UTRECHT UNIVERSITY

INSTITUTE FOR MARINE AND ATMOSPHERIC RESEARCH UTRECHT

CLIMATE PHYSICS

MSC THESIS

Evaluation of methane leakage rates from the oil and gas infrastructure of Romania

Authors: Stavropoulou Foteini 9285520

Supervisors: Prof. Dr. Thomas Röckmann Dr. Daniel Zavala-Araiza

October, 2022



Abstract

Methane (CH_4) emissions into the atmosphere from the oil and natural gas (O&G) sector are a significant source of greenhouse gas emissions, but the magnitude and location of these emissions remain highly uncertain for a number of production regions. This is especially relevant for Romania, which has one of the highest reported annual emissions of CH₄ from the energy sector in Europe and, therefore, plays an important role in reaching the CH₄ emission reduction targets of the European Union. In this study, CH₄ emissions from O&G production sites in the southern, mainly oil-producing part of Romania were investigated on a component and facility level using a combination of various ground-based measurement techniques. On the component scale, an Optical Gas Imaging camera for the detection of individual leaking components and a Hi-Flow Sampler device for the quantification of the emissions were used. On the facility scale, four methods were used to measure CH_4 emissions, namely Gaussian Plume Modelling (GPM), Other Test Method (OTM-33A), Tracer Dispersion Method (TDM) and Mass Balance Method (MBA) using Unmanned Aerial Vehicle (UAV)- based measurements. Emissions derived from measured concentrations ranged from 0.0006 kg CH_4 h⁻¹ to 73 kg CH_4 h⁻¹ for individual oil wells. Derived emissions were characterised by heavily skewed distributions, with 10% of sites accounting for more than 85% of total emissions. Combining the results from all site-level quantification approaches, we derive a mean emission factor of 8.3 kg h^{-1} of CH₄ (3.8 - 19, 95% confidence interval). Comparing our estimated emission factor to those from other O&G production regions in North America, we find that Romania presents one of the highest emission factors and levels of skewness. When our estimated emission factor for the subset of oil wells only is used to scale up to the national scale, the estimate of 2019 total emissions from oil wells is 240 ktons $CH_4 \text{ yr}^{-1}$ (min = 110 ktons yr^{-1} and max = 555 ktons yr^{-1}), approximately 3 times higher than the total 2020 upstream O&G sector reported emissions. A subset of sites was screened with an infrared (IR) camera and the analysis of recorded IR videos suggests that roughly half of surveyed sites had identified emissions, and more than three quarters of the detected emissions from oil wells are vented by facility design. This is in agreement with no reported gas production for the majority of these wells. Our results suggest that previously reported reductions of O&G related CH₄ emissions in Romania, derived from a reduction in production, were over-optimistic. On the other hand, the present O&G production infrastructure in Romania holds a massive mitigation potential. Specifically, implementing regulations for unreported vented emissions serve as key mitigation opportunities for reducing CH_4 emissions in the EU.

Keywords: Methane emissions; Oil and gas sector; Emissions distributions; Ground-based measurements; Romania;

Acknowledgements

First and foremost, I would like to express my gratitude and deep appreciation to my supervisors Prof. Dr. Thomas Röckmann and Dr. Daniel Zavala-Araiza for guiding me throughout my entire research project and for letting me be a part of this incredible project. I would like to thank Dr. Jaroslaw Necki, Dr. Hossein Maazallahi, Katarina Vinković, Danut Dumitru and the rest of the ROMEO team for all the meetings and conversations, and for always being willing to answer my questions and providing me with guidance, feedback and valuable input about the measurements and the datasets. I am very grateful for this amazing experience of working together with such a large team of admirable researchers from multiple worldwide research institutes and for making me feel as if I was present during the ROMEO campaign. I would also like to thank Dr. Guus Velders for taking part in my thesis examination and spending time reading and evaluating my report. Finally, but most importantly, I would like to thank my family, friends and partner for always believing in me and continuously encouraging me during this master thesis research project and throughout my studies.

 $\mathbf{4}$

Contents

1	Intr	oducti	on	5						
	1.1	The ro	ble of methane on climate change	5						
	1.2	Oil an	d gas sector	5						
	1.3	Bottor	n up - top down approaches	6						
	1.4	Romai	nia and ROMEO campaign	7						
	1.5	.5 Research objectives								
2	Mai	torial a	nd methods	a						
-	2 1	Investi	rated area	g						
	2.1	Moasu	rement methods	10						
	2.2	2 2 1	Component scale	10						
		2.2.1	Exclipte scale	10						
	9 3	Dotorr	racinty scale	11						
	2.5	0 2 1	Log normal distributions	14						
		2.3.1	Statistical actimator	14						
		2.3.2 9.2.2	"Non detects" and Detection Limit	10						
		2.3.3	Non-detects and Detection Limit	17						
3	\mathbf{Res}	ults		18						
	3.1	Compo	onent Scale	18						
		3.1.1	Dataset overview	18						
		3.1.2	Oil wells	19						
		3.1.3	Sources of observed emissions	21						
	3.2	Facilit	y-level quantifications	22						
		3.2.1	Dataset overview	22						
		3.2.2	Oil wells	23						
		3.2.3	Emissions distributions and emission factors	25						
		3.2.4	Screenings	27						
4	Dis	cussion		29						
	4.1	Compa	arison with \mathbf{CH}_4 emissions reported from other studies $\ldots \ldots \ldots \ldots \ldots \ldots$	30						
	4.2	Compa	arison with $\mathbf{CH_4}$ emissions reported from national inventories	31						
5	Cor	nclusio	15	34						
Re	efere	nces		35						
C				10						
Su	ipple	ementa	ry Material	40						

Acronyms and Abbreviations

Abbreviations and units	
IPCC	Intergovernmental Panel on Climate Change
CO ₂	Carbon Dioxide
GWP	Global Warming Potential
CH ₄	Methane
H ₂ S	Hydrogen Sulfide
O&G	Oil and Gas
IEA	International Energy Agency
US EPA	United States Environmental Protection Agency
EEA	European Environment Agency
UNFCCC	United Nations Framework Convention on Climate Change
ROMEO	Romanian Methane Emissions from Oil Gas
CCAC	Climate and Clean Air Coalition
OGI	Optical Gas Imaging
HFS	Hi-Flow Sampler
GPM	Gaussian Plume Method
OTM	Other Test Method
TDM	Tracer Dispersion Method
MBA	Mass Balance Approach
UAV	Unmanned Aerial Vehicle
BDL	Below Detection Limit
EF	Emission Factor
pdf	probability density function
kg/h	kilograms per hour
ppm	parts per million
scm	standard cubic meter

