Machinery & Energetics
Vol. 16, No. 1. 2025

Journal homepage:
https://technicalscience.com.ua/en

UDC 89:528.8
Doi: 10.31548/machinery/1.2025.146

Kamala Aghayeva
PhD
Azerbaijan State Oil and Industry University
AZ1010, 34 Azadliq Ave., Baku, Azerbaijan
https://orcid.org/0009-0006-6779-2550

German Krauklit*
Postgraduate Student
Azerbaijan National Aerospace Agency
AZ1115, 1 Akhundov Str., Baku, Azerbaijan
https://orcid.org/0000-0002-6171-9348

Automated methane emission monitoring systems based on satellite data: Radiation transfer model analysis

Abstract. The study aimed to assess the accuracy of automated methane emission monitoring systems at oil and gas fields in Azerbaijan using satellite data and a radiation transfer model. The methodology included analysing Sentinel-5P and GHGSat satellite data for 2024, applying MODTRAN and SCIATRAN models to incorporate atmospheric factors, and validating the results with ground measurements using Los Gatos Research spectrometers and Picarro G2401 gas analysers. The results demonstrated that GHGSat determined localised emissions ($R^2 = 0.89$, RMSE = 4.7 ppb) with greater accuracy, while Sentinel-5P demonstrated underestimation of concentrations in high humidity ($R^2 = 0.72$, RMSE = 12.4 ppb). Data correction using the MODTRAN and SCIATRAN models improved the accuracy of the measurements: RMSE decreased to 7.8 ppb for Sentinel-5P and 4.2 ppb for GHGSat. The highest methane emissions were detected on the Apsheron Peninsula (2.8 ppb), which is associated with leaks and gas processing processes. The seasonal analysis demonstrated an increase in concentrations in winter (3 ppm) and summer (2.7 ppm) due to a decrease in the dispersion rate and intensification of mining. Machine learning methods (XGBoost, Random Forest) improved the forecasting accuracy ($R^2 = 0.91$ for XGBoost) by identifying key factors: wind speed, temperature, mining intensity and humidity. The findings highlight the need to integrate satellite data with ground-based measurements and radiative transfer models to improve monitoring accuracy and develop emission reduction strategies

Keywords: machine learning; spatial distribution; oil and gas industry; ground-based measurements; climate processes

INTRODUCTION

Methane emissions monitoring is central in the assessment of the environmental and climate impact of the oil and gas industry. Methane is one of the most substantial greenhouse gases with a high ability to accumulate heat in the atmosphere and significantly affect climate processes. The main sources of anthropogenic emissions are pipeline leaks, oil and gas production and processing processes, and associated gas discharges. However, accurate and timely control of emissions remains a challenge, as traditional monitoring methods, including ground-based measurements, have limited spatial and temporal resolution. In this regard, automated systems using satellite data are a promising solution that can be used to monitor methane emissions over large areas with high frequency and accuracy.

One of the main challenges in the use of satellite data is the impact of atmospheric factors such as humidity, temperature and aerosol load on measurement accuracy. Studies demonstrated that high humidity could cause a

Article's History: Received: 24.10.2024; Revised: 16.01.2025; Accepted: 27.02.2025.

Aghayeva, K., & Krauklit, G. (2025). Automated methane emission monitoring systems based on satellite data: Radiation transfer model analysis. *Machinery & Energetics,* 16(1), 146-156. doi: 10.31548/machinery/1.2025.146.

*Corresponding author

systematic underestimation of methane concentrations recorded by satellite instruments. A. Tanzharikov *et al*. (2023) addressed the impact of meteorological conditions on the accuracy of satellite measurements and found that changes in the level of water vapour in the atmosphere significantly affect the infrared absorption coefficient. However, their model did not incorporate the variability of aerosol characteristics, which limited the accuracy of the data obtained.

The accuracy of satellite methane monitoring depends not only on atmospheric conditions but also on the data processing algorithms used (Rusho *et al*., 2024). The development of machine learning methods has significantly improved the analysis of satellite measurements. K. Aghayeva & G. Krauklit (2024) proposed the use of XGBoost and Random Forest algorithms to correct satellite data considering meteorological factors such as wind speed, temperature, and mining intensity. During the experiments, the author determined that machine learning can reduce the root mean square error (RMSE) of measurements and increase the consistency of data with ground observations. However, the proposed models required a large amount of pre-labelled data, which made them difficult to apply in practice.

Detection of localised sources of methane emissions is one of the key monitoring tasks, as oil and gas fields are often characterised by an uneven distribution of leaks (Deryaev, 2023). L.R. Bekirova & E.M. Bunyatova (2024) conducted a comparative analysis of the accuracy of Sentinel-5P and GHGSat, determining that Sentinel-5P demonstrates lower accuracy in localising small emission sources due to its lower spatial resolution. However, their study did not consider the impact of seasonal changes on the quality of measurements, which leaves open the question of the need for additional data correction in different periods of the year.

Seasonal changes in methane emissions are an important factor affecting the accuracy of emissions forecasting. Fluctuations in temperature and wind speed can lead to significant changes in the concentration of gas in the atmosphere (Satin & Kutsyi, 2024). B. Schuit *et al*. (2023) studied the impact of seasonal factors on methane distribution and determined that in winter, concentrations increase due to low turbulence, while in summer, emissions increase due to increased production. However, the study did not address possible abnormal emissions associated with emergencies, which limits the ability to use the data for operational emissions control.

Ground-based measurements are valuable for the validation of satellite data, but their use is subject to technical limitations due to the spatial discreteness of the observations. To address this problem, methods for the combined use of spectrometers and gas analysers have been developed. B. Andrews *et al*. (2023) compared the performance of Los Gatos Research spectrometers and Picarro G2401 gas analysers in calibrating satellite data, finding that combining the two methods reduced measurement error. However, the study did not include the possible influence of atmospheric turbulence, which could distort the results.

The use of radiation transfer models addresses the influence of atmospheric factors on infrared radiation propagation and improves the accuracy of satellite measurements. G. Domènech-Gil *et al*. (2023) investigated the effectiveness of MODTRAN and SCIATRAN models for data correction and found that their use reduces systematic errors, especially in high humidity conditions. However, the study did not consider the impact of different types of underlying surfaces on the characteristics of reflected radiation, which can lead to additional uncertainties in the calculations.

