atmosphere

MDPI

Article

Addressing Low-Cost Methane Sensor Calibration Shortcomings with Machine Learning

Elijah Kiplimo *, Stuart N. Riddick *, Mercy Mbua, Aashish Upreti, Abhinav Anand and Daniel J. Zimmerle

The Methane Emission Technology Evaluation Center (METEC), Energy Institute, Colorado State University, Fort Collins, CO 80524, USA; mercy.mbua@colostate.edu (M.M.); aashish.upreti@colostate.edu (A.U.); a.anand@colostate.edu (A.A.); dan.zimmerle@colostate.edu (D.J.Z.)

* Correspondence: elijah.kiplimo@colostate.edu (E.K.); stuart.riddick@colostate.edu (S.N.R.)

Abstract: Quantifying methane emissions is essential for meeting near-term climate goals and is typically carried out using methane concentrations measured downwind of the source. One major source of methane that is important to observe and promptly remediate is fugitive emissions from oil and gas production sites but installing methane sensors at the thousands of sites within a production basin is expensive. In recent years, relatively inexpensive metal oxide sensors have been used to measure methane concentrations at production sites. Current methods used to calibrate metal oxide sensors have been shown to have significant shortcomings, resulting in limited confidence in methane concentrations generated by these sensors. To address this, we investigate using machine learning (ML) to generate a model that converts metal oxide sensor output to methane mixing ratios. To generate test data, two metal oxide sensors, TGS2600 and TGS2611, were collocated with a trace methane analyzer downwind of controlled methane releases. Over the duration of the measurements, the trace gas analyzer's average methane mixing ratio was 2.40 ppm with a maximum of 147.6 ppm. The average calculated methane mixing ratios for the TGS2600 and TGS2611 using the ML algorithm were 2.42 ppm and 2.40 ppm, with maximum values of 117.5 ppm and 106.3 ppm, respectively. A comparison of histograms generated using the analyzer and metal oxide sensors mixing ratios shows overlap coefficients of 0.95 and 0.94 for the TGS2600 and TGS2611, respectively. Overall, our results showed there was a good agreement between the ML-derived metal oxide sensors' mixing ratios and those generated using the more accurate trace gas analyzer. This suggests that the response of lower-cost sensors calibrated using ML could be used to generate mixing ratios with precision and accuracy comparable to higher priced trace methane analyzers. This would improve confidence in low-cost sensors' response, reduce the cost of sensor deployment, and allow for timely and accurate tracking of methane emissions.

Keywords: methane; quantification; metal-oxide; sensor; calibration; machine learning

check for updates

Citation: Kiplimo, E.; Riddick, S.N.; Mbua, M.; Upreti, A.; Anand, A.; Zimmerle, D.J. Addressing Low-Cost Methane Sensor Calibration Shortcomings with Machine Learning. *Atmosphere* **2024**, *15*, 1313. https://doi.org/10.3390/atmos15111313

Academic Editor: Esther Hontañón

Received: 23 August 2024
Revised: 23 October 2024
Accepted: 29 October 2024
Published: 31 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Methane (CH_4), a gas with a global warming potential 25 times greater than carbon dioxide, has increased in atmospheric composition by a factor of three since preindustrial times, largely due to changes in emissions from anthropogenic activities [1,2]. Nearly all countries globally have entered treaties or agreements to reduce methane emissions but it has recently been suggested that methods used to quantify methane emissions are insufficiently precise to measure emission reduction of mitigation strategies [3–6].

Measuring methane emissions from oil and gas production activities has been the focus of many studies since the 1990s [7–12]. Recently, approaches have developed from survey methods, where short-duration measurements are made by instruments mounted on a mobile platform [13–15], to continuous monitoring approaches, where measurements are made by instrumentation fixed in location [6,16–18]. The advantage of continuous monitors

is that observations are more likely to detect the short-duration, large-emission events that are typical of oil and gas emission distributions and are essential to capture if realistic emissions estimates are to be generated [19,20]. Typically, on-site continuous monitoring systems comprise several (a minimum of four) methane measurement/meteorological instrumentation installed at fixed points around the site coupled with a dispersion model to infer a rate of emission [6,21,22]. One difficulty in designing a continuous monitoring system that is to be installed at the fence line of an oil and gas production site is the choice of methane measurement instrumentation. Methane sensing technologies include metal oxide sensors (USD 15 per sensor), integrated infrared sensors (USD 300 per sensor), tunable diode laser absorption spectrometers (USD 1000 per sensor), and optical cavity instrumentation (>USD 10,000 per sensor). As there are thousands of production sites across the US alone, costly (>USD 1000) methane measurement systems are prohibitively expensive to deploy at every site; therefore, metal oxide sensors are currently being widely deployed across oil and gas production sites across the US [6,22].

As metal oxide sensors are being widely distributed to measure near-background methane concentrations in air (typically below 100 ppm), many studies have described different methods of calibration, deployment, and response. However, most studies report that these sensors are sensitive to temperature and relative humidity and require individual calibration [17,23–30]. Typical metal oxide sensor models used in these studies are the Taguchi Gas Sensors (TGS) 2600 and 2611 models produced by Figaro Engineering Inc. (Osaka, Japan). TGS2600 uses a tin dioxide (SnO_2) sensing layer that reacts with the detected gases. The detection mechanism relies on the change in electrical conductivity through the SnO_2 in the presence of methane, which is a reducing gas. Absorbance of oxygen molecules occurs on the surface of SnO_2, forming O^{2-} ions leading to electron capture that creates a depletion layer at the surface of the SnO_2, increasing its resistance [23,24]. Methane interacts with the adsorbed oxygen ions, releasing the trapped electrons back into the SnO_2 conduction band, decreasing its resistance. The change in resistance is proportional to the concentration of CH_4. The TGS2611 has the same operation basis as the TGS2600 but has been optimized for methane detection by the integration of a heating element to maintain optimal temperature for methane detection and features a filter material that selectively permits methane to reach the sensing element while blocking other gases [23,24].

Studies using Figaro TGS sensors have measured the resistance change across the metal oxide strip to infer changes in concentration. It is assumed the metal oxide sensor has a relatively constant resistance in clean air (R_0, Ω), which becomes lower in the presence of methane (R_s, Ω), and the ratio of these resistances gives a measure of the methane mixing ratio in air. The resistance of the metal oxide sensor is also affected by the air temperature (T_a, °C) and relative humidity (rH, %), and the ratio of resistance can be corrected to account for these factors using an empirically derived algorithm (Equation (1)) [25].

$$\left(\frac{R_s}{R_0}\right)_{corr} = \frac{R_s}{R_0} \times (0.024 + 0.0072 \times rH + 0.0246 \times T_a) \quad (1)$$

The temperature and humidity corrected ratio of R_s and R_0 is then typically converted to a methane mixing ratio using a sensor-specific calibration algorithm [25,30]. As they perform badly in low relative humidity, less than 40% RH [30], the TGS sensors are typically calibrated by comparing $(R_s/R_0)_{corr}$ to methane mixing ratios measured using a sub-ppm reference instrument measuring methane contemporaneously. An algorithm is then generated to translate $(R_s/R_0)_{corr}$ to a calibrated mixing ratio ($[CH_4]_{cal}$). Published algorithms have been linear, power functions, and exponential functions [17,23,25,26,30,31], and there is not a good current understanding of differences between individual sensors' responses to changing methane concentrations apart from possible differences in manufacturing [17].