1 Introduction

1.1 The role of methane on climate change

Climate change is a major threat to our world and one of the most important topics of public debate [1]. Anthropogenic emissions of greenhouse gasses to the atmosphere are the main cause of global warming. Reducing these emissions is therefore crucial in order to minimise the negative impacts of climate change on people and the environment. To address this issue, a legally binding international treaty, called the Paris Agreement, was adopted by 196 Parties in 2015 [2]. The goal of this agreement is to drastically cut global greenhouse gas emissions to limit global temperature rise to 2 C°, preferably to 1.5 C°, compared to pre-industrial levels [2]. The most recent Intergovernmental Panel on Climate Change (IPCC) report has stressed the importance of rapidly reducing methane (CH₄) emissions in the atmosphere in addition to reducing carbon dioxide (CO₂). They have estimated that a 50% reduction in anthropogenic CH₄ emissions over the next 30 years would lead to a reduction in warming of about 0.20 C° globally, increasing the feasibility of achieving the Paris Agreement goal [1]. However, CH₄ emissions have more than doubled since the pre-industrial era and its concentration continues to rise rapidly mostly due to human-related activities [1].

Methane is the second most abundant and significant greenhouse gas after CO₂, accounting for at least 25% of current global warming [1]. Even though CH₄ has a shorter lifetime in the atmosphere of around 9.1 \mp 0.9 years, compared with centuries for CO₂, CH₄ is more effective at trapping radiation than CO₂ [1]. CH₄ has a Global Warming Potential (GWP) of around 30 times greater compared to CO₂ over a 100-year time horizon and more than 80 over a 20-year time horizon. Additionally, CH₄ contributes to the formation of tropospheric ozone, a potent regional air pollutant that worsens local air quality and causes serious health problems and damage on agricultural yields [3]. Due to its short lifetime, the strong global warming potential and the local negative impacts, reducing CH₄ would lead to substantial climate benefits in the near- and long-term future.

1.2 Oil and gas sector

CH₄ is emitted from a variety of anthropogenic and natural sources. The largest natural sources are from wetlands, wildfires, freshwater and geological processes, while the largest anthropogenic emissions, accounting for 50-65% of total CH₄ emissions, include fossil fuel exploitation, agricultural activities, landfills and wastewater treatment [4]. Approximately one third of global anthropogenic CH₄ emissions come from the fossil fuel-based energy sector, which includes emissions from coal, oil and natural gas extraction and transport, and use of natural gas [5]. Therefore, even though it is crucial to tackle all sources of CH₄ emissions related to human activity, the oil and gas (O&G) sector offers the most cost-effective potential for methane abatement. Considering the current high natural gas prices, the International Energy Agency (IEA) estimates that 75% of emissions reductions from the energy sector can be achieved at no net monetary cost and could even result in economic savings, given that CH₄ is the main component of natural gas and has commercial value [6]. Reducing CH₄ emissions from O&G operations is one of the most substantial, easily accessible, and affordable mitigation actions governments can take to address the climate change issue [5].

CH₄ can be emitted from a variety of sources along the entire O&G supply chain. Emissions are often divided into three categories: vented emissions (i.e. "intended" or "operational" emissions), fugitive emissions (i.e. "leaks" or "unintended" emissions) and incomplete combustion emissions related to gas flaring. Venting is the intentional release of CH_4 into the atmosphere and can occur as part of routine equipment maintenance or normal operations of certain equipment such as gas driven pneumatic controllers. Flaring is the process of burning associated gases released during normal operations or unplanned processes. Both venting and flaring are two commonly used methods for disposal of unwanted or without value gases which are produced during exploration of fields, production of oil and gas, and oil transport and refining. Leaks and other irregular or unplanned releases of gases can occur in the O&G infrastructure from flanges, valves and other malfunctioning equipment [7]. Reducing both fugitive and vented emissions offers a great opportunity for mitigation of total anthropogenic CH_4 emissions to the atmosphere. Depending on the different characteristics and processes related to the site type (oil or gas), different possibilities may exist. However, the total magnitude and location of both fugitive and vented emissions is uncertain, thus, reduction strategies cannot be effectively implemented. Therefore, improving our understanding of CH₄ emissions from the O&G sector requires accurate emissions quantification and a combination of approaches.

1.3 Bottom up - top down approaches

O&G CH₄ emissions can be quantified by using approaches based on ambient CH₄ concentration measurements and models (top-down method) and approaches based on direct measurement of emissions from individual sources or standard emission factors (bottom-up method). Top-down approach relies on measurements from aircrafts, tall towers, weather stations or satellites, and models to determine the total CH₄ flux rate from a specific region with multiple sites and sources. It uses a combination of atmospheric observations of CH₄ concentrations and modeling, and infer fluxes by accounting for atmospheric transport. The greatest challenge associated with this method is attributing observed CH_4 concentrations to specific sources (both anthropogenic and natural) and accurately representing atmospheric transport [8]. The bottom-up approach relies on measurements taken directly from leaking components and equipment or at a site level. Then, the measurements are extrapolated to a larger scale, for example national, by multiplying emission factors (emissions per component or per site per unit time) by activity factors (total number of components or sites). The goal in this approach is to measure emissions from a statistically representative sample of sources, in order to capture the entire emissions distribution. In most cases emissions distributions from the O&G infrastructure have non Gaussian distributions with a small fraction of the total population of the sources responsible for a disproportionately large share of emissions (right-skewed or "fat tailed") [9, 10, 11, 12]. These sites have often been referred to as super-emitters [9]. Their locations are difficult to predict and their occurrence can be stochastic. Additionally, some of these high emission facilities may have ongoing emissions, while others may have sporadic episodes of major releases [13].