Identification of methane leaks requires the integration of different types of data, including satellite observations, ground-based measurements and meteorological data. Q. Chen *et al*. (2023) developed an integrated methodology for analysing emissions by combining satellite imagery, gas analyser data and atmospheric models. The results demonstrated that this approach identifies leaks with high accuracy and estimates their intensity. However, the proposed methodology requires significant computational resources, which limits its operational application in the field.

The development of strategies to reduce methane emissions is a relevant area of research, as emission control is essential to fulfilling international environmental obligations. J. Dooley *et al*. (2024) proposed methods for optimising the operation of oil and gas infrastructure, including methane capture and processing technologies. The authors emphasised the importance of integrating satellite data with ground-based measurements for rapid response to emergency leaks. However, their approach did not address the economic aspects of implementing such technologies, which is an important factor for their practical application.

Thus, automated methane emission monitoring systems based on satellite data and radiative transfer models can improve the accuracy of emission estimation and identify key factors affecting their dynamics. The study aimed to assess the accuracy of such systems in the oil and gas fields of Azerbaijan, considering the climatic and technological peculiarities of the region. The objectives of the study included analysis of Sentinel-5P and GHGSat satellite data, application of MODTRAN and SCIATRAN radiative transfer models to correct measurements, and validation of the results using ground-based measurements.

MATERIALS AND METHODS

The study was conducted in 2024 and covered oil and gas fields in Azerbaijan, including the Absheron Peninsula, Shah Deniz, Azeri-Chirag-Guneshli and Umid-Babek fields. This region was selected due to the high concentration of oil and gas production activities, historical cases of significant methane emissions, and the strategic importance of these facilities in the context of the country's energy security. Additionally, the study covered the climatic features of the region, such as the presence of the Caspian Sea wind currents, which facilitate the transport of gas emissions, requiring comprehensive modelling of their distribution.

To monitor methane emissions, satellite data obtained from the Sentinel-5P and GHGSat platforms (Canada) was used, which provide high spatial resolution (up to 50 m) and sensitivity to low methane concentrations in the atmosphere. The data were analysed using a radiative transfer model that quantified the spread of methane emissions and their contribution to the regional dynamics of air pollution. Satellite data from TROPOMI spectrometers (Sentinel-5P) and GHGSat hyperspectral sensors were collected from January to December 2024. Primary processing included data filtering to exclude cloud cover, atmospheric pollution and other factors that could distort measurements. For this purpose, meteorological data provided by the European Centre for Medium-Range Weather Forecasts was used, which contained information on wind speed and direction, humidity and air temperature. Several radiative transfer models were used to quantify the emissions and their distribution, including MODTRAN (Spectral Sciences Inc., USA) and SCIATRAN (German Aerospace Centre, DLR). These models were used to calculate the interaction of electromagnetic radiation with the gas components of the atmosphere and to correct satellite measurements for solar scattering. Additionally, the WRF-Chem inverse model (National Centre for Atmospheric Research, USA) was used to compare satellite data with real measurements and to model the spatial and temporal dynamics of methane emissions. To confirm the data obtained, ground-based methane concentration measurements conducted by automated atmospheric monitoring stations equipped with Los Gatos Research (LGR) laser spectrometers were conducted. Additionally, mobile ground-based monitoring systems equipped with Picarro G2401 gas analysers were used to conduct measurements near oil and gas facilities and in areas with high emission concentrations. Data validation was performed by comparing satellite and ground measurements using the coefficient of determination (R^2) and RMSE analysis.

Furthermore, machine learning methods were used to improve the accuracy of emissions modelling. In particular, Random Forest and Gradient Boosting (XGBoost) algorithms were used, trained on satellite measurements and ground-based monitoring stations. The models included key meteorological and technological parameters, such as wind speed, air temperature, humidity, atmospheric pressure, and oil and gas production intensity, as input variables. These models improved the interpretation of satellite images and incorporated non-linear relationships between these factors and methane concentration.

The analysis of the spatial distribution of emissions was conducted using the geostatistical method, which is an interpolation by the kriging method. To ensure adequate scaling of the calculations, spatial interpolation was performed with a grid spacing of 500 m, which provided a detailed visualisation of the dynamics of methane emissions and an assessment of their relationship with local pollution sources. The SPSS Statistics software package (IBM, USA) was used for statistical data processing. Mean values, standard deviations and confidence intervals were calculated. Analysis of variance (ANOVA) and Mann-Whitney U test were used to identify statistically significant differences in methane concentration between different areas. Correlation analysis was conducted to identify the dependence of methane concentration on weather conditions and oil and gas production intensity.

RESULTS

Satellite monitoring of methane emissions at oil and gas fields in Azerbaijan in 2024 was carried out using data from two satellite platforms: Sentinel-5P (European Space Agency, ESA) and GHGSat (Canada) (Table 1). The Sentinel-5P satellite is equipped with the TROPOMI spectrometer, which provides a wide coverage area and sensitivity to methane at the level of ≈1.8 ppb (parts per billion) at a spatial resolution of 7 × 5.5 km. This recorded background methane concentrations at the regional level. At the same time, the GHGSat satellites are equipped with hyperspectral sensors that provide a spatial resolution of up to 50 m, which enables detailed analysis of local emission sources, including specific oil and gas production facilities.

Table 1. Main parameters of Sentinel-5P and GHGSat satellites)

Parameter	Sentinel-5P (TROPOMI)	GHGSat (hyperspectral sensor)
Spatial resolution	7 × 5.5 km	Up to 50 m
Sensitivity to methane	≈1.8 ppb	≈10-20 ppb
Territory coverage per day	2600 km in width	Up to 12 km^2 per picture
Main objective	Global monitoring	Localised emissions determination

Source: compiled by the authors

Comparison of Sentinel-5P and GHGSat data with field measurements demonstrated a difference in the accuracy of methane concentrations and emissions location (Table 2). GHGSat provided a better match to the ground measurements due to its high spatial resolution, which recorded emissions at the level of individual oil and gas facilities. At the same time, Sentinel-5P data demonstrated a high degree of discrepancy with field measurements at the local level but showed stability in global monitoring of the overall emissions dynamics.

