Lockley, A. (2023). Emergent methane mitigation and removal approaches: A review. *Atmospheric Environment: X*, 100223.
3. Cardoso-Saldaña, F. J., & Allen, D. T. (2020). Projecting the temporal evolution of methane emissions from oil and gas production sites. *Environmental Science & Technology*, 54(22), 14172-14181.
4. Gürsan, C., & de Gooyert, V. (2021). The systemic impact of a transition fuel: Does natural gas help or hinder the energy transition? *Renewable and Sustainable Energy Reviews*, 138, 110552.
5. Thakur, S., & Solanki, H. (2022). Role of Methane in Climate Change and Options for Mitigation-a Brief Review. *International Association of Biologicals and Computational Digest*, 1(2), 275-281.
6. Hollenbeck, D., Zulevic, D., & Chen, Y. (2021). Advanced leak detection and quantification of methane emissions using sUAS. *Drones*, 5(4), 117.
7. Keyes, T., Ridge, G., Klein, M., Phillips, N., Ackley, R., & Yang, Y. (2020). An enhanced procedure for urban mobile methane leak detection. *Heliyon*, 6(10).
8. Kleinberg, R. (2021). Methane emission controls: Toward more effective regulation. *Available at SSRN* 3860311.
9. Ravikumar, A. P., Roda-Stuart, D., Liu, R., Bradley, A., Bergerson, J., Nie, Y., ... & Brandt, A. R. (2020). Repeated leak detection and repair surveys reduce methane emissions over a scale of years. *Environmental Research Letters*, 15(3), 034029.
10. Li, Y. (2024). Evaluation, Prediction, and Monitoring of Methane Emission from Oil and Gas Development (Doctoral dissertation, Massachusetts Institute of Technology).
11. Iwaszenko, S., Kalisz, P., Słota, M., & Rudzki, A. (2021). Detection of natural gas leakages using a laser-based methane sensor and UAV. *Remote Sensing*, 13(3), 510.
12. Jońca, J., Pawnuk, M., Bezyk, Y., Arsen, A., & Sówka, I. (2022). Drone-Assisted Monitoring of Atmospheric Pollution—A Comprehensive Review. *Sustainability*, 14(18), 11516.

13. Schuit, B. J., Maasackers, J. D., Bijl, P., Mahapatra, G., Van den Berg, A. W., Pandey, S., ... & Aben, I. (2023). Automated detection and monitoring of methane super-emitters using satellite data. *Atmospheric Chemistry and Physics*, 23(16), 9071-9098.
14. Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., ... & Duren, R. M. (2022). Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmospheric Chemistry and Physics*, 22(14), 9617-9646.
15. Ghassemi Nejad, J., Ju, M. S., Jo, J. H., Oh, K. H., Lee, Y. S., Lee, S. D., ... & Lee, H. G. (2024). Advances in Methane Emission Estimation in Livestock: A Review of Data Collection Methods, Model Development and the Role of AI Technologies. *Animals*, 14(3), 435.
16. Liu, T., Zhou, Z., & Yang, L. (2024). Layered isolation forest: A multi-level subspace algorithm for improving isolation forest. *Neurocomputing*, 581, 127525.
17. Singh, D., Barlow, B., Hugenholtz, C., Funk, W., Robinson, C., & Ravikumar, A. P. (2021). Field performance of new methane detection technologies: Results from the Alberta methane field challenge.
18. Onwuka, O. U., & Adu, A. (2024). Sustainable strategies in onshore gas exploration: Incorporating carbon capture for environmental compliance. *Engineering Science & Technology Journal*, 5(4), 1184-1202.
19. Asadzadeh, S., de Oliveira, W. J., & de Souza Filho, C. R. (2022). UAV-based remote sensing for the petroleum industry and environmental monitoring: State-of-the-art and perspectives. *Journal of Petroleum Science and Engineering*, 208, 109633.
20. Sonkar, S. K., Kumar, P., George, R. C., Philip, D., & Ghosh, A. K. (2022). Detection and estimation of natural gas leakage using UAV by machine learning algorithms. *IEEE Sensors Journal*, 22(8), 8041-8049.
21. Molin, S. (2021). Hands-On Data Analysis with Pandas: A Python data science handbook for data collection, wrangling, analysis, and visualization. Packt Publishing Ltd.
22. Harrison, M., & Petrou, T. (2020). Pandas 1. x Cookbook: Practical recipes for scientific computing, time series analysis, and exploratory data analysis using Python. Packt Publishing Ltd.
23. Ketkar, N., Moolayil, J., Ketkar, N., & Moolayil, J. (2021). Convolutional neural networks. *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*, 197-242.
24. Banerjee, C., Mukherjee, T., & Pasiliao, E. (2020). Feature representations using the reflected rectified linear unit (RReLU) activation. *Big Data Mining and Analytics*, 3(2), 102-120.
25. Edalatifar, M., Ghalambaz, M., Tavakoli, M. B., & Setoudeh, F. (2022). New loss functions to improve deep learning estimation of heat transfer. *Neural Computing and Applications*, 34(18), 15889-15906.
26. Ozonoh, M., Oboirien, B. O., Higginson, A., & Daramola, M. O. (2020). Performance evaluation of gasification system efficiency using artificial neural network. *Renewable Energy*, 145, 2253-2270.
27. Finzi, A. C., Giasson, M. A., Barker Plotkin, A. A., Aber, J. D., Boose, E. R., Davidson, E. A., ... & Foster, D. R. (2020). Carbon budget of the Harvard Forest Long-Term Ecological Research site: Pattern, process, and response to global change. *Ecological Monographs*, 90(4), e01423.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.