Nested within the conversion from measured data to a calibrated methane mixing ratio are several potential sources of uncertainty: 1. The size of the measured resistance (R_s) appears to be highly variable between individual sensors, and through trial and error, we have found that the output of some of the sensors in near background mixing ratios

(<10 ppm)—the methane signal—has become almost indistinguishable from the noise; 2. The calculation of $[CH_4]_{cal}$ is heavily dependent on determining a value for the TGS resistance in clean air, R_0; 3. Given the variability in response of sensors to changes in methane concentration, it seems questionable that all sensors would respond equally to changes in relative humidity and temperature (as defined in Equation (1)), especially in sensors with low signal-to-noise ratios. Overall, these uncertainties raise concerns about the historical methods for converting signals from TGS methane sensors to a calibrated methane mixing ratio.

One alternative approach could be to use a novel machine learning approach that accounts for low signal-to-noise ratios and differences in sensor response to relative humidity and temperature changes. Recent developments in machine learning algorithms have transformed sensor calibration and enhanced their reliability in environmental monitoring [32]. One example of machine learning is the random forest (RF) regressor developed by Leo Breiman in 2001. Random forest is an ensemble machine learning model [33], which implies that the model is built by the combination of predictions from multiple simpler models [33]. Random forest uses a collection of decision trees, which are the models that split the data into branches based on the input features to make a prediction. Training the RF regression model involves presenting the model with the input variables (metal oxide sensors' resistances, temperature, and relative humidity) that are passed through multiple decision trees, and the output aggregation predicts the target variable (in this case the trace gas analyzer CH_4 concentrations).

To investigate this, we will use two TGS sensors with low signal-to-noise ratios and report on the RF machine learning approach to generating calibrated methane mixing ratios. Specifically, we will 1. collect R_s values from low-response TGS2600 and TGS2611 sensors contemporaneously with a sub-ppm methane analyzer; 2. generate calibration curves using (a) the traditional method (presented above) and (b) the RF machine learning approach; and 3. compare the calibrated methane mixing ratios reported by the TGS2600 and TGS2611 to the methane analyzer to generate an understanding of sensor type response, bias, and drift over time. Ultimately, we aim to provide evidence to understand if the RF machine learning approach can generate better calibration algorithms than currently used methods.

2. Materials and Methods

2.1. Methane Measurement

2.1.1. Figaro TGS Sensors

Both the TGS 2600 and 2611 sensors were connected to a Raspberry Pi 4 (London, UK) via an ADS1115 16 bit analog to digital converter, with an inclusion of a DHT22 temperature and humidity sensor. The Raspberry Pi was powered by a 5 V 3 A USB-C power supply adapter, and the sensors were powered by the Raspberry Pi's MaxLinear MxL7704 regulator. The TGS 2600, 2611, and DHT22 were connected to the A0, A1, and A2 pins of the module, respectively. Figure 1 illustrates the system architecture, providing an overview of its structure and components. A Python script was used to enable continuous reading of data at 1 s intervals. The script leveraged the Inter-Integrated Circuit (I2C) communication protocol to interface with the ADS1115, deploying the Adafruit ADS1x15 library for data acquisition. The acquired data were timestamped and written to daily individual files.

2.1.2. Reference Instrument—Aeris MIRA Ultra Mobile LDS

The Aeris Technologies, Inc. (Hayward, CA, USA) Ultra is a mid-infrared laser absorption analyzer [34]. This instrument integrates a laser absorption spectrometer operating in the mid-infrared spectrum, achieving simultaneous quantification of methane between 10 ppb and 10,000 ppm and ethane between 1 ppb and 1000 ppm [34]. The device operates at standard rates of 1 or 2 Hz, with an option to increase it to 10 Hz. For these experiments, the instrument was set to measure at 5 Hz.

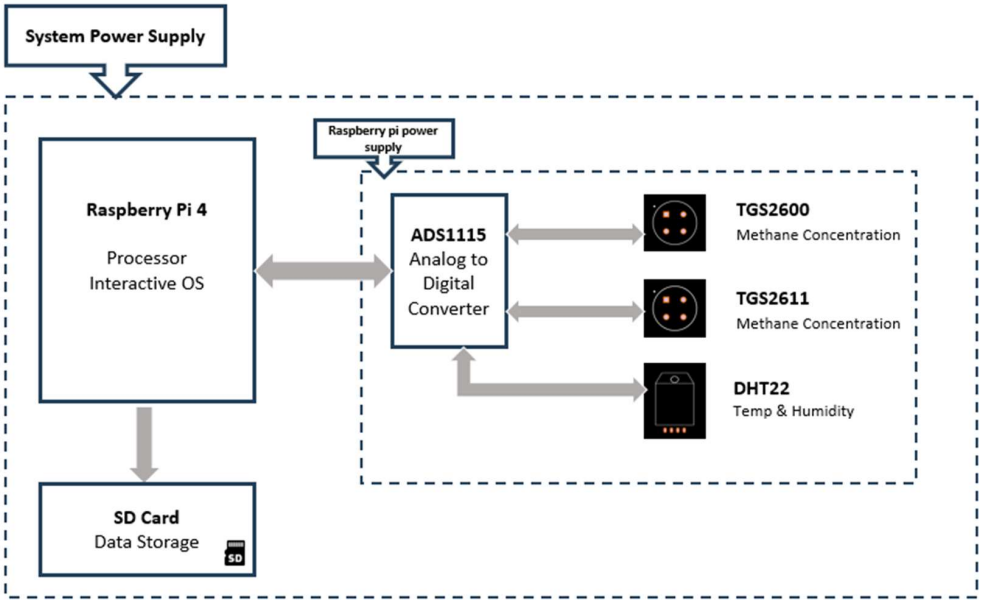


Figure 1. System architecture of sensor integration, power supply, and data acquisition.

2.2. Controlled Methane Release Experiments

Between 1 March 2024 and 2 May 2024, controlled methane release experiments were conducted at Colorado State University's Methane Emission Technology Evaluation Center (METEC) in Fort Collins, Colorado. The experiments simulated real-world oil and gas production emissions by releasing natural gas of known composition from multiple infrastructure points, including wellheads, separators, and tanks. The releases took place at multiple sites (pads 4 and 5), with emissions ranging from 0.005 to 8 kg of methane per hour (CH_4 h^{-1}). The methane releases were both single and multi-source, imitating specific operational scenarios and emissions patterns (Figure 2).

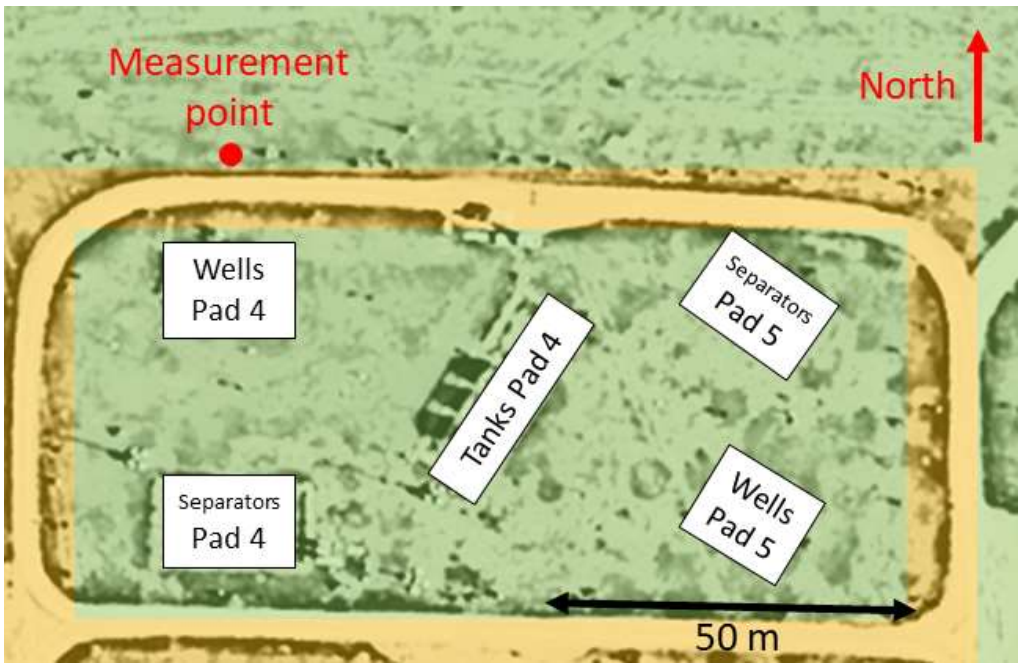


Figure 2. Field layout of the METEC site in Fort Collins with the measurement location on the north of the site.