Table 2. Comparison of the accuracy of satellite data with ground measurements

Parameter	Sentinel-5P (TROPOMI)	GHGSat (hyperspectral sensor)	Ground measurements
Average methane concentration in the vicinity of the fields (ppm)	12-15	18-21	18-22
RMSE, ppb	12.4	4.7	–
Determination coefficient (R^2)	0.72	0.89	1
Spatial resolution	7 × 5.5 km	Up to 50 m	Precision measurements
Exposure to climatic factors	High	Average	Low

Source: compiled by the authors

Field measurements revealed higher methane concentrations in the vicinity of oil and gas platforms compared to Sentinel-5P satellite observations. At the Azeri-Chirag-Guneshli field, mobile ground gas analysers demonstrated average methane concentrations in the vicinity of the installations in the range of 18-22 ppm, while Sentinel-5P satellite data showed average values of 12-15 ppm for the same area. In contrast, the GHGSat satellite recorded concentrations close to field measurements with a deviation of less than 10%, indicating its higher accuracy in detecting localised emissions.

The analysis revealed that the main discrepancies between satellite and ground measurements were observed in areas with strong wind currents and high air humidity. These meteorological factors contributed to the rapid dispersion of methane, resulting in the underestimation of concentrations recorded by Sentinel-5P. In addition, the relatively large spatial resolution of Sentinel-5P (7 × 5 km) contributed to the averaging of values, reducing the accuracy of emission source localisation. In contrast, GHGSat, possessing significantly higher spatial resolution (≈25 m) could detect local methane emissions even before their significant dispersion in the atmosphere, which provided a more accurate correspondence with ground data.

To quantify the accuracy of satellite observations, the coefficient of determination (R^2) between satellite and field measurements was calculated. For GHGSat, it was 0.89, indicating a high degree of correlation with ground data, while for Sentinel-5P this figure reached only 0.72, indicating a greater variability of data. Additionally, the RMSE was calculated, which for Sentinel-5P was 12.4 ppb, reflecting significant differences with field measurements. At the same time, for GHGSat, this figure was much lower – 4.7 ppb, which confirmed its higher accuracy in assessing methane concentrations. Thus, the analysis demonstrated that automated GHGSat satellite systems provide more accurate monitoring of local methane emissions, while Sentinel-5P provides a less detailed but stable picture of background air pollution. However, the identified deviations of satellite data from ground-based measurements indicate the need to correct them using radiative transfer models, which will improve the accuracy of methane concentrations and incorporate the impact of climatic factors.

Correction of Sentinel-5P and GHGSat satellite data using MODTRAN and SCIATRAN radiative transfer models improved the accuracy of measurements by considering the interaction of solar radiation with atmospheric gases (Table 3). The analysis demonstrated that without the use of these models, Sentinel-5P data systematically underestimated methane concentrations in conditions of high humidity and cloud cover. This was particularly noticeable in the Apsheron Peninsula and fields located near the Caspian Sea, where the coefficient of determination (R^2) between satellite and ground measurements before correction was 0.72, and after correction using SCIATRAN was 0.83.

Table 3. Influence of radiative transfer models on the accuracy of satellite measurements

Parameter	Average discrepancy with ground measurements (%)	RMSE, ppb	Determination coefficient (R^2)	Efficiency in high humidity conditions	Efficiency in areas with heterogeneous ground conditions
Sentinel-5P (before correction)	15	12.4	0.72	Low	Low
Sentinel-5P (MODTRAN)	9	8.1	0.81	High	Average
Sentinel-5P (SCIATRAN)	7	7.8	0.83	Average	High
GHGSat (before correction)	8	6.3	0.85	Average	High
GHGSat (MODTRAN)	5	4.9	0.89	High	High
GHGSat (SCIATRAN)	3	4.2	0.92	High	High

Source: compiled by the authors

MODTRAN and SCIATRAN radiative transfer models were used to correct satellite measurements of methane concentration, incorporating complex atmospheric processes such as scattering and absorption of solar radiation, as well as the influence of cloud cover and the underlying surface. As part of the study, a detailed calibration of

Sentinel-5P data was conducted, including the correction of atmospheric and climatic factors, which improved the accuracy of the data interpretation. The MODTRAN model incorporated the influence of molecular scattering and absorption in the atmosphere, which is especially important for regions with high humidity and dense cloud cover. As a result of applying this model, the adjusted estimates of methane concentration in such areas increased by an average of 9-12%, which ensured better matching of satellite data with ground measurements. This confirms the importance of the humidity and optical properties of the atmosphere when interpreting satellite data.

SCIATRAN, in turn, demonstrated higher efficiency in complex terrain and heterogeneous underlying surfaces. Its algorithms improved the analysis of the interaction of solar radiation with atmospheric aerosols and surface albedo, which is relevant for areas with rough terrain, such as the Apsheron Peninsula, as well as for offshore fields, where the influence of reflected signals from the water surface makes it difficult to accurately interpret data. The combination of MODTRAN and SCIATRAN data significantly reduced the RMSE of satellite measurements of methane concentration. For MODTRAN, the RMSE decreased from 12.4 ppb to 8.1 ppb, and for SCIATRAN – to 7.8 ppb, which indicates a significant increase in measurement accuracy. The greatest correction effect was observed in regions with high cloud cover, where the original satellite data showed the most significant discrepancies with ground measurements.

One example of a successful correction was the Umid-Babek field. Before the application of the MODTRAN and SCIATRAN models, methane concentrations recorded by Sentinel-5P were on average 15% lower than those from ground monitoring stations. After the correction, the discrepancy was reduced to 9% for MODTRAN and 7% for SCIATRAN, which provided a more reliable picture of the emissions distribution. Thus, the use of radiative transfer models has demonstrated high efficiency in improving the accuracy of satellite monitoring of methane emissions, especially in difficult climatic and geographical conditions. MODTRAN proved to be the most appropriate for regions with high humidity and dense cloud cover, while SCIATRAN provided better results in areas with heterogeneous underlying surfaces, such as offshore fields and regions with complex terrain. An analysis of the spatial distribution of methane emissions based on GHGSat and Sentinel-5P satellite data after correction with radiative transfer models identified four main areas with the highest concentrations of emissions (Table 4). The highest levels of methane were recorded in areas of intensive oil and gas production, indicating that these emissions are linked to anthropogenic sources, including gas leaks, flaring operations and infrastructure depressurisation.