Recent measurement-based studies, mostly in the United States and a few in Canada and Mexico, have consistently shown that across years, scales, and methods, estimates of total as well as $O\&G CH_4$ emissions from top-down approaches exceed bottom-up estimates based on emission inventories [10, 14, 15, 16, 17]. Alvarez et al. (2018) determined CH_4 emissions from the U.S. O&G 2015 supply chain by using a combination of ground-based, site-level measurements and aircraft observations and found that the annual national-scale estimate is approximately 60% higher than the U.S. Environmental Protection Agency (EPA) inventory estimate [15]. This percentage increases even more when considering only the production sector, with estimates by Alvarez et al. (2018) being almost 2.2 times higher than the EPA estimates [15]. One of the main possible reasons for these discrepancies is that published estimates are primarily based on outdated emission factors from the 1990s which may not reflect current technologies and practices. Measurements for generating emission factors are expensive, thus bottom-up estimates often rely on standard emission factors. Second, counts and location of facilities and equipment used in inventories are inaccurate, contradictory and incomplete. Lastly, sampling might be insufficient and not representative of the population of sites, especially if emissions distributions are positively skewed [16]. It is, therefore, necessary to reconcile and verify top-down approaches and bottom-up estimates based on emission inventories.

1.4 Romania and ROMEO campaign

Romania is one of the oldest O&G producers in Europe, dating back to 1857. According to BP's statistical Review of Energy Sector in 2021, Romania is the fourth largest oil producer and the fifth largest natural gas producer in the EU, including the UK [18]. The emissions reported for the year 2020 to the United Nations Framework Convention on Climate Change (UNFCCC) show that Romania has one of the highest annual emissions of CH_4 from the energy sector in Europe [19]. These emissions are especially relevant for the European Union's goals to urgently tackle CH_4 emissions across all sectors by 2030, under the EU Methane Strategy [20]. To achieve this, the strategy presents specific legal actions in the energy sector, such as an obligation to improve measurement and reporting of CH_4 emissions, improve leak detection and repair in the O&G infrastructure and ban venting and flaring [20]. Therefore, limiting these CH_4 emissions from the O&G sector of Romania could play a key role in reaching the greenhouse gas emission reduction targets of the European Union. However, there are concerns related to the accuracy of the reported emissions since the emission rates are estimations derived using standard (not country specific) emission factors and there are not enough atmospheric observations of CH_4 emissions from the O&G sector of Romania.

The Romanian Methane Emissions from Oil & Gas (ROMEO) measurement campaign aimed to address this gap in knowledge and investigate CH_4 emissions from O&G production in Romania [21]. ROMEO was initiated by the European H2020 project Methane goes Mobile - Measurements and Modelling (MEMO²) and is part of the international Climate and Clean Air Coalition's (CCAC's) Methane Science Studies [22]. From September 30th to October 20th, 2019, a three-week measurement campaign took place in southern Romania with up to 70 participants from 14 research institutes. By using a variety of measurement platforms and emission quantification methods, the goal of this project was to investigate CH_4 emissions at a component, facility and basin scale, thus providing a combined bottom-up and top-down quantification of CH_4 emissions related to O&G exploration, natural gas distribution and gas use from Romania [21].

1.5 Research objectives

In this paper we synthesize and analyse the ground-based measurements of CH_4 emissions collected during the ROMEO campaign and offers insights into the CH_4 emission patterns from the onshore O&G sector in Romania. The aim is to bring together several critical aspects of CH_4 emissions on a component and facility scale and to, ultimately, help reduce the uncertainties in the emission estimates and improve national inventories. We provide a comprehensive overview of the aggregated ground-based CH_4 emissions data and we answer the following research questions:

- 1. What are the major equipment sources of detected emissions across the O&G production sector?
- 2. How can we effectively characterize the emissions distributions and what are the differences between the quantification methods?
- 3. What are the estimated emission factors derived from the ground-based measurements, and how do these results compare to CH_4 emissions from production sites across other regions?
- 4. What can we learn about emission patterns so that CH₄ emissions can be reduced?

The paper is organized as follows. Section 2 consists of an overview of the sampling area, the measurement techniques used for the quantification of CH_4 emissions, and the methodology used for deriving emissions distributions and emission factors. Section 3 describes the results of our analysis with all the relevant figures divided into two separate sections, one for the component scale and one for the facility scale. Section 4 closes the paper with a discussion of the results and draws together the study's conclusions. Finally, the Supplementary material provides additional information and figures about the analysis and the results.

2 Material and methods

2.1 Investigated area

The main campaign covered the southern part of the country around the cities Bucharest, Ploesti, Pitesti, Targoviste and Craiova. The majority of the country's oil reservoirs are located in this area making it a mainly oil production region. Using information about the location of the sites provided by the O&G operator of the production sites, the investigated area was divided into regions and clusters. Fig. 1 shows a map of Romania, with target regions and clusters. The assigned regions correspond to areas with a high or low location density and magnitude of O&G production of gas and oil wells [23]. Due to their high density of O&G production sites, regions 2, 4, 5A, 6, 7 and 8 were selected as the main research areas. Each region was divided into smaller clusters of sites, each covering areas between 2 and 120 km² and consisting of a different number of sites (between 10 and 583 oil and gas related sites such as oil wells, gas wells, gas compressor stations and oil parks) (Korben et al., 2022, in preparation). Clusters were defined in order to obtain flying authorisations for the aerial measurements, coordinate the ground-based measurements and aircraft measurements and facilitate the comparison between the results from the top-down and bottom-up approaches. In general, each measurement team focused on different regions and clusters in order to visit as many O&G production sites as possible and to avoid a spatial sampling bias or duplicate measurements by different teams [23].

A few sites outside these regions and clusters were also measured. A few of these measured sites did not have a match within the full data set containing information about the total population of sites from Romania provided by the operators. In these cases, the site type on those sites either could not be confirmed, or a site type was assigned based on visual inspection. Fig. S1 shows a map of the location of quantified oil wells by each method, including a table providing the number of the surveyed sites and the total population of sites per region. In our analysis we combine the quantifications regardless the region. Finally, as will be further discussed later on, due to the nature of the region, the majority of the measurements during the ROMEO campaign was from oil wells, and thus, the rest of the analysis will focus on this specific subset.



Fig. 1. Map of investigated regions and clusters in Romania during ROMEO campaign. Regions are indicated with the purple and red colored areas. Clusters are illustrated as yellow outlined areas inside each region. The values correspond to the region number.

To assess how representative the measured sites were in comparison to the characteristics of the total population of sites in Romania and to determine possible differences between the characteristics of sites measured with different quantification methods, age, oil and gas production were compared. For the gas production, the majority of visited oil wells report zero gas production or had no gas production in 2019. We calculate the average gas production per site for the rest of the oil wells which report a non-zero value for their production. For the average well age, which is defined as the number of years in production, we use the reported spud dates from the operators. We perform this analysis for both the component and the facility scale measurements.