Throughout this study, the TGS methane sensors and the inlet of the Aeris Sentinel analyzer were located at the north-western part of the facility (approximately 40.60° N, 105.14° W) at a height of 2 m above the ground. The inlets of the TGS methane sensors and the Aeris Sentinel analyzer were collocated to ensure simultaneous sampling of the same air mass (Figure 3). This configuration was intended to minimize spatial variability and provide consistent comparative data from both the low-cost sensors and the high-precision reference analyzer. The experimental period spanned a broad range of environmental conditions, with external temperatures varying significantly from −22 °C earlier in the year to +30 °C towards the end of the deployment.

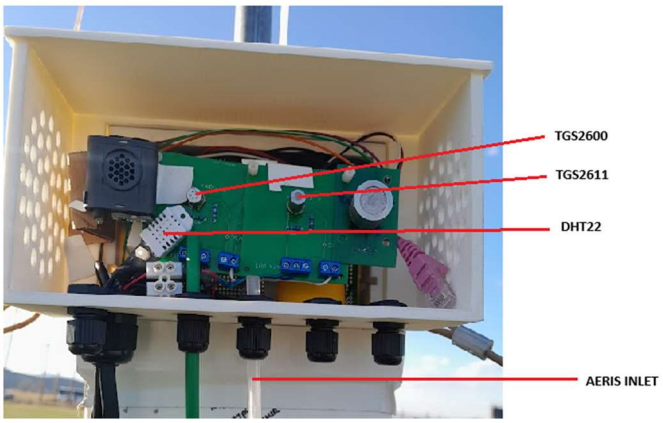


Figure 3. Experimental sampling setup for TGS2600, TGS2611, DHT22, and Aeris Sentinel.

The two sensors were investigated for their resistance in clean air to give insights on their stability. Using the reference analyzer (background of 1.8–2.1 ppm), the resistance of each TGS sensor in air was determined and normalized. The TGS2600 and TGS2611 sensors' raw readings from the analog digital converter had an average of 29,076.85 and 4714.29 within the baseline-to-air range. Within this range, TGS2600 had mean deviations of −65.59, indicating that the sensor resistance was slightly lower than baseline-to-clean air, while TGS2611, with a deviation of 88.60, had readings that were slightly higher than the average clean air resistance. Temperature correlation analysis for TGS2600 was 0.26, showing a slight effect of temperature on the baseline resistance, while for TGS2611 the correlation was −0.28. Humidity had a negligible correlation for TGS2600 (−0.02), while TGS2611 showed a positive correlation of 0.17 with humidity. In both instances, the correlations were weak, and therefore there was no need to compensate for the effects at baseline resistance. Signal-to-noise ratio was also analyzed for both TGS2600 and TGS2611 at baseline to air resistance, and the values were 26.98 and 21.24, respectively, indicating high signal qualities and low noise levels for both sensors in clean air.

2.3. Eugster and Kling (2012) Metal Oxide Sensor Calibration

A calibration following published methods [25,30] was performed on a subset of the TGS2600 and TGS2611 data collected during the controlled release experiments. As described above, the voltage measured across the high tolerance resistor was recorded every second and used to calculate the resistance of the metal oxide strip normalized. This resistance (R_s) was then normalized against the resistance of the metal oxide strip in background methane (R_0) and corrected for temperature and relative humidity (Equation (1)). The $(R_s/R_0)_{corr}$ was plotted against a subset of the contemporaneous methane mixing ratios measured by the Aeris sentinel to generate a calibration algorithm. This algorithm was then used to generate calibrated mixing ratios ($[CH_4]_{cal}$) for every measurement.

2.4. Machine Learning Calibration

2.4.1. Data Preprocessing

In this study, the random forest (RF) algorithm was used to predict methane concentrations from the TGS sensor outputs, temperature, and relative humidity measured by the DHT22. In the RF algorithm, a group of decision trees acts as flowcharts, where the internal nodes represent decisions made on features and branches represent the decision outcome, and the leaf nodes are the final prediction (Figure 4). The first step in RF is the creation of subsets of the original dataset, where data are drawn with replacement from the original dataset. This approach reduces overfitting by enabling each tree to be created with different subsets of data and minimizes the correlation between the trees without increasing the variance [33]. The second stage randomly selects a subset of features to construct decision trees, and this ensures diverse trees since not all trees make the same splits and prevents similarity between trees. By generating a diverse and uncorrelated group of trees, this process helps to minimize predictive errors, ultimately producing a stronger and more reliable final prediction. To make a prediction, the individual tree outputs are aggregated.

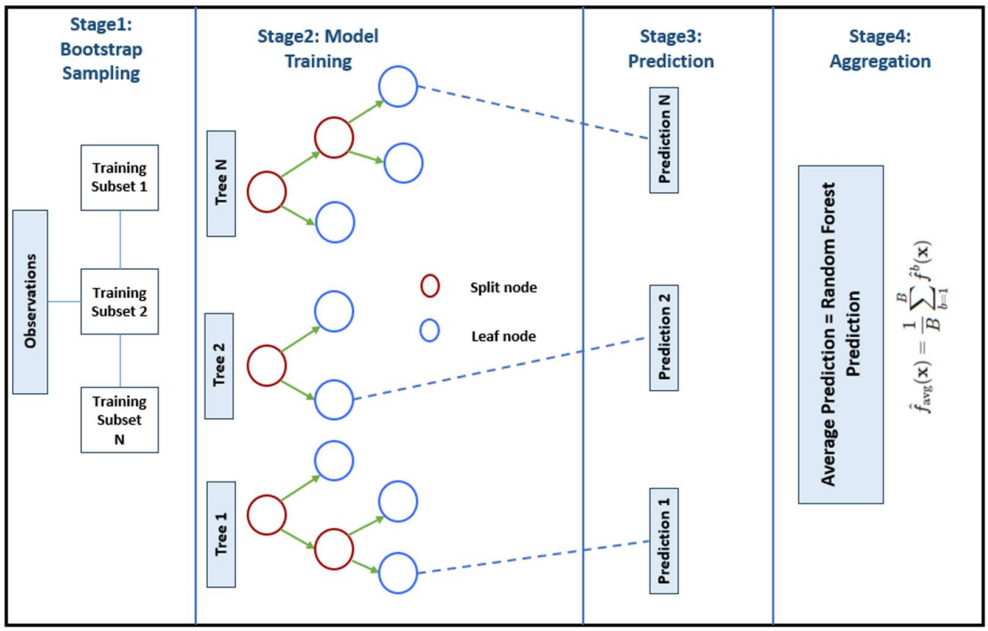


Figure 4. Graphical representation of the random forest (RF) regression process.