Table 4. Geographical distribution of methane emissions at oil and gas fields in Azerbaijan

Region	Average methane concentration (ppm)	Maximum concentration (ppm)	Main sources of emissions
Apsheron Peninsula	2.8	3.2	Wells, gas processing, pipeline leaks
Azeri-Chirag-Guneshli	2.5	3.1	Platforms, flaring, process leaks
Shah Deniz	2.3-2.6	2.9	Compressor stations, gas pipelines
Umid Babek	2.1-2.4	2.7	Gas production, pipeline depressurisation

Source: compiled by the authors

The most pronounced pollution hotspot was recorded on the Apsheron Peninsula, where the highest density of oil and gas infrastructure is concentrated. Satellite and ground-based measurements showed that the average methane concentration in this region reached 2.8 ppm, which was 35% higher than the background values recorded in neighbouring areas. An analysis of the spatial distribution of emissions revealed significant local anomalies due to the high concentration of pollution sources. The main factors contributing to the elevated methane levels were the active exploitation of oil and gas fields, pipeline leaks, hydrocarbon storage and transportation processes, and emissions from oil refineries.

In addition to onshore sources, significant methane concentrations were detected in the offshore area of the Azeri-Chirag-Guneshli field, where the average emission level was 2.5 ppm, with peak values of up to 3.1 ppm recorded near certain platforms. The main sources of pollution in the area are process leaks arising from hydrocarbon production and transportation, as well as gas flaring. The incomplete combustion of hydrocarbons resulted in additional methane emissions, which increased the overall level of pollution in the region. Additionally, the influence of meteorological factors, such as wind direction and speed, played a significant role in forming the spatial distribution of concentrations. The analysis of atmospheric circulation showed that under low wind conditions, local emissions accumulated in the surface layer of the atmosphere, leading to an increase in methane concentrations. On the contrary, when the wind increased, more active dispersion of pollutants occurred, which reduced concentrations in coastal areas but resulted in their spread to more remote areas.

At the Shah Deniz field, methane concentrations ranged from 2.3 to 2.6 ppm, which also indicates a significant contribution to overall emissions. The main sources of pollution here are compressor stations and gas pipelines used to transport the gas produced. Analysis of GHGSat data confirmed that the highest emissions were observed in the vicinity of gas processing facilities and at pipeline junctions, where depressurisation and leakages are frequent. At the Umid-Babek field, average methane concentrations ranged from 2.1-2.4 ppm, with maximum levels reaching

2.7 ppm. This area has more stable emissions dynamics due to the nature of gas production and lower flaring intensity. The main sources of emissions here are pipeline leaks and process emissions from production platforms.

A comparison of the data obtained demonstrated that methane concentrations in offshore fields (Shah Deniz and Azeri-Chirag-Guneshli) were lower than in onshore areas (Apsheron Peninsula). This is determined by the peculiarities of airflow circulation over the Caspian Sea, which contributed to the accelerated dispersion of methane. In contrast, there was a tendency for emissions to accumulate in coastal areas due to slower wind flows and peculiar atmospheric processes. Thus, the analysis of the geographical distribution of emissions confirmed that the Apsheron Peninsula is the largest contributor to methane pollution, with the highest density of oil and gas infrastructure and the largest number of potential emission sources. Offshore fields also demonstrate high levels of pollution, but the influence of wind flows leads to faster dispersion of methane, reducing its local concentration in the atmosphere. An analysis of the seasonal dynamics of methane emissions based on GHGSat and Sentinel-5P satellite data after correction by radiative transfer models showed a clear dependence of methane concentrations on the climatic conditions of the region (Table 5). During 2024, significant fluctuations in emissions levels were observed due to changes in temperature, humidity and wind flows typical of the coastal areas of the Caspian Sea.

Table 5. Seasonal changes in methane concentration at oil and gas fields in Azerbaijan

Season	Average methane concentration (ppm)	Main reasons for changes in concentrations
Winter (December-February)	3 (max)	Slower dispersion increased flaring
Spring (March-May)	2.4	High wind speeds, improved dispersion
Summer (June-August)	2.7	Increased production intensity, increased leakage
Autumn (September-November)	2.3	High wind activity, reduced production

Source: compiled by the authors

The highest methane concentrations were recorded in the winter (December-February) and summer (June-August) periods, while emission levels were lower in the off-season (spring and autumn). During the winter months, average methane concentrations on the Absheron Peninsula reached 3 ppm, which is 15-20% higher than in summer. This increase is attributed to a combination of factors: increased flaring due to increased demand for heating, as well as slower atmospheric convection, which reduces the rate of emissions dissipation. In addition, low temperatures contribute to the condensation of moisture in the air, which creates conditions for the retention of methane in the surface layer of the atmosphere.

The summer peak in methane concentrations observed at offshore fields such as Azeri-Chirag-Guneshli and Shah Deniz is determined by the increased intensity of drilling and gas production during this period. Average methane concentrations at these fields reached 2.7 ppm in the summer, which is about 12% higher than in spring and autumn. The main factor influencing this increase is the increase in process leakage during production stimulation.

In spring and autumn, methane concentrations were 10-15% lower than in winter and summer. This is determined by the fact that wind activity over the Caspian Sea reaches its maximum in spring and autumn, which contributes to the intensive dispersion of methane emissions. For instance, in April and October, average concentrations at the Shah Deniz field dropped to 2.2 ppm, which is 0.4-0.5 ppm lower than in the summer months.

Thus, the seasonal dynamics of methane emissions at Azerbaijan's oil and gas fields are determined by a combination of natural and climatic factors and the specifics of hydrocarbon production in different periods of the year. Winter and summer periods are characterised by the highest methane concentrations, with the main factor being a decrease in the rate of gas dissipation in winter and intensified gas production and leakage in summer. In contrast, spring and autumn show relatively low methane concentrations due to high air velocities that help disperse pollutants.