2.2 Measurement methods

A number of independent and complementary techniques were applied to measure CH_4 emissions on a component and facility scale. On the component scale, the combination of an Optical Gas Imaging (OGI) camera for the detection of potential leak sources and a Hi-Flow Sampler (HFS) device for the quantification of the emissions was implemented. On the facility scale, four methods were used to measure CH_4 emissions, namely Gaussian Plume Modelling (GPM), Other Test Method (OTM-33A), Tracer Dispersion Method (TDM) and Mass Balance Method (MBA) using Unmanned Aerial Vehicle (UAV)based measurements. The following section offers a brief introduction of the methods used.

2.2.1 Component scale

To investigate the origin of the emissions at the component scale, the first step was to screen the facilities using an Optical Gas Imaging (OGI) camera to locate CH_4 sources. The most frequently used

OGI technology for the detection, location and visualisation of CH_4 emissions involves the use of infrared imaging. An infrared camera visualizes a narrow range of the infrared spectrum where CH_4 and other hydrocarbons actively absorb radiation (between 3.3 - 3.4 µm). These cameras allow users to screen non-accessible components from a distance, quickly locate gas leaks from a variety of equipment and visualise emission plumes that are invisible to the naked eye [24]. A total number of 181 sites were visited and screened with a FLIR GasFindIR infrared camera. A total of 234 IR videos of the leaking components were recorded to document detected emissions. These videos were reviewed to verify the number of emission points and identify the type of emitting equipment.

While infrared cameras have been useful in detecting leaks, their use in quantifying leaks is still being studied, and is an area of ongoing research. Therefore, after the detection and location of potential leak sources with OGI, a secondary measurement device is needed for accurate quantification. Thus, once a site was scanned with the infrared camera, CH_4 emission concentrations from accessible identified leaks were measured with a Hi-Flow Sampler (HFS). The HFS is a portable, battery-operated instrument used to determine the rate of gas leakage from a variety of different components in the O&G infrastructure [25]. A component's leak rate is determined by sampling at a high flow rate to collect all the gas emitted from the component as well as a certain amount of surrounding air. The gas leak rate can then be calculated using the flow rate of the sampling stream and the gas concentration within that stream. In some cases, especially when the leaking component is located on top of a tank or it is abnormally shaped, access and safety restrictions prevent the use of the HFS. As a consequence, CH_4 emission concentrations might not be quantified from the total number of emitting sources from a single site, resulting in an underestimation of site-level emissions.

2.2.2 Facility scale

Facility scale measurements were carried out with multiple vehicles and drones equipped with methane analysers, and were divided into two phases: screening and quantification. During the screening phase, the vehicles drove from site to site, circling the target site if possible and recording CH_4 atmospheric concentrations. Screenings were performed in order to identify potential sources at the site, measure CH_4 concentrations above background, check site accessibility and determine whether off-site sources could interfere with subsequent emission quantification, thereby ensuring the proper implementation of the measurement methods. Also, once CH_4 atmospheric concentrations were collected via screenings, a Gaussian plume algorithm was applied to locate the sources and determine normalized CH_4 enhancements. The algorithm uses a radius of 100 m to look for any sites close to the maximum CH_4 concentration observed and then this emission is attributed to this particular site. Peaks are additionally scaled to 1 m width by keeping the peak area constant and considering the shape of the plume as Gaussian. This was done because sites were screened from a variety of distances and the maximum signal might not be representative of the actual emissions. Therefore, scaling the peaks allowed to compare all plumes easily. The final results of this algorithm determined normalized CH_4 enhancements.

In our analysis, we also investigate how different parameters in the model for the Gaussian fit affect the results. Specifically, we want to further explore what causes certain plumes, especially the ones in the zero mode of the maximum CH_4 distribution (Fig. 8 of the results section), to be rejected or accepted for the Gaussian fit. The team, that developed this algorithm, performed a sensitivity analysis and found two possible parameters, the first derivative of the concentration and the background level. The effect of each parameter was evaluated by changing the values for one parameter and keeping the value of the other parameter constant. The analysis was carried out for two of the instruments used during the screening phase. The analysis of only one of the instruments is reported here since both of them had similar results. Results are shown in S3.

In this study, the results of the screening phase were used to integrate and verify the information obtained through the different measurement methods. A total of 1043 sites were screened using five cars. The majority of those sites, namely 85%, were oil production sites, 10% were gas production sites, and the remaining 5% were facilities not included in the file provided by the O&G operators (Korben et al, 2022, in preparation). Methods deployed during the quantification phase are summarized in the following sections.

2.2.2.1 Tracer Dispersion Method

The Tracer gas Dispersion Method (TDM) or tracer release method is a ground-based remote sensing technique developed in the 1990s to quantify CH₄ emissions from a variety of natural gas systems and since then has widely been used in the O&G sector [26]. TDM involves the release of tracer compounds at a controlled, constant rate, at locations as close as possible to a suspected emission point at a site. Atmospheric concentrations of both the gas with unknown emission rate, in this case CH₄, and the tracer gas can then be measured downwind. The underlying concept of this method is that the tracer gas disperses in a manner equivalent to the CH₄ in the atmosphere, so they undergo the same atmospheric transport processes. Therefore, even when the plume dilutes, the ratios of their emission rates will be the same as their concentration ratios. When there was no site access, the tracer was released from the side of the fence protecting the wells, so that the distance between the expected source of methane emissions and the tracer gas would affect the quality of the measurements as little as possible. This method also requires the use of long-lived atmospheric tracer gases in order to maintain a steady concentration ratio between CH₄ and tracer gases during atmospheric dispersion. In this study, acetylene (C₂H₂) and nitrous oxide (N₂O) were used as tracer gases to measure CH₄ emissions.

If the concentrations of CH_4 was the same within the analytical uncertainty upwind and downwind of a certain site, then the emission rates in these sites were defined as being Below Detection Limit (BDL). Upper estimates for emission rates were assigned to these sites based on the lowest measurable emission rate that would have been detectable with the mobile equipment.