The data from the two methane sensors and the DHT22 were compiled in one file. Since the data were collected synchronously from all sensors, each timestamp corresponded to a complete set of sensor readings, hence eliminating the need for imputation and guaranteeing that there were no missing data points. The data gathering script also ensured that all sensor signals were available before being written to the main file but introduced a few extra seconds of delay in cases where the DHT22 sensor signals were not received. This approach guaranteed the completeness of the datasets. With the collocated reference instrument (Aeris MIRA) sampling at a higher frequency than the metal oxide sensor, the Aeris data were averaged to 1s for comparison with the metal oxide sensors' data. A total of 1.6 million data points were collected during this study. However, there were instances of power loss and high winds tipping our data collection mast, which necessitated a 48 h preheating period for the sensors each time this occurred. In addition, we adhered to the manufacturer's specifications by running the sensors for at least a two-day preheating period before starting the tests again [35]. During our analysis, there were instances of high emissions, and we corrected for the sensor time constants, ensuring that the sensors were fully flushed before the next reading. These factors combined led to the elimination of 424,000 data points. Consequently, our model was trained and tested on 518,168 data points,

which constituted 44% of the resultant data points. The datasets used in our model training align with recommendations since tens of thousands of points are generally sufficient to achieve a reliable and accurate result [33,36].

2.4.2. Model Training

The Google Colab environment was employed as the primary environment for training, saving, and making predictions using the model. This environment provided an ideal Graphical Processing Unit (GPU) to efficiently handle the dataset and run the Random Forest Regressor model. Within the environment, the Pandas library was used for data manipulation and analysis, the Scikit-Learn library for machine learning tools, the Joblib library for object serialization, and Matplotlib for visualizations. The RF model input data were split into training and testing using 70–30 criteria, and the split was performed randomly to eliminate bias and ensure a good representation of the data distribution.

Denoting the original dataset as D with N observations, each sample x_i has a probability $\frac{1}{N}$ of being selected each time, and the sample can be selected several times (Equation (2)).

$$D = \{x_1, x_2, \ldots\ldots, x_N\} \quad (2)$$

The bootstrapped dataset, D^*, is expressed as a function of each bootstrapped sample, x_i^*, drawn independently from D with replacement (Equation (3)).

$$D^* = \{x_1^*, x_2^*, \ldots, x_N^*\} \quad (3)$$

There were three input features to the model, and the maximum features hyperparameter of three was defined in the code. Previous studies show that the near-optimum results are given by the square root of the total number of features [33], but in our case we used all the features to create the trees as there were a smaller number of features. The created trees enhance the model's predictive performance by reducing the variances through the creation of child nodes (Figure 5) that are more homogeneous with respect to the target variable.

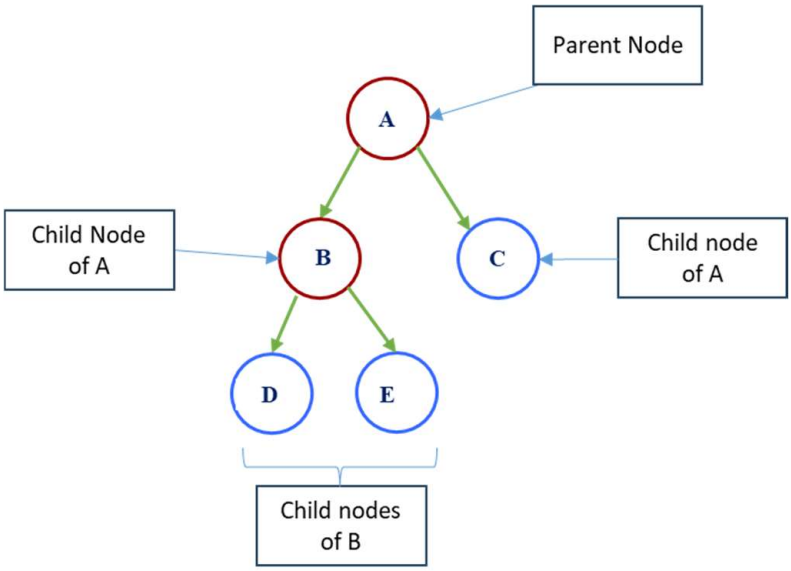


Figure 5. Tree structure showing hierarchical relationships, where A is the root node and parent of B and C. B further branches into two child nodes, D and E, while C has no children, highlighting a typical parent-child relationship within a tree diagram.

The variance of each observation (V_y) is calculated from each concentration observed by the Aeris analyser (y_i), the total number of observations (n), and the mean of the observed values (y^*, Equation (4)).

$$V_y = \frac{1}{n} \sum_{i=1}^{n} (y_i - y^*)^2 \quad (4)$$

Once the values are split into child nodes based on the features, each of the tree nodes (left and right) has an associated variance. The aim of tree splitting is to minimize the overall variance, and this is achieved by evaluating the reduction in variance (V_r) achieved through the comparison of parent node variance (P_{Vy}) to the weighted average of the child nodes' variance for the total number of observations in the parent node (n; Equation (5)). The weighted average of the child nodes' variance is calculated from the number of measurements in the left (n_l) and right (n_r) nodes and the variance of the measurements in the left (V_{yl}) and right (V_{yr}) nodes.

$$V_r = P_{v_y} - \left(\left(\frac{n_l}{n} \times V_{y_l} \right) + \left(\frac{n_r}{n} \times V_{y_l} \right) \right) \quad (5)$$

The final prediction of the RF regression model was determined through the aggregation of the predictions from individual trees in the forest. K-fold cross-validation was employed to evaluate the performance of the RF models for the two sensors. Due to the size of the available computing resources, a stratified k-fold approach with 5 folds was used, providing a reliable estimate of the model performance.

The trained random forest regressor model was stored and used to calculate mixing ratios based on the metal oxide and DHT22 temperature and humidity outputs.

2.4.3. Model Evaluation Metrics

Methane mixing ratios calculated using the ML approach were compared to mixing ratios measured by the Aeris Ultra analyzer output using the coefficient of determination (R^2) of the linear regression between the data, the mean absolute error (MAE), the mean squared error (MSE), the root mean squared error (RMSE), the mean absolute percentage error (MAPE), and the explained variance score. The feature importance, which is the measure of how each input variable contributes to the model prediction, was also evaluated for the ML calculated data. Feature importance is determined by evaluating the reduction in impurity (a measure of how mixed or uncertain the target variable is within a node) at each node, regulated by the probability of reaching that node [36]. The node probability is the number of nodes reaching that node divided by the total number of samples. The higher the value, the more important the feature, and higher values indicate greater feature importance. Histograms of measured and modeled mixing ratio data, consisting of 0.2 ppm bins truncated at <2 ppm and >5 ppm, were compared using the overlap coefficient to investigate the overlap of data.

3. Results

Methane mixing ratios observed by the Aeris Ultra analyzer ranged from background (2 ppm) to 147.6 ppm (Figure 6). Throughout the experiment, the temperature varied with large diurnal cycles from a maximum of 27 °C to a minimum of −5 °C, with relative humidity cycling similarly between a maximum of 91% and a minimum of 18%.