The use of machine learning methods to analyse satellite data on methane emissions has significantly improved the accuracy of identifying pollution sources and forecasting gas concentrations (Table 6). The use of Random Forest and XGBoost algorithms trained on a combination of Sentinel-5P and GHGSat satellite data, ground measurements and meteorological parameters identified hidden patterns between methane emissions and the region's climatic conditions.

Table 6. Comparison of the accuracy of machine learning algorithms in predicting methane emissions

Parameter	Random forest	XGBoost
RMSE, ppb	6.1	4.2
Determination coefficient (R^2)	0.82	0.91
Sensitivity to non-linear dependencies	Average	High
Model training time (min)	12	18

Source: compiled by the authors based on H. Salman *et al*. (2024)

Comparative analysis of the models demonstrated that XGBoost features the highest accuracy of emissions forecasting compared to Random Forest. The RMSE for XGBoost was 4.2 ppb, which is significantly lower than that of Random Forest (6.1 ppb), and the coefficient of determination (R^2) reached 0.91, indicating a high degree of agreement between predictions and actual measurements. This is because XGBoost is more efficient at handling non-linear relationships between methane emissions and meteorological factors such as wind speed and direction, temperature, humidity and atmospheric pressure.

An additional analysis of the importance of variables in predicting methane emissions identified the key factors that have the greatest impact on atmospheric concentrations. The largest contribution to the model's predictions was made by wind speed (28%), which determines how quickly methane dissipates in the atmosphere. Air temperature (22%) was the second most important factor, affecting convection processes and methane retention in the surface layer. Gas production intensity (19%) was also of high importance, as an increase in production volumes is usually accompanied by an increase in process leaks and flared gas volumes. Air humidity (16%) was substantial, as methane can remain in the lower atmosphere for longer in high humidity conditions. Atmospheric pressure (15%) also significantly affected methane distribution, affecting its vertical migration in the atmosphere.

The analysis showed that XGBoost was more resistant to noise in the data and better able to cope with non-linear dependencies, which is especially important when forecasting methane emissions in the challenging meteorological conditions of the Caspian region. However, the training time of this model was 18 minutes, which is higher than that of Random Forest (12 minutes), making XGBoost more resource-intensive.

Thus, the use of machine learning has improved the accuracy of satellite monitoring of methane emissions, identified key climatic and technological factors affecting pollution levels, and improved the predictability of gas concentrations in the atmosphere. The XGBoost algorithm has demonstrated the best results, which confirms the prospects of its use in automated methane emission monitoring systems in oil and gas fields.

DISCUSSION

The study confirmed the effectiveness of satellite monitoring of methane emissions at Azerbaijan's oil and gas fields, revealing significant differences in the accuracy and resolution of data obtained from the Sentinel-5P and GHGSat platforms. Sentinel-5P data demonstrated high stability in global monitoring, but its accuracy at the local level was lower than that of GHGSat. This correlated with the results of O. Schneising *et al*. (2023), which also noted that large-scale satellite platforms demonstrate high accuracy in assessing general pollution trends but are inferior in local emission detection. At the same time, J. Churchill *et al*. (2023) highlighted the possibility of combining data from different satellite platforms to obtain more accurate results. In this study, this is confirmed by comparison with ground-based measurements, which indicates the prospects of integrating Sentinel-5P and GHGSat data.

Additional analysis revealed that the differences between the platforms are due not only to spatial resolution but also to the sensitivity of the sensors to methane concentrations. The study showed that GHGSat can record emissions with high accuracy due to its hyperspectral sensors, which was also previously noted by J. Churchill *et al*. (2022). However, according to B. Rouet-Leduc *et al*. (2023), Sentinel-5P may be more effective in detecting global pollution trends, despite its limited ability to detect local leaks. The study determined that Sentinel-5P data requires additional correction to improve its accuracy.

Comparison of satellite data with ground measurements revealed systematic discrepancies, especially in conditions of high humidity and strong winds. Similar results were obtained by M. Komarudin *et al*. (2022), noting that climatic factors significantly affect the accuracy of remote sensing. However, G. Domènech-Gil *et al*. (2024) highlighted that satellite measurements can be corrected using climate models. In the present study, the use of MODTRAN and SCIATRAN radiative transfer models improved the accuracy of satellite data, which confirms the possibility of adapting correction methods.

The results demonstrated that the use of radiative transfer models significantly improves the accuracy of satellite monitoring of methane emissions. The use of SCIATRAN proved to be more effective in areas with a heterogeneous underlying surface, which is confirmed by A.A. Bloom *et al*. (2021) in an analysis of the impact of albedo on the accuracy of satellite measurements. At the same time, W. Collins *et al*. (2022) argue that SCIATRAN's effectiveness is limited in regions with high humidity. However, this study determined that SCIATRAN successfully corrected Sentinel-5P data, providing more accurate measurements in difficult climatic conditions.

The analysis of the spatial distribution of methane emissions confirmed that the highest concentrations were recorded in areas with a high density of oil and gas infrastructure. This is consistent with the findings of A. Andrews *et al*. (2023), noted the connection between the intensity of hydrocarbon production and the level of air pollution. At the same time, A. Groshenry *et al*. (2024) indicated that gas transportation has a significant impact on emissions. The study determined that pipeline leaks do indeed make a significant contribution to pollution, but the main source of emissions remains gas flaring.

Seasonal changes in methane concentrations have shown that emission levels reach their maximum values in winter and summer, while in spring and autumn, they decrease (Ismanzhanov *et al*., 2012). These data confirm the results of L. Gushungo *et al*. (2022), noting a substantial impact of climatic conditions on methane concentrations. However, D. Varon *et al*. (2022) argued that seasonal fluctuations are less pronounced in regions with stable meteorological conditions. The present study determined that meteorological factors are substantial in the seasonal dynamics of emissions, which emphasises the need to incorporate them in the interpretation of satellite data.

The use of machine learning methods has significantly improved the accuracy of satellite data analysis. The XGBoost algorithm demonstrated the best results compared to Random Forest, which is consistent with the findings of A. Ferrari *et al*. (2024), also noted the high efficiency of XGBoost in environmental modelling tasks. At the same time, W. Daniels *et al*. (2024) highlighted that XGBoost requires significant computing resources and is inferior to Random Forest in terms of data processing speed. However, this study confirmed that XGBoost has higher accuracy despite its computational complexity.