On several days of the ROMEO campaign, the TDM could not be applied due to technical maintenance required for the gas analyzer detecting the tracer gas. During these days, the GPM was applied. Additionally, because of site accessibility and appropriate wind conditions constrains, some emitting sites could not be successfully quantified by using the TDM. In these cases, the emission rates were calculated by applying the Gaussian plume method (GPM), using the peak CH_4 concentration recorded a few meters downwind of the site. This approach mostly uses only one concentration record, and not a proper GPM measurement and, emission rates from this approach are referred to as "Estimate". Based on the results of a statistical model developed by the TDM team, emission rates quantified using GPM or estimates were corrected by a factor of 2 or more. Specifically, by applying all three methods at 41 O&G sites, they found that for the majority of those sites the estimated CH_4 emission rates from GPM and estimate were lower than the emission rates quantified using TDM.

In summary, three types of emission rate assessment were used by the TDM team to investigate the sites: quantification by either TDM or GPM, estimation, and assessing whether the emission rate was BDL. Table S2 of Supplementary material provides a detailed overview of the investigated sites from the

TDM team in terms of type of site and type of evaluation. In total, 200 sites were sampled, including the sites defined as BDL. Overall, estimates accounted for 35% of the quantifications, tracer release experiments for 26%, GPM for 11.5% and the rest 27.5% were sites with emissions below the detection limit. More information about the TDM and its application during the ROMEO campaign can be found in the paper of Antonio et al. 2022 [23].

2.2.2.2 Gaussian Plume Method

The Gaussian plume method (GPM) uses an idealized calculation for the average local-scale CH_4 dispersion assuming constant meteorological conditions in time and space over a flat region to derive emission rate estimates from plume observations [27]. When a gas is released from an emission point, it is entrained in the prevailing ambient air flow in (defined as the x direction) and the dispersion from the emission point creates an idealized cone while it disperses in the y and z direction over time. Assuming the gas to be well mixed within the volume of the cone, the mixing ratio of the gas at any point, and eventually the emission rate, can be calculated by using information about the height of the source, wind speed and wind dispersion parameters [28] and applying Eq. 1 (Korben et al., 2022, in preparation).

$$\mathbf{Q} = 2\pi \cdot \boldsymbol{\sigma}_{\mathbf{v}} \cdot \boldsymbol{\sigma}_{\mathbf{z}} \cdot \mathbf{U} \cdot \mathbf{C} \tag{1}$$

Where σ_y and σ_z are the horizontal and vertical dispersion coefficients, U is the horizontal mean wind speed, and C is the maximum CH₄ concentration from the Gaussian fit algorithm. This method can be used on public roads without site access and offers an important advantage in terms of the limited need for equipment and time. However, GPM modeling can introduce systematic errors that are difficult to quantify and can give errors on emission magnitudes of at least a factor of three, if not more [29]. During the ROMEO campaign, multiple transects were carried out downwind from the source at locations marked as suitable for GPM. Then, based on the comparison between the results of the actual measured concentrations and the results of the GPM, the emission rate for each location was estimated. A total of 20 measurements were performed at a variety of sites using GPM.

Compared to the GPM measurements conducted by the TDM team, no scaling or correction was performed for these measurements. Since one of the objectives of this study is to evaluate and integrate the results from each different measurement method, and because of the small sample size of this GPM dataset, we decide to combine all quantifications performed with the GPM and the estimates into one dataset. Therefore, we decide to use the raw data of the GPM and the "Estimates" from the TDM team before the correction was performed. Hereafter, when we mention GPM, we will refer to this combined dataset. Similarly, when we mention TDM, we refer to the sites evaluated by actual tracer release experiments (column TDM from Table S2).

2.2.2.3 Other Test Method 33A

Other Test Method (OTM) 33A is one of the Geospatial Measurement of Air Pollution Remote Emission Quantification (GMAP-REQ) approaches developed by the United States Environmental Protection Agency (EPA) [30]. This test method refers to the use of ground-based vehicles to detect and assess emissions from a variety of sources located near-field and at ground level, and to estimate emissions in using a "Direct Assessment" approach. The idea is based on stationary observations of the concentration of trace gases in relation to the direction of the wind. OTM-33A involves detection of emissions by driving downwind of possible emission sources in order to transect an emissions plume and measure the ambient background CH_4 mixing ratio. If enhancements of CH_4 are detected, then the vehicle is safely parked downwind of the emission location, in the plume, and source emissions are evaluated in a short measurement time (approximately 20 minutes). However, one major restriction of this approach is that the emissions must come from a single point and no surrounding trees or other obstacles should be between the measurement point and the source [31]. Therefore, when the OTM-33A could not be applied, either because the topographic conditions were not suitable or because the wind conditions were not appropriate, the GPM was applied. A total of 77 quantifications were performed at different sites using OTM-33A.

2.2.2.4 Mass Balance Approach

The Mass Balance Approach (MBA) has been applied widely to aircraft-based measurements of CH_4 and other trace gas fluxes from the facility scale up to the basin scale. This method considers the conservation of the mass of CH_4 within a system (or volume), which is typically represented as a box. It involves flying downwind and/or around a region containing a possible source at a single vertical height or multiple heights and measuring what goes into and out of a volume of air. Emission rates are then estimated by taking the difference of the measurements as the net surface flux within that volume [31].

An emerging approach for deriving CH_4 emission fluxes from industrial regions using the MBA involves deploying integrated unmanned-aerial-vehicle (UAV) (typically known as drone) systems. Compared to air-crafts, UAVs are affordable, simple to use, much easier to deploy and have been used to quantify CH_4 emission fluxes from a wide range of emission sources such as landfills [32], dairy farms [33] and natural gas compressor stations [34]. They also allow transecting the plume over its entire vertical and horizontal extent compared to ground-based measurements that typically capture a certain portion of the plume [35]. During the ROMEO campaign, two teams performed quantifications using drones as their measurement tool. By combining these two UAV-based datasets, a total of 45 quantifications were carried out.

2.3 Determination of emissions distributions and emission factors

2.3.1 Log-normal distributions

A number of studies of CH_4 emissions from O&G sites have found that different types of O&G sites have highly positively skewed emission distributions with -as mentioned before- a small fraction of sites ("super-emitters") accounting for a large fraction of the total emissions. These distributions often become symmetric and normal when plotted as the logarithm of emissions. To account for this behaviour, log-normal distributions have been widely used in the literature [10, 12, 15, 16].