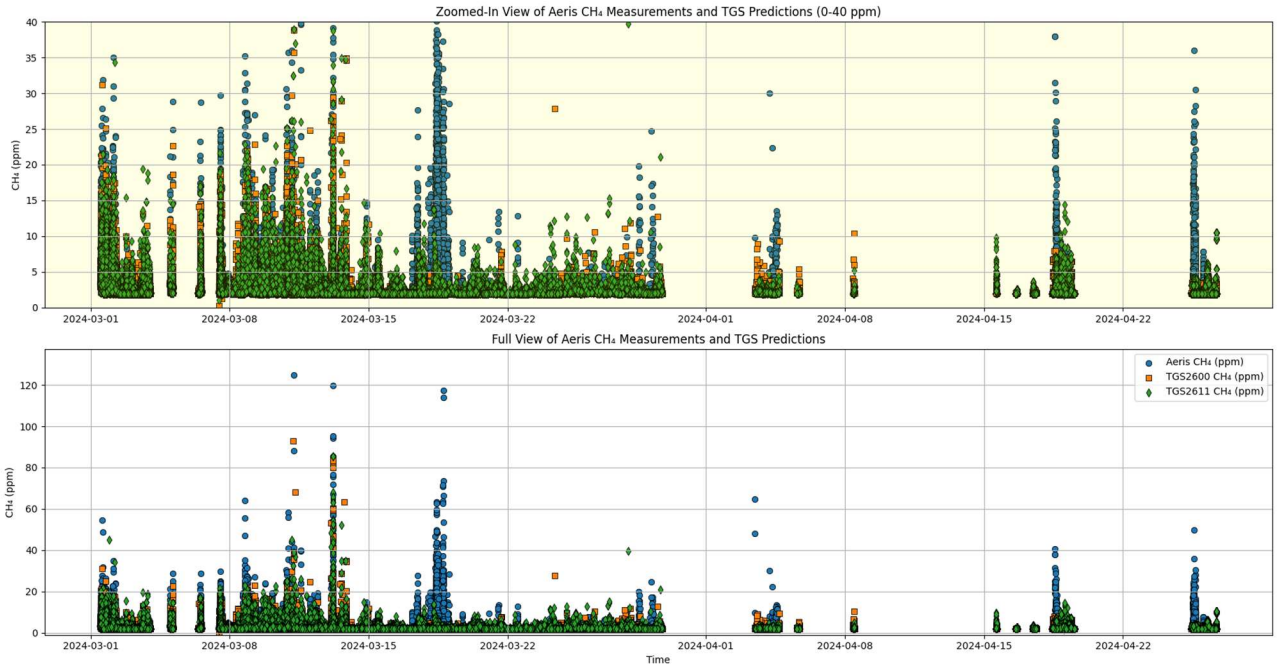


Figure 6. Time series of methane mixing ratios measured by the Aeris Sentinel, alongside predictions from the TGS2600 and TGS2611 sensors. The top panel shows a zoomed-in view (0–40 ppm) of the methane mixing ratios for detailed comparison, while the bottom panel provides the full range of data. Methane mixing ratios for the Aeris Sentinel are represented in blue, predictions from the TGS2600 sensor with a machine learning calibration curve are in orange, and predictions from the TGS2611 sensor with a machine learning calibration curve are in green.

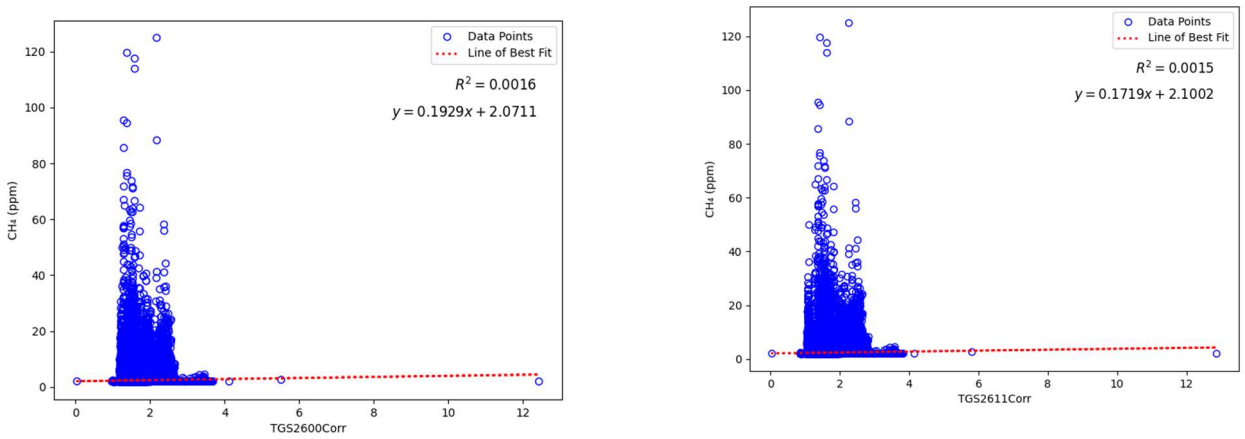
3.1. Eugster and Kling (2012) Calibration Method

Calibration Curves

Previous studies describing the use of metal oxide sensors to measure methane concentrations in the air have reported either linear or exponential relationships between $(R_s/R_0)_{corr}$ and methane mixing ratios using methane analyzer [17,23–26,30,31,37]. Unlike these studies, the TGS sensors used in this study did not respond in the same way, and there was no discernible relationship between $(R_s/R_0)_{corr}$ and methane mixing ratio with R^2 of 0.0016 and 0.0016 for the TGS2600 and TGS2611, respectively (Figure 7A,B).

It was later discovered that the resistor used to form the Wheatstone bridge with the TGS sensors was 20 times larger than suggested by other studies [25,30] and resulted in a very small signal-to-noise ratio that had not been observed by the other studies. This meant that the data reported by the TGS sensors used in this study could not be calibrated using the methods described by other papers. When the linear calibration algorithms were used to calculate the methane mixing ratio and compared against the measurement data for the rest of the campaign (12 March–2 May 2024), the correlation is zero.

The sensors were analyzed for cross-sensitivity during the period of deployment. Data from the reference analyzer were used to identify 10 min periods where the sensors were subject to background methane concentrations before and after deployment. The resistance of the sensors in ambient methane concentrations in relatively constant temperature and humidity (±0.5 °C and ±2%, respectively) were used to derive a deviation of the signal. Both the TGS2600 and TGS2611 sensors showed low cross-sensitivity indices of 0.000063 and 0.00027, respectively. Regarding selectivity of gas detection and following the sensors' datasheets, the metal oxide sensors are cross-sensitive to carbon monoxide, iso-butane, ethanol, and hydrogen. However, these gases are unlikely to be a significant source of contamination [30,31,35].



(**A**) (**B**)

Figure 7. Data used to generate a calibration curve following the methods of Eugster and Kling (2012) for the (**A**) TGS2600 and (**B**) the TGS2611. $(R_s/R_0)_{corr}$ plotted on the *x*-axis and methane mixing ratios measured using the Aeris Sentinel on the *y*-axis.

3.2. Machine Learning Methods

A time series plot of the RF predicted methane concentrations showed a similar pattern to the actual measured concentrations from the Aeris analyzer (Figure 6). The model demonstrated a high sensitivity to the minor variations in input features, enabling noticeable concentration levels with slight changes in sensor and environmental readings. The average measured concentration for the Aeris analyzer (during the measurement period) was 2.40 ppm with a maximum of 147.6 ppm compared to the average calculated mixing ratios of 2.42 ppm and 2.40 ppm with a maximum value of 117.5 ppm and 106.3 ppm for the TGS2600 and TGS2611, respectively. The gradient of the regression between the measured mixing ratios and those calculated from the TGS2600 data was 0.71 with an R^2 of 0.34, while the gradient for the TGS2611 was 0.69 and an R^2 of 0.38.

The mean squared error values of 1.40 for the TGS2600 and 1.32 for the TGS2611 show the presence of extensive prediction errors, with RMSE values of 1.18 ppm and 1.14 ppm suggesting that typical concentration prediction errors were around these figures (Table 1). The mean absolute percentage error suggests that the model's predictions were, on average, off by approximately 7.21% and 8.11% from the actual values for the TGS2600 and 2611, respectively. Finally, the explained variance scores, indicated that the model accounted for about 33% and 38% of the variance in CH_4 concentrations for TGS2600 and TGS2611, respectively.

Table 1. Performance metrics for the mixing ratios generated using the TGS2600 data with the RF algorithm (TGS2600-RF) and mixing ratios generated using the TGS2611 data with the RF algorithm (TGS2611-RF) when compared to measured mixing ratios.