Additionally, the importance of variables affecting methane concentrations was analysed. Wind speed, air temperature and production intensity were found to be key factors, which is consistent with the findings of B. Erland *et al*. (2022), which also noted the importance of these parameters. An increase in wind speed promotes faster methane dissipation, reducing its local concentration, while weak winds can contribute to its accumulation near emission sources. Air temperature is also substantial: during cold periods, methane remains in the surface layer of the atmosphere for longer, while at high temperatures it rises faster to the upper layers, affecting its detection by satellites.

J. Wang *et al*. (2022) and L. Shen *et al*. (2022) noted that air humidity has different effects depending on the region. In high humidity conditions, water vapour particles can absorb and scatter infrared radiation, which affects the accuracy of satellite measurements. At the same time, in arid regions, more stable atmospheric conditions contribute to better detection of emission sources (Golinko & Nedosnovanyi, 2024). L. Thompson & M. Beck (2024) and N. Isnaini *et al*. (2024) emphasised the importance of atmospheric pressure and solar radiation, which were also considered in this study. High atmospheric pressure can increase air density, which decreases the vertical movement of methane, while solar radiation promotes photochemical processes, affecting its concentration. In addition, S. Chen *et al*. (2023) noted that local emission sources can have a disproportionate impact on regional methane concentrations. For instance, leaks from individual oil and gas installations can cause significant local anomalies that exceed the background pollution level by several times (Doroshenko *et al*., 2023). This highlights the need to combine satellite observations with ground-based measurements for more accurate emissions mapping.

Additionally, the analysis of the spatial distribution of emissions, incorporating technological and climatic factors, demonstrated that the highest concentrations are observed in areas of active oil and gas production. This is confirmed by the results of M. Gålfalk *et al*. (2024), noting that the production infrastructure plays a key role in the formation of local pollution. Significant concentrations of methane are observed near drilling rigs, compressor stations and transport hubs, where leaks occur most frequently. Cases of uncontrolled emissions from damaged pipelines were also identified, which underscores the importance of regular monitoring and maintenance of equipment.

M. Watine-Guiu *et al*. (2023) indicated that monitoring efficiency can be improved by improving data interpretation methods, which is also consistent with the findings of this study. In particular, the use of combined analysis of satellite and ground-based measurements, as well as the application of machine learning algorithms for automated emission source detection, can improve the accuracy of methane detection. In addition, the integration of data from different satellite platforms, such as Sentinel-5P and GHGSat, can incorporate differences in spatial resolution and sensor sensitivity, which makes monitoring more accurate and comprehensive.

Thus, the study confirmed that satellite monitoring is efficient for controlling methane emissions in oil and gas fields. The application of radiation transfer models and machine learning methods has significantly improved the accuracy of satellite measurements. Integration of data from various satellite platforms and the use of modern data analysis algorithms provide a comprehensive approach to emissions monitoring, covering the influence of climatic factors and technological processes. The results confirm the prospects for further development of satellite-based observation methods and their integration with ground-based measurements to create more accurate environmental monitoring systems.

CONCLUSIONS

The study conducted satellite monitoring of methane emissions at oil and gas fields in Azerbaijan in 2024 using Sentinel-5P and GHGSat data. The analysis demonstrated that the GHGSat satellite, due to its high spatial resolution (up to 50 m), provided accurate localisation of emission sources, while Sentinel-5P (7 × 5.5 km resolution) demonstrated high stability in monitoring background methane concentrations. Comparison of satellite data with ground measurements revealed significant differences in the accuracy of methane concentrations. GHGSat showed the best agreement with the ground data ($R^2 = 0.89$, RMSE = 4.7 ppb), while Sentinel-5P showed significant deviations ($R^2 = 0.72$, RMSE = 12.4 ppb), especially in conditions of high humidity and strong winds. The main differences were due to the averaging of values due to the low spatial resolution of Sentinel-5P and the influence of climatic factors.

The use of MODTRAN and SCIATRAN radiative transfer models improved the accuracy of satellite measurements. Correction of Sentinel-5P data using SCIATRAN increased the coefficient of determination to 0.83 and reduced the RMSE by 37%, which demonstrates the importance of considering atmospheric and climatic factors when analysing satellite observations. The geographical analysis identified four areas with the highest methane emissions: Apsheron Peninsula, Azeri-Chirag-Guneshli, Shah Deniz and Umid Babek. The highest concentrations were recorded near oil and gas facilities, which confirms the anthropogenic nature of the emissions. The highest pollution was recorded on the Apsheron Peninsula (up to 3.2 ppm), which is determined by the high density of oil and gas infrastructure and technological leaks.

The seasonal analysis demonstrated that the maximum methane concentrations were observed in winter (3 ppm) and summer (2.7 ppm), while in spring and autumn, emissions decreased by 10-15% due to more active methane dissipation. In winter, the increase in concentrations is explained by the low rate of dissipation and the increase in gas flaring. The use of machine learning methods (Random Forest and XGBoost) improved the accuracy of methane emissions forecasting. The XGBoost algorithm demonstrated the best results ($R^2 = 0.91$, RMSE = 4.2 ppb) and identified key factors affecting methane concentrations, including wind speed, air temperature and gas production intensity. Further research could focus on the integration of satellite and drone technologies to improve the accuracy of local methane emissions monitoring.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Aghayeva, K., & Krauklit, G. (2024). Innovative software for analysing satellite data and methane emissions using radiative transfer model. *Bulletin of Cherkasy State Technological University*, 29(3), 65-76. doi: 10.62660/bcstu/3.2024.65.

[2] Andrews, A., Boucher, C., Chakrabarti, A., Dauphin, M., Doshi, M., Rashid, K., Speck, A., Van Pelt, A., Yuan, J., & Zielinski, L. (2023). Quantitative mapping of methane emissions in oil & gas facilities. In *Proceedings of the SPE annual technical conference and exhibition*. San Antonio: SPE. doi: 10.2118/214909-ms.

[3] Andrews, B., Chakrabarti, A., Dauphin, M., & Speck, A. (2023). Application of machine learning for calibrating gas sensors for methane emissions monitoring. *Sensors*, 23(24), article number 9898. doi: 10.3390/s23249898.