We first examined if our datasets are likely to have been derived from the log-normal distribution by using two statistical tests. To carry out these tests, we first log-transform the measured site-level emissions. The Shapiro-Wilk [36] and Lilliefors tests [37] for normality are then used to determine if the log-transformed data are normally distributed. These two tests are appropriate in a situation where the parameters (μ and σ) of the null distribution are unknown. Previous studies have found that the ShapiroWilk test is the most powerful normality test and the performance of Lilliefors test is quite comparable with Shapiro-Wilk test [38]. We perform the tests for the complete datasets as well as for the subset of oil wells including measurements above the detection limit of each method. The null hypothesis for the tests is that the log transformed emissions data comes from a normal distribution, with critical P-value of 0.05. The statistical tests were performed in *Python* using the scientific computation libraries SciPy [39] and statsmodels [40].

Table 1 shows the results from both statistical tests for each tested dataset. For the subset of oil wells, the null hypothesis of lognormality is accepted by both the Shapiro-Wilk and Lilliefors test for all four measurement methods. Therefore, we conclude that for oil wells, the assumption that the distribution of site-level emissions rates above the detection limit follows a log-normal distribution is valid. For the complete datasets, including every measurement for every type of site, the null hypothesis is accepted by the Lilliefors test for every method applied apart from the TDM, whereas the hypothesis is accepted by the Shapiro-Wilk test for every method applied apart from the MBA. We did not apply the statistical test to any other subset or different type of site because the sample sizes were too small.

For the screenings, we run the statistical tests only for the subset of oil wells. The null hypothesis of lognormality cannot be rejected by both tests only for two out of the five screening datasets, whereas for the other three datasets and the dataset combining all measurements, the hypothesis is rejected. We decide to apply the statistical estimator for the subset of oil wells to qualitatively compare the results between the quantifications and the screenings. However, we acknowledge that the log-normal distribution might not characterize the distribution from the screenings accurately.

Chouning	Shapiro - W	ilk test	Lilliefors test		
Grouping	P-value	Result	P-value	Result	
OTM-33A					
Complete dataset	0.824	Pass	0.970	Pass	
Oil wells	0.723	Pass	0.229	Pass	
GPM					
Complete dataset	0.098	Pass	0.627	Pass	
Oil wells	0.177	Pass	0.504	Pass	
TDM					
Complete dataset	0.093	Pass	0.045	Fail	
Oil wells	0.100	Pass	0.096	Pass	
MBA					
Complete dataset	< 0.0001	Fail	0.289	Pass	
Oil wells	0.494	Pass	0.682	Pass	
Screenings - oil w	ells				
Vehicle 1	0.018	Fail	0.001	Fail	
Vehicle 2	0.940	Pass	0.573	Pass	
Vehicle 3	0.377	Pass	0.722	Pass	
Vehicle 4	0.036	Fail	0.015	Fail	
Vehicle 5	0.002	Fail	0.013	Fail	
Combined vehicles	0.002	Fail	0.050	Pass	

Tab. 1. Results from the Shapiro-Wilk test and the Lilliefors test of lognormality for each tested dataset.

2.3.2 Statistical estimator

In this study, we use the statistical estimator developed by Zavala-Araiza et al. (2015) to estimate emissions probability density functions (pdfs) that follow a log-normal distribution [16]. These pdfs are then used to derive representative site-level emission factors. As a result, we obtain pdfs and emission factors that consider the effect of the low probability but high-emission sites that describe skewed distributions.

As a prerequisite (see above), for the subset of oil wells which passed the log-normality statistical tests, we assume that the emission rate distributions follow a log-normal distribution. Let x be the natural logarithm of CH4 emissions (in kg h^{-1}) measured at a site. Since x is normally distributed, the pdf of observing a single data point x, is given by:

$$p(\mathbf{x}|\boldsymbol{\mu}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mathbf{x}-\boldsymbol{\mu})^2}{2\sigma^2}}$$
(2)

Where μ and σ denote the mean and the standard deviation of the log-transformed data. We define $\Phi(x)$ as the cumulative standard normal:

$$\Phi(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\vartheta^2}{2}} d\vartheta$$
(3)

And:

$$\int_{-\infty}^{x} \mathbf{p}(\vartheta|\boldsymbol{\mu}, \sigma) \mathrm{d}\vartheta = \Phi\left(\frac{\mathbf{x} - \boldsymbol{\mu}}{\sigma}\right) \tag{4}$$

The natural logarithm of the likelihood function, or log-likelihood function is:

$$l(\mu, \sigma) = S_o \ln \Phi\left(\frac{DL - \mu}{\sigma}\right) - S_r \ln \sigma - \sum_{i=1}^{S_r} \frac{(x_i - \mu)^2}{2\sigma^2}$$
(5)

Where DL is the Detection Limit, or the lowest detectable emission rate, of each quantification method, S_o is the number of measurements at or below the detection limit and S_r is the number of measurements above the detection limit.

We use Maximum Likelihood Estimation (MLE) [41] to derive the parameters μ and σ by performing an optimisation routine which maximises Eq. 5. MLE is a popular method that allows us to use the observed data to estimate the parameters of the probability distribution that generated this observed sample. We also use a direct search algorithm to calculate 95% confidence intervals by inverting the Likelihood Ratio Test, a statistical test used to compare the goodness of fit between two models [42]. We can then use the maximum likelihood estimated parameters to derive a central, site-level emission factor on the arithmetic scale, EF, defined as:

$$EF = e^{\mu + \frac{1}{2}\sigma^2} \tag{6}$$

We can determine site-level emission factors by using the statistical estimator for sites that have sufficiently large sample sizes. Previous studies have successfully applied this approach with sample sizes equal or greater than 25 [15]. Zavala-Araiza et al. (2015) provide an extensive description of the statistical estimator approach as well as additional variations or constrains of this method [16].

2.3.3 "Non-detects" and Detection Limit

The implementation of the statistical estimator for the log-normal fits requires information about the detection limit of each method and the number of sites emitting at an emission rate below this detection limit, the so called "non-detects". However, even when using the same analytical platform to measure emissions, the lowest detectable emission rate will be affected by the measuring distance and the meteorological conditions for a given measurement [43]. Therefore, this parameter is specific to a given site at a given measurement time and cannot be applied broadly across different sites. However, since the detection limit is a necessary parameter for the log-normal fit and the calculation of the emission factors, we estimate an average detection limit for each quantification approach based on information from the datasets and the teams conducting the measurements. We do this for the subset of oil wells.