Metric	TGS2600-RF	TGS2611-RF
R-squared (R^2)	0.34	0.38
Mean absolute error (MAE)	0.24	0.26
Mean squared error (MSE)	1.40	1.32
Root mean squared error (RMSE)	1.18	1.14
Mean absolute percentage error (MAPE)	7.21	8.11
Explained variance score	0.33	0.38

The relative contribution of each feature was also assessed, and from these, the TGS2600 feature importance scores indicate that the relative humidity feature is the most influential, accounting for approximately 35.8% of the model's predictive power, the sensor

resistance at 33.5%, and temperature at 30.7%. In the case of the TGS2611, humidity again emerges as the most significant feature, contributing 39.5% to the model's predictions, while the resistance accounts for 32.1% and temperature contributes 28.4%. The consistent significance of humidity across both sensor models suggests its critical role in the accuracy of sensor measurements.

The model was also tested against the training and test datasets, and it demonstrated strong performance on the training data, with an R-squared value of 0.81 and 0.79 for TGS2600 and TGS2611, respectively. However, the R-squared values on the test data were 0.69 and 0.71, revealing a drop in predictive accuracy when applied to new data. This suggests that the model has become too closely aligned with the specific patterns in the training data, reducing its ability to perform as well on unseen data. This pattern is also evident when new and considerably large datasets with new instances are applied to the model and the R-squared value decreases. While the test set performance is still reasonably good, there is room for improvement in enhancing the model's generalizability. Future improvements could be achieved by applying techniques such as hyperparameter tuning or increasing the size of the training dataset. Additionally, exploring other ensemble models might lead to better overall performance.

Histograms, comprising 17 0.2 ppm bins truncated at <2 ppm and >5 ppm (Figure 8), show all datasets are heavily right skewed. Comparisons using the overlap coefficient, also known as the Szymkiewicz–Simpson coefficient, show an overlap between the measured mixing ratios and those mixing ratios generated by TGS2600 and TGS2611 of 0.95 and 0.94, respectively. This means that the measured mixing ratios and mixing ratios generated by TGS2600 have 95% of distribution overlapping, whereas with TGS2611, it is 94%.

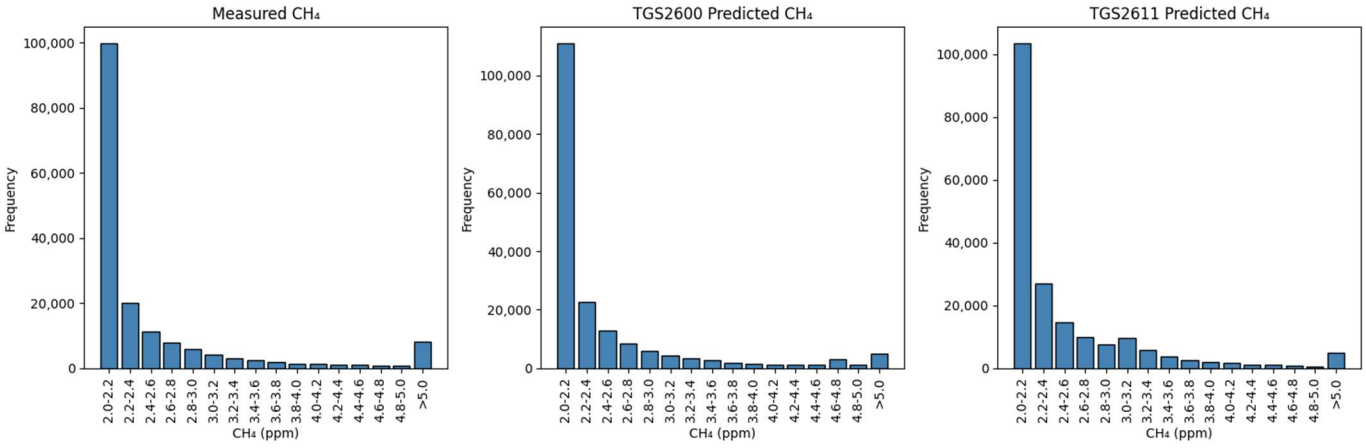


Figure 8. Histograms, comprising 15 0.2 ppm bins truncated at <2 ppm and >5 ppm, for the measured mixing ratios (**left pane**), the mixing ratios calculated using the TGS2600 data with the RF algorithm (**center pane**), and mixing ratios calculated using the TGS2611 data with the RF algorithm (**right pane**).

4. Discussion

This study reports the response of two metal oxide sensors, Figaro TGS2600 and TGS2611, to methane emissions at Colorado State University's METEC site between 1 March and 2 May 2024. Methane mixing ratios observed by the reference analyzer ranged from background to 119 ppm (Figure 7A), while both TGS sensor responses showed diurnal variation in response to changes in temperature and relative humidity but no obvious response to changes in methane mixing ratio. When the TGS response data were used to generate calibration curves following published methods [25,30], there was no correlation between the calculated mixing ratios and the measured mixing ratios (Figure 7A,B). A set of training data were then used with the Random Forest machine learning algorithm using the resistance, temperature, and relative humidity data to map to the measured reference methane mixing ratio data.

4.1. Random Forest Calibration Versus Linear Regression Calibration

The random forest algorithm performed better than the linear regression model. The R^2 for the linear regression model for TGS2600 and TGS2611 are 0.004 and 0.005, respectively, compared to the R^2 obtained by the RF model, 0.34 and 0.38, respectively (Table 1). The improvement in R^2 can be attributed to the ability of the RF algorithm to capture non-linear and convoluted interrelations between the sensor outputs and the target variable. In addition, the ability of the RF model to implement bootstrapping and feature randomness reduces the risk of overfitting while allowing the model to popularize better to new datasets. This study also assessed the effect of the incorporation of larger datasets and advanced techniques such as k-fold cross validation on the model performance. By expanding the model's training dataset, the model was able to capture the underlying patterns and improve the accuracy of the unseen data. These improvements, however, often require increased computational power for extensive model evaluation.

4.2. Influence of Humidity and Temperature

The signal-to-noise ratio of the metal oxide data was too low for previously used calibration methods to be employed. The sensors were deployed through periods of low and high temperatures and also at high fluctuations in humidity. Therefore, the RF algorithm was able to identify the underlying complex relationship between the sensor resistances and the variable conditions that the sensors were exposed to. Humidity was found to be the most influential factor for both sensors, and this could be as a result of the formation of a layer of water molecules on the sensing element, altering the sensor's response greatly. The quality of methane mixing ratio data calculated using the machine-learning-derived algorithms is similar to those reported previously, even though there was a much lower signal-to-noise ratio of input data. Eugster and Kling (2012) reported a R^2 between modeled and measured data between 1.85 and 1.00 ppm of 0.19 [25] for TGS2600, while this study reports an R^2 over the same range of mixing ratios of 0.13. Collier-Oxandale et al. (2018) reported an RMSE of 0.38 for mixing ratios between 2 and 7 ppm [38] for TGS2600, while this study reports an RMSE of 1.18. Riddick et al. (2020a) reported an accuracy of $\pm$0.01 ppm between 1.85 and 5.85 ppm, while this study reports an accuracy of $\pm$0.24 ppm.

While our study generated 1.6 million data points, Bastviken et al. (2020) used a simpler calibration model but generated a smaller dataset, between 619 and 930 datum per sensor. They placed a Figaro NGM2611-E13, which is a calibrated version of the TGS2611, with a humidity and temperature sensor in a real-world deployment scenario but did not explore the interactions between the input variables [39]. Collier-Oxandale et al. (2018) used methane sensors with reference instruments during field deployment [40]. Our study used a similar approach, only that we used controlled releases in a field setting. That study noted longer deployment to capture high emission events, which we had control over in our study due to the nature of our releases. Rivera Martinez et al. (2021) calibrated a TGS2611-E00 methane sensor across relatively narrow temperature and humidity ranges (15–30 °C, 40–80%) and 2–9 ppm methane ranges and co-located with an Picarro G2301 Cavity Ringdown Gas Analyzer [41]. They used a linear model, whereas our study used the RF algorithm, which allows for calibration across a broader range of methane levels. As a result of the varying models used, the setup, and concentration ranges, it is hard to compare their metrics with ours. However, all these studies report the importance of factoring in the humidity and temperature for sensor calibration and the use of a precise reference analyzer.