[4] Bekirova, L.R., & Bunyatova, E.M. (2024). Intelligent ecological monitoring system for remote assessment of surface condition in processes of petrochemistry and oil refining. *Processes of Petrochemistry and Oil Refining*, 25(1), 284-300. doi: 10.62972/1726-4685.2024.1.284.

[5] Bloom, A.A., Bowman, K.W., Lee, M., Turner, A.J., Schroeder, R., Worden, J.R., Weidner, R.J., McDonald, K.C., & Jacob, D.J. (2021). CMS: Global 0.5-deg wetland methane emissions and uncertainty (WetCHARTs v1.3.1). doi: 10.3334/ORNLDAAC/1915.

[6] Chen, Q., Schissel, C., Kimura, Y., McGaughey, G., McDonald-Buller, E., & Allen, D. (2023). Assessing detection efficiencies for continuous methane emission monitoring systems at oil and gas production sites. *Environmental Science & Technology*, 57(4), 1788-1796. doi: 10.1021/acs.est.2c06990.

[7] Chen, S., Frish, M., & Aubut, N. (2023). Long-open-path fixed continuous methane emission monitor. *Next-Generation Spectroscopic Technologies XV*, 12516, 125160U-1-25160U-11. doi: 10.1117/12.2663633.

[8] Churchill, J., Wise, B., Cooper, I., & Smith, B. (2022). Automated unmanned systems perform safe emissions quantification & reduction in Permian operations. In *Proceedings of the SPE annual technical conference and exhibition* (article number SPE-210173-MS). Houston: SPE. doi: 10.2118/210173-ms.

[9] Churchill, J., Wise, B., Cooper, I., & Smith, B. (2023). Methane emissions quantification and resulting methane emissions reduction in the Permian basin enabled by automated unmanned systems. *SPE Production & Operations*, 38(3), 391-405. doi: 10.2118/210173-pa.

[10] Collins, W., Orbach, R., Bailey, M., Biraud, S., Coddington, I., DiCarlo, D., Peischl, J., Radhakrishnan, A., & Schimel, D. (2022). Monitoring methane emissions from oil and gas operations. *Optics Express*, 30(14), 24326-24351. doi: 10.1364/oe.464421.

[11] Daniels, W., Jia, M., & Hammerling, D. (2024). Detection, localization, and quantification of single-source methane emissions on oil and gas production sites using point-in-space continuous monitoring systems. *Elementa: Science of the Anthropocene*, 12(1), article number 00110. doi: 10.1525/elementa.2023.00110.

[12] Deryaev, A. (2023). Features of forecasting abnormally high reservoir pressures when drilling wells in areas of south-western Turkmenistan. *Innovaciencia*, 11(1), 1-16. doi: 10.15649/2346075X.3605.

[13] Domènech-Gil, G., Duc, N., Wikner, J., Eriksson, J., Påledal, S., Puglisi, D., & Bastviken, D. (2023). Electronic nose for improved environmental methane monitoring. *Environmental Science & Technology*, 58(1), 352-361. doi: 10.1021/acs.est.3c06945.

[14] Domènech-Gil, G., Duc, N., Wikner, J., Eriksson, J., Puglisi, D., & Bastviken, D. (2024). Efficient methane monitoring with low-cost chemical sensors and machine learning. *Proceedings*, 97(1), article number 79. doi: 10.3390/proceedings2024097079.

[15] Dooley, J., Minschwaner, K., Dubey, M., Abbadi, S., Sherwin, E., Meyer, A., Follansbee, E., & Lee, J. (2024). A new aerial approach for quantifying and attributing methane emissions: Implementation and validation. *Atmospheric Measurement Techniques*, 17(17), 5091-5111. doi: 10.5194/amt-17-5091-2024.

[16] Doroshenko, Yu., Liuta, N., & Lisafin, D. (2023). Improving the standardisation of oil product losses. *Prospecting and Development of Oil and Gas Fields*, 23(3), 24-33. doi: 10.69628/pdogf/3.2023.24.

[17] Erland, B., Thorpe, A., & Gamon, J. (2022). Recent advances toward transparent methane emissions monitoring: A review. *Environmental Science & Technology*, 56(23), 16567-16581. doi: 10.1021/acs.est.2c02136.

[18] Ferrari, A., Laneve, G., Pampanoni, V., Carvajal, A., & Rossi, F. (2024). Monitoring methane emissions from landfills using Prisma imagery. In *Proceedings of the IEEE international geoscience and remote sensing symposium* (pp. 3663-3667). Athens: IEEE. doi: 10.1109/IGARSS53475.2024.10642079.

[19] Gålfalk, M., Påledal, S., Yngvesson, J., & Bastviken, D. (2024). Measurements of methane emissions from a biofertilizer storage tank using ground-based hyperspectral imaging and flux chambers. *Environmental Science & Technology*, 58(8), 3766-3775. doi: 10.1021/acs.est.3c06810.

[20] Golinko, V., & Nedosnovanyi, O. (2024). Improvement of automation of geoinformation data processing using neural network technology. *Technologies and Engineering*, 25(4), 19-28. doi: 10.30857/2786-5371.2024.4.2.

[21] Groshenry, A., Giron, C., Hessel, C., Franchis, C., Facciolo, G., Ehret, T., & Sas, K. (2024). Methane emissions monitoring using geostationary satellites. In *Proceedings of the IEEE international geoscience and remote sensing symposium* (pp. 10376-10380). Athens: IEEE. doi: 10.1109/IGARSS53475.2024.10642856.

[22] Gushungo, L., Mupfumira, P., & Von Kallon, D. (2022). Development of an automated coal mine methane extraction system. In *Proceedings of the international conference on industrial engineering and operations management* (pp. 263-270). Nsukka: IEOM Society International. doi: 10.46254/af03.20220038.

[23] Ismanzhanov, A.I., Murzakulov, N.A., & Azimzhanov, O.A. (2012). Investigation on heat exchange in interlayer space of multilayer greenhouses. *Applied Solar Energy*, 48(2), 118-120. doi: 10.3103/S0003701X12020107.

[24] Isnaini, N., Muntini, M., & Sugriwan, I. (2024). Website-based monitoring system for methane (CH_4) concentration, humidity, and temperature. *Journal of Physics: Conference Series*, 2780, article number 012035. doi: 10.1088/1742-6596/2780/1/012035.