Korben et al. (2022, in preparation) used the screening data of two of the five screening vehicles to estimate the number of sites below the detection limit for the OTM-33A method. By evaluating the screening measurements from sites where CH₄ enhancements could not be detected or the CH₄ enhancement was below 200 ppb and, therefore, close to the background level, they determined that the non-detects correspond to a mean fraction of 35% for the subset of oil wells. As will be discussed in Section 3, we performed a similar independent evaluation of the screenings data and our results are in line with Korben et al. (2022, in preparation) findings. Therefore, we use the fraction of 35% of non-detects for our analysis. Brantley et al. (2014) determined the detection limit of OTM-33A method equal to 0.036 kg h⁻¹ [44]. Robertson et al. (2020) performed a sensitivity analysis using different detection limits but since no significant effect on the results was found, they also determined the detection limit as 0.11 kg h⁻¹, which is the lowest emission rate measured using OTM-33A in this study. Therefore, we also use this value for our analysis.

For the UAV-based measurements, the detection limit is also set equal to the lowest quantified value, which is the same as OTM-33A method, 0.11 kg h^{-1} . Since these teams visited approximately the same regions and their detection limit is the same, we determine the percentage of non-detects to be equal as the one for OTM-33A method, meaning 35%. We follow a similar idea and we determine the same values for GPM.

For the TDM measurements, we use the number of BDL values to determine an average detection limit and the percentage of "non-detects". By taking the average of the assigned BDL emission rates and calculating the fraction of these BDL values from the total number of measurements, we determine the detection limit to be equal to 0.07 kg h⁻¹ and the non-detects to be accounting for a fraction of 27% for oil wells. Roscioli et al. (2015) reported the detection limit of TDM equal to 0.02 kg h⁻¹ [45]. However, because of unfavorable meteorological conditions during the three-week campaign in Romania, we expected this value to be higher.

The effect on the log-normal fit and the final emission factors was further evaluated by testing several different values for the detection limit and the fraction of "non-detectc" (see S2 of Supplementary Material). We find that by decreasing the value of the detection limit or by increasing the fraction of non-detects, the estimated emission factors increase, due to the widening of the distribution towards the lower end.

3 Results

Here, we present the results of the analysis described in Section 2. We divide the section into two parts, one for the component scale and one for the facility scale. For both parts we provide a comprehensive overview of the ground-based CH_4 emissions data and we assess the samples' representativeness in comparison to the total population of sites in the country. For the component scale, we review the recorded videos of the leaking components, and we identify major equipment sources of detected emissions across the O&G production sector. For the facility scale, we characterize the emissions distributions, and we present the estimated emission factors derived from the different measurement methods. Additionally, we present a brief overview and analysis of the screening data which are used to integrate the information obtained through the ground based measurements.

3.1 Component Scale

3.1.1 Dataset overview

A total of 181 sites were screened with the infrared camera, corresponding to approximately 3% of the total population of sites provided by the operator. Table 2 shows the different type of sites visited, the number of sites with detected CH_4 emissions, the number of identified and quantified individual leaks from each site type and the range of quantified CH_4 emission rates per individual component with the HFS method. CH_4 emissions were detected from approximately half (49%) of these sites. A total of 231 individual leaks were identified and the emission rates of 62 (27%) of them were measured using the HFS method, whereas the rest 169 (73%) were assessed to be not accessible for quantification.

Site Description	$\# ext{ of } ext{sites } ext{visited}$	# of emitting sites	# of identified leaks	# of quantified leaks	$\begin{array}{c} {\rm Range \ of \ CH_4} \\ {\rm emission \ rates} \\ {\rm [kg/h]} \end{array}$
Oil wells	155	74	86	14	0.09 - 6.5
Gas wells	6	3	3	3	0.07 - 0.2
Oil parks	5	5	28	7	0.21 - 6.5
Gas compressor stations	2	2	85	33	0.02 - 1.6
Other facilities ¹	13	6	30	5	0.14 - 0.6
Total	181	89	231	62	0.07 - 6.5

Tab. 2. Overview of screened sites with infrared camera and measured CH_4 emission rates per individual leaking component.

¹ "Other facilities" include 4 oil production batteries, 2 disposal injection wells, 1 oil deposit, 1 random location and 5 sites mentioned as "other facilities" in the data provided by the O&G production operators.

Approximately 86% of sites visited were oil wells. At least one leak was detected at 74 out of the 155 screened oil wells with an average of 1.5 leaks detected per site. Only two gas compressor stations were monitored with an OGI technology and both of them had a very high number of emission sources.

One individual gas compressor station had 58 leaks and the other 27 leaks. After quick repairs at the first station, approximately half of the identified leaks were repaired. We find that oil parks and other facilities show a high number of leaks as well, with an average of 5.6 leaks and 5 leaks detected per site, respectively.

A total emission rate of 88 kg h⁻¹ was measured with the HFS method. These emissions were the result of many small individual sources. The breakdown of these emissions among various types of sites are shown in Fig. 2. Even though only 14 out of the 62 quantified individual leaks were from oil wells compared to the 33 quantified leaks from the gas compressor stations, total emissions were strongly dominated by oil wells. Specifically, these 14 emission points, with a range of emission rates between 0.09 kg h⁻¹ and 6.5 kg h⁻¹, accounted for more than 63% of total quantified emissions. Emissions from the 7 quantified components from oil parks ranged between 0.21 kg h⁻¹ to 6.5 kg h⁻¹ and contributed to approximately 18% of total emissions. The high number of leaking components from the two gas compressor stations represented only 16% of the total emissions and showed a small range of emission rates between 0.02 kg h⁻¹ to 1.6 kg h⁻¹.

Among the 181 investigated sites, there were 7 non-producing sites according to the operator (5 oil wells and 2 disposal injection wells). A non-producing site is defined as a facility that was not producing or operating at the time of measurement due to low reservoir pressure or technical problems [23]. Emissions were detected from 2 of the 5 non-producing oil wells. The first location had one identified leak and the second location had two identified and quantified leaks with emission rates equal to 0.2 kg h⁻¹ and 0.1 kg h⁻¹.



Fig. 2. Breakdown of CH₄ contributions to total measured emissions by type of site.