5. Conclusions

This study provides evidence that machine learning approaches could be used to generate sensor calibration algorithms to better constrain methane mixing ratios measured using low-cost methane sensors. Practically, this would mean that machine learning could be used to derive algorithms to better define complex relationships between sensor

outputs and environmental conditions to enable better calibration and improve the metal oxide sensor's accuracy. The main advantages of using machine learning algorithms for generating calibration algorithms are that they can model complex relationships between sensor outputs and environmental conditions, and they reduce the effects of overfitting by using multiple trees. The shortcomings are that with larger datasets, the computational costs increase significantly and that the model interaction between the features is less interpretable compared to simpler models, like linear regression models. The development of machine learning algorithms is increasing exponentially, and it is expected that in the future the model can perform better even with less training data while being less affected by noise or sensor anomalies.

This study investigated the use of a machine learning model to calibrate poor-quality, low-cost, metal oxide sensor data to generate representative methane mixing ratio data. Currently, calibration of metal oxide sensors requires a sensor output with a high signal-to-noise ratio. The high importance of humidity in both TGS2600 and TGS2611 sensors underscores the importance of environmental variability when developing predictive models for gas detection.

Our results also showed that the machine learning derived mixing ratios were as good as those reported by other studies using better signal-to-noise ratio data. While our findings show that machine learning techniques can improve the predictive capabilities of lower-cost sensors such as the TGS2600 and TGS2611 for methane detection, the current performance, with R^2 values of around 0.34 and 0.38, indicates there is still considerable work to be carried out before these sensors can match the precision and accuracy of high-end trace gas analyzers. This would improve the low-cost sensors' responses, give more confidence to calculated emissions, and reduce the uncertainty in greenhouse gas emission inventories.

Author Contributions: E.K.: conceptualization, investigation, methodology, supervision, writing—original draft preparation, review, and editing; S.N.R.: funding acquisition, conceptualization, investigation, methodology, supervision, writing—original draft preparation, review and editing; M.M.: investigation, review, and editing; A.U.: review and editing; A.A.: review and editing; D.J.Z.: funding acquisition, project administration, conceptualization, supervision, review, and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This project has been funded by the U.S. Department of Energy's Office of Fossil Energy and Carbon Management (FECM), project # DE-FE0032276 "Capabilities Enhancement for Methane Emissions Technology Evaluation Center (METEC) to Decarbonize Natural Gas Resources".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in this article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. IPCC. *Climate Change 2013—The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2014; ISBN 978-1-107-41532-4.
2. IPCC. *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2022; ISBN 978-1-00-932584-4.
3. Global Methane Pledge. Global Methane Pledge—Fast Action on Methane to Keep a 1.5 °C Future Within Reach. Available online: www.globalmethanepledge.org (accessed on 15 September 2022).
4. UNFCCC. Paris Agreement. United Nations Framework Convention on Climate Change. FCCC/CP/2015/L.9/Rev.1. Available online: https://unfccc.int/documents/9064 (accessed on 16 June 2023).
5. UNEP/CCAC (United Nations Environment Programme and Climate and Clean Air Coalition). *Global Methane Assessment: Benefits and Costs of Mitigating Methane Emissions*; United Nations Environment Programme: Nairobi, Kenya, 2021. Available online: https://www.ccacoalition.org/en/resources/global-methane-assessment-full-report (accessed on 25 October 2022).