[25] Komarudin, M., Septama, H., & Yulianti, T. (2022). Autonomous cyber physical systems for monitoring of methane gas in rice field. In *Proceedings of the 2nd international conference on smart and innovative agriculture* (pp. 138-143). London: Atlantis Press. doi: 10.2991/absr.k.220305.020.

[26] Rouet-Leduc, B., Kerdreux, T., Tuel, A., & Hulbert, C. (2023). Autonomous detection of methane emissions in multispectral satellite data using deep learning. *ARXIV*. doi: 10.48550/arXiv.2308.11003.

[27] Rusho, M.A., Azizova, R., Mykhalevskiy, D., Karyonov, M., & Hasanova, H. (2024). Advanced earthquake prediction: Unifying networks, algorithms, and attention-driven LSTM modelling. *International Journal of GEOMATE*, 27(119), 135-142. doi: 10.21660/2024.119.m2415.

[28] Salman, H., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024, 69-79. doi: 10.58496/BJML/2024/007.

[29] Satin, I., & Kutsyi, D. (2024). Greenhouse gas emissions from municipal waste disposal sites and measures for its reduction. *Scientific Reports of the National University of Life and Environmental Sciences of Ukraine*, 20(3). doi: 10.31548/dopovidi.3(109).2024.003.

[30] Schneising, O., Buchwitz, M., Hachmeister, J., Vanselow, S., Reuter, M., Buschmann, M., Bovensmann, H., & Burrows, J.P. (2023). Advances in retrieving XCH_4 and XCO from Sentinel-5 Precursor: Improvements in the scientific TROPOMI/WFMD algorithm. *Atmospheric Measurement Techniques*, 16(2), 669-694. doi: 10.5194/amt-16-669-2023.

[31] Schuit, B., *et al.* (2023). Automated detection and monitoring of methane super-emitters using satellite data. *Atmospheric Chemistry and Physics*, 23(16), 9071-9098. doi: 10.5194/acp-23-9071-2023.

[32] Shen, L., *et al.* (2022). Satellite quantification of oil and natural gas methane emissions in the US and Canada including contributions from individual basins. *Atmospheric Chemistry and Physics*, 22(17), 11203-11215. doi: 10.5194/acp-22-11203-2022.

[33] Tanzharikov, A., Dairabayev, M., Kim, A., Sosa, J., & Bissembay, A. (2023). Real-time greenhouse gas emissions visibility in Tengiz production operations. In *Proceedings of the SPE caspian technical conference and exhibition* (SPE-217553-MS). Baku: IADC. doi: 10.2118/217553-MS.

[34] Thompson, L., & Beck, M. (2024). Use of new technology to measure methane emissions. *Journal of Animal Science*, 102(2), 220-221. doi: 10.1093/jas/skae102.250.

[35] Varon, D., *et al*. (2022). Integrated Methane Inversion (IMI 1.0): A user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations. *Geoscientific Model Development*, 15(14), 5787-5805. doi: 10.5194/gmd-15-5787-2022.

[36] Wang, J., Daniels, W., Hammerling, D., Harrison, M., Burmaster, K., George, F., & Ravikumar, A. (2022). Multiscale methane measurements at oil and gas facilities reveal necessary frameworks for improved emissions accounting. *Environmental Science & Technology*, 56(20), 14743-14752. doi: 10.1021/acs.est.2c06211.

[37] Watine-Guiu, M., Varon, D., Irakulis-Loitxate, I., Balasus, N., & Jacob, D. (2023). Geostationary satellite observations of extreme and transient methane emissions from oil and gas infrastructure. *Proceedings of the National Academy of Sciences of the United States of America*, 120(52), article number e2310797120. doi: 10.1073/pnas.2310797120.

Камала Агаєва
Доктор філософії
Азербайджанський державний університет нафти та промисловості
AZ1010, просп. Азадлик, 34, м. Баку, Азербайджан
https://orcid.org/0009-0006-6779-2550

Герман Краукліт
Аспірант
Азербайджанське національне аерокосмічне агентство
AZ1115, вул. Ахундова, 1, м. Баку, Азербайджан
https://orcid.org/0000-0002-6171-9348

Автоматизовані системи для моніторингу викидів метану за допомогою супутникових даних: аналіз із використанням моделі радіаційного переносу

Анотація. Метою дослідження було оцінювання точності автоматизованих систем моніторингу викидів метану на нафтогазових родовищах Азербайджану з використанням супутникових даних і моделі радіаційного переносу. Методологія включала аналіз супутникових даних Sentinel-5P і GHGSat за 2024 рік, застосування моделей MODTRAN і SCIATRAN для врахування атмосферних чинників і валідацію результатів наземними вимірами з використанням спектрометрів Los Gatos Research і газоаналізаторів Picarro G2401. Результати показали, що GHGSat точніше фіксував локальні викиди ($R^2 = 0{,}89$, RMSE = 4,7 ppb), тоді як Sentinel-5P продемонстрував заниження концентрацій за високої вологості ($R^2 = 0{,}72$, RMSE = 12,4 ppb). Корекція даних із застосуванням моделей MODTRAN і SCIATRAN дала змогу підвищити точність вимірювань: RMSE знизилося до 7,8 ppb для Sentinel-5P і до 4,2 ppb для GHGSat. Найбільші викиди метану виявлено на Апшеронському півострові (2,8 ppm), що пов'язано з витоками і технологічними процесами переробки газу. Сезонний аналіз показав збільшення концентрацій взимку (3 ppm) і влітку (2,7 ppm) внаслідок зниження швидкості розсіювання та інтенсифікації видобутку. Методи машинного навчання (XGBoost, Random Forest) підвищили точність прогнозування ($R^2 = 0{,}91$ для XGBoost), виявивши ключові фактори: швидкість вітру, температуру, інтенсивність видобутку та вологість. Висновки наголошують на необхідності інтеграції супутникових даних із наземними вимірами та моделями радіаційного переносу для підвищення точності моніторингу та розроблення стратегій зниження викидів

Ключові слова: машинне навчання; просторовий розподіл; нафтогазова галузь; наземні вимірювання; кліматичні процеси