3.1.2 Oil wells

Fig. 3 shows the number of oil wells visited by the infrared camera per region and per cluster divided in sites with identified leaks and sites without identified leaks. The majority, 74 %, of the measurements were taken in region 6, indicating a very strong sampling bias towards this region (Fig. 2a). The percentage of emitting sites versus the percentage of non-emitting sites was higher in regions 7 (60%) and 5A (67%) compared to region 6 (43%). The sampling bias can also be seen in the visited clusters contained in each region (Fig. 2b). Most screened sites were in cluster C6 03. One interesting finding is that in

comparison to other clusters, in cluster $C6_03$, the percentage of non-emitting sites versus emitting sites is much higher. However, because of the lack of measurements in other clusters, it is unclear whether this distinction is robust.



Fig. 3. Number of screened oil wells per region and cluster divided by sites with identified leaks and sites without identified leaks.

A summary of the characteristics from the screened oil wells and from the total population of oil wells in Romania are shown in Table 3. No significant differences were found between emitting and non-emitting sites. For the gas production, approximately 70% of emitting and 82% of non-emitting oil wells visited report zero gas production or had no gas production in 2019. These percentages are higher than the average percentage of the total population of oil wells in the country, equal to 52%, already indicating a sampling bias towards low gas producing sites. Emitting oil wells had an average age of 36 years, average gas production of 9,500 scm per year and average oil production of 48 tons per year. We find similar range of values for non-emitting oil wells with an average age of 37 years, average gas production in 2019 based on information received from the operators were 37 years,

27,400 scm and 32 tons per year, respectively. Doing this comparison, we find that the sites visited were representative of the total population of sites in the country only in terms of age, with a slight focus on newer sites. However, measurements leaned more towards the high oil but very low gas producing end of the spectrum.

Tab. 3. Summary of characteristics (production and age) from screened oil wells and from the total population of oil wells in Romania.

Characteristics	Emitting oil wells	Non-emitting oil wells	Total population
Age [years]	36	37	37
Gas production $[10^3 \text{ scm per year}]$	9.5	7.5	27
Zero gas production [% of sites]	70	82	52
Oil production [tons per year]	48	52	32

3.1.3 Sources of observed emissions

From the manual inspection of the recorded infrared videos of the leaking components, 226 out of 231 detected emission sources could be attributed to specific major equipment types. Fig. 4 shows the frequency of the identified leaking components per type of site. For oil wells, among the seven different types of emitting equipment that could be identified, the most frequently detected sources were openended lines, accounting for more than half (55%) of the detected components, followed by inaccessible components located below the ground (25%) and other malfunctioning equipment such as flanges and threaded connections (20%). For gas compressor stations, 34% of the leaking components were classified as flanges, 20% as valves, 13% as threaded connections, 13% as compressor seals and the rest as open-ended lines, pressure gauges and inaccessible components below the ground. We find that oil parks and facilities share the same major source types, therefore we combine them into one category. The most frequently detected sources are oil tanks accounting for 55%, followed by open-ended lines, containers and a few other devices.



Fig. 4. Frequency of identified leaking components per type of site.

3.2 Facility-level quantifications

3.2.1 Dataset overview

A total of 342 sites were investigated by using one of the measurement methods outlined in Section 2, covering approximately 5.7% of the total population of sites in Romania. If we distinguish between different types of production infrastructure, the vast majority of the measurements, namely 218 sites including the sites assessed as BDL from the TDM team, were carried out at oil wells, which is the focus of our analysis. Table 4 provides an overview of the number of sampled sites according to the type of site for each measurement method used.

CH₄ emissions were investigated at 77 O&G sites with the OTM-33A method, 45 O&G sites with the MBA, 50 O&G sites with the TDM and 111 O&G sites with the GPM. CH₄ emission rates ranged from 0.11 kg h⁻¹ to 73 kg h⁻¹ for the OTM-33A, 0.0006 to 139 kg h⁻¹ for the GPM, 0.001 kg h⁻¹ and 107 kg h⁻¹ for the TDM, and 0.0004 kg h⁻¹ and 51 kg h⁻¹ for the MBA for all type of sites. Among all the 342 investigated sites, emissions were still detected from 12 non-producing sites at the time of measurements. Emission rates ranged between 0.006 kg h⁻¹ and 19 kg h⁻¹.

Site Description	Number of sites						
Site Description	OTM-33A	GPM^1	TDM^2	MBA			
Oil wells	54	68	25	33			
Gas wells	11	12	6	2			
Other facilities ³	6	30	19	8			
Unknown	6	6 1		2			
Total	77	111	50	45			

Tab. 4. Overview of the number of sampled sites according to the site type for each measurement method used.

 1 This category includes both GPM and "Estimates" based on one concentration record.

 $^2\ BDL$ values estimated from the TDM team are not included in this table.

3 "Other facilities" include oil parks, gas compressor stations, oil deposits, oil and gas production batteries, disposal injection wells and sites mentioned as "other facilities" in the data provided by the O&G production operators.

3.2.2 Oil wells

Table 5 provides basic statistics such as number of measurements, arithmetic mean, median, minimum, and maximum values for all methods used for the subset of oil well sites. Fig. 5 illustrates the distributions of the quantified emission rates from oil wells per method. CH₄ emission rates ranged from 0.11 kg h⁻¹ to 73 kg h⁻¹ for the OTM-33A, 0.0006 to 46 kg h⁻¹ for the GPM, 0.0012 to 27 kg h⁻¹ for the TDM, and 0.0004 kg h⁻¹ to 18 kg h⁻¹ for the MBA. The difference between the arithmetic mean and median estimates demonstrate that the emission rates were positively skewed for all methods, with a few measured wells responsible for the majority of the emissions. The more positively skewed a distribution is, the greater the mean estimates will be compared to the median estimates. We find that the GPM has the larger difference between mean and median, with a mean value of 5.9 kg h⁻¹ and a median value of 1.2 kg h⁻¹, whereas the MBA had the smallest difference, with a mean value of 2.3 kg h⁻¹ and a median value of 1.5 kg h⁻¹. OTM-33A and TDM had comparable estimates, with a mean value of 4.1 kg h⁻¹ and a median of 1.9 kg h⁻¹, for OTM-33A and, a mean value of 3.5 kg h⁻¹ and a median value of 0.6 kg h⁻¹ for TDM.

Tab. 5. Descriptive statistics of measured CH_4