6. Bell, C.; Ilonze, C.; Duggan, A.; Zimmerle, D. Performance of Continuous Emission Monitoring Solutions under a Single-Blind Controlled Testing Protocol. *Environ. Sci. Technol.* **2023**, *57*, 5794–5805. [CrossRef] [PubMed]
7. Caulton, D.R.; Lu, J.M.; Lane, H.M.; Buchholz, B.; Fitts, J.P.; Golston, L.M.; Guo, X.; Li, Q.; McSpiritt, J.; Pan, D.; et al. Importance of Superemitter Natural Gas Well Pads in the Marcellus Shale. *Environ. Sci. Technol.* **2019**, *53*, 4747–4754. [CrossRef] [PubMed]
8. Riddick, S.N.; Cheptonui, F.; Yuan, K.; Mbua, M.; Day, R.; Vaughn, T.L.; Duggan, A.; Bennett, K.E.; Zimmerle, D.J. Estimating Regional Methane Emission Factors from Energy and Agricultural Sector Sources Using a Portable Measurement System: Case Study of the Denver–Julesburg Basin. *Sensors* **2022**, *22*, 7410. [CrossRef] [PubMed]
9. Albertson, J.D.; Harvey, T.; Foderaro, G.; Zhu, P.; Zhou, X.; Ferrari, S.; Amin, M.S.; Modrak, M.; Brantley, H.; Thoma, E.D. A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production. *Environ. Sci. Technol.* **2016**, *50*, 2487–2497. [CrossRef]
10. Harrison, M.R.; Cowgill, R.M.; Campbell, L.M.; Lott, R.A. *Methane Emissions from the Natural Gas Industry*; Final Report, GRI-94/0257 and EPA-600/R-96-080; Gas Research Institute and US Environmental Protection Agency: Washington, DC, USA, 1996; Volumes 1–15.
11. Campbell, L.M.; Campbell, M.V.; Epperson, D.L. *Methane Emissions from the Natural Gas Industry, Volume 2: Technical Report*; Final Report, GRI-94/0257.1 and EPA-600/R-96-080b; Gas Research Institute and U.S. Environmental Protection Agency: Washington, DC, USA, 1996.
12. Riddick, S.N.; Mauzerall, D.L. Likely Substantial Underestimation of Reported Methane Emissions from United Kingdom Upstream Oil and Gas Activities. *Energy Environ. Sci.* **2023**, *16*, 295–304. [CrossRef]
13. Caulton, D.R.; Li, Q.; Bou-Zeid, E.; Fitts, J.P.; Golston, L.M.; Pan, D.; Lu, J.; Lane, H.M.; Buchholz, B.; Guo, X.; et al. Quantifying Uncertainties from Mobile-Laboratory-Derived Emissions of Well Pads Using Inverse Gaussian Methods. *Atmos. Chem. Phys.* **2018**, *18*, 15145–15168. [CrossRef]
14. Peischl, J.; Eilerman, S.J.; Neuman, J.A.; Aikin, K.C.; de Gouw, J.; Gilman, J.B.; Herndon, S.C.; Nadkarni, R.; Trainer, M.; Warneke, C.; et al. Quantifying Methane and Ethane Emissions to the Atmosphere From Central and Western U.S. Oil and Natural Gas Production Regions. *J. Geophys. Res. Atmos.* **2018**, *13*, 7725–7740. [CrossRef]
15. Pétron, G.; Karion, A.; Sweeney, C.; Miller, B.R.; Montzka, S.A.; Frost, G.J.; Trainer, M.; Tans, P.; Andrews, A.; Kofler, J.; et al. A New Look at Methane and Nonmethane Hydrocarbon Emissions from Oil and Natural Gas Operations in the Colorado Denver-Julesburg Basin. *J. Geophys. Res. Atmos.* **2014**, *119*, 6836–6852. [CrossRef]
16. Barkley, Z.; Davis, K.; Miles, N.; Richardson, S.; Deng, A.; Hmiel, B.; Lyon, D.; Lauvaux, T. Quantification of Oil and Gas Methane Emissions in the Delaware and Marcellus Basins Using a Network of Continuous Tower-Based Measurements. *Atmos. Chem. Phys.* **2023**, *23*, 6127–6144. [CrossRef]
17. Riddick, S.N.; Ancona, R.; Cheptonui, F.; Bell, C.S.; Duggan, A.; Bennett, K.E.; Zimmerle, D.J. A Cautionary Report of Calculating Methane Emissions Using Low-Cost Fence-Line Sensors. *Elem. Sci. Anthr.* **2022**, *10*, 00021. [CrossRef]
18. Cho, Y.; Smits, K.M.; Riddick, S.N.; Zimmerle, D.J. Calibration and Field Deployment of Low-Cost Sensor Network to Monitor Underground Pipeline Leakage. *Sens. Actuators B Chem.* **2022**, *355*, 131276. [CrossRef]
19. Vaughn, T.L.; Bell, C.S.; Pickering, C.K.; Schwietzke, S.; Heath, G.A.; Pétron, G.; Zimmerle, D.J.; Schnell, R.C.; Nummedal, D. Temporal Variability Largely Explains Top-down/Bottom-up Difference in Methane Emission Estimates from a Natural Gas Production Region. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11712–11717. [CrossRef]
20. Riddick, S.N.; Mbua, M.; Santos, A.; Hartzell, W.; Zimmerle, D.J. Potential Underestimate in Reported Bottom-up Methane Emissions from Oil and Gas Operations in the Delaware Basin. *Atmosphere* **2024**, *15*, 202. [CrossRef]
21. Bell, C.S.; Vaughn, T.L.; Zimmerle, D.; Herndon, S.C.; Yacovitch, T.I.; Heath, G.A.; Pétron, G.; Edie, R.; Field, R.A.; Murphy, S.M.; et al. Comparison of Methane Emission Estimates from Multiple Measurement Techniques at Natural Gas Production Pads. *Elem. Sci. Anth.* **2017**, *5*, 79. [CrossRef]
22. Bell, C.; Rutherford, J.; Brandt, A.; Sherwin, E.; Vaughn, T.; Zimmerle, D. Single-Blind Determination of Methane Detection Limits and Quantification Accuracy Using Aircraft-Based LiDAR. *Elem. Sci. Anthr.* **2022**, *10*, 00080. [CrossRef]
23. Shah, A.; Laurent, O.; Lienhardt, L.; Broquet, G.; Rivera Martinez, R.; Allegrini, E.; Ciais, P. Characterising the Methane Gas and Environmental Response of the Figaro Taguchi Gas Sensor (TGS) 2611-E00. *Atmos. Meas. Tech.* **2023**, *16*, 3391–3419. [CrossRef]
24. Shah, A.; Laurent, O.; Broquet, G.; Philippon, C.; Kumar, P.; Allegrini, E.; Ciais, P. Determining Methane Mole Fraction at a Landfill Site Using the Figaro Taguchi Gas Sensor 2611-C00 and Wind Direction Measurements. *Environ. Sci. Atmos.* **2024**, *4*, 362–386. [CrossRef]
25. Eugster, W.; Kling, G.W. Performance of a Low-Cost Methane Sensor for Ambient Concentration Measurements in Preliminary Studies. *Atmos. Meas. Tech.* **2012**, *5*, 1925–1934. [CrossRef]
26. Eugster, W.; Laundre, J.; Eugster, J.; Kling, G.W. Long-Term Reliability of the Figaro TGS 2600 Solid-State Methane Sensor under Low-Arctic Conditions at Toolik Lake, Alaska. *Atmos. Meas. Tech.* **2020**, *13*, 2681–2695. [CrossRef]
27. Lin, J.J.Y.; Buehler, C.; Datta, A.; Gentner, D.R.; Koehler, K.; Zamora, M.L. Laboratory and Field Evaluation of a Low-Cost Methane Sensor and Key Environmental Factors for Sensor Calibration. *Environ. Sci. Atmos.* **2023**, *3*, 683–694. [CrossRef]
28. Nagahage, I.S.P.; Nagahage, E.A.A.D.; Fujino, T. Assessment of the Applicability of a Low-Cost Sensor–Based Methane Monitoring System for Continuous Multi-Channel Sampling. *Environ. Monit. Assess.* **2021**, *193*, 509. [CrossRef] [PubMed]
29. Sugriwan, I.; Soesanto, O. Development of TGS2611 Methane Sensor and SHT11 Humidity and Temperature Sensor for Measuring Greenhouse Gas on Peatlands in South Kalimantan, Indonesia. *J. Phys. Conf. Ser.* **2017**, *853*, 012006. [CrossRef]

30. Riddick, S.N.; Mauzerall, D.L.; Celia, M.; Allen, G.; Pitt, J.; Kang, M.; Riddick, J.C. The Calibration and Deployment of a Low-Cost Methane Sensor. *Atmos. Environ.* **2020**, *230*, 117440. [CrossRef]
31. Figaro Production Information. TGS 2611—For the Detection of Methane. Available online: https://www.figarosensor.com/product/docs/tgs%202611c00(1013).pdf (accessed on 26 August 2020).
32. Andrews, B.; Chakrabarti, A.; Dauphin, M.; Speck, A. Application of Machine Learning for Calibrating Gas Sensors for Methane Emissions Monitoring. *Sensors* **2023**, *23*, 9898. [CrossRef]
33. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
34. AERIS TECHNOLOGIES. AERIS TECHNOLOGIES. MIRA Mobile Methane/Ethane Analyzers. Aeris Technologies. Available online: https://aerissensors.com/mira-pico-mobile-lds/ (accessed on 23 August 2024).
35. FIGARO TGS 2600—for the Detection of Air Contaminants. 2013. Available online: https://www.figarosensor.com/product/docs/TGS2600B00%20(0913).pdf (accessed on 23 September 2024).
36. Han, S.; Kim, H. Optimal Feature Set Size in Random Forest Regression. *Appl. Sci.* **2021**, *11*, 3428. [CrossRef]
37. Ellis, C. Hyperparameter Tuning in Random Forests. Available online: https://crunchingthedata.com/hyperparameter-tuning-in-random-forests/ (accessed on 13 August 2024).
38. Furuta, D.; Sayahi, T.; Li, J.; Wilson, B.; Presto, A.A.; Li, J. Characterization of Inexpensive Metal Oxide Sensor Performance for Trace Methane Detection. *Atmos. Meas. Tech.* **2022**, *15*, 5117–5128. [CrossRef]
39. Bastviken, D.; Nygren, J.; Schenk, J.; Parellada Massana, R.; Duc, N.T. Technical Note: Facilitating the Use of Low-Cost Methane (CH_4) Sensors in Flux Chambers—Calibration, Data Processing, and an Open-Source Make-It-Yourself Logger. *Biogeosciences* **2020**, *17*, 3659–3667. [CrossRef]
40. Collier-Oxandale, A.; Casey, J.G.; Piedrahita, R.; Ortega, J.; Halliday, H.; Johnston, J.; Hannigan, M.P. Assessing a Low-Cost Methane Sensor Quantification System for Use in Complex Rural and Urban Environments. *Atmos. Meas. Tech.* **2018**, *11*, 3569–3594. [CrossRef]
41. Rivera Martinez, R.; Santaren, D.; Laurent, O.; Cropley, F.; Mallet, C.; Ramonet, M.; Caldow, C.; Rivier, L.; Broquet, G.; Bouchet, C.; et al. The Potential of Low-Cost Tin-Oxide Sensors Combined with Machine Learning for Estimating Atmospheric CH_4 Variations around Background Concentration. *Atmosphere* **2021**, *12*, 107. [CrossRef]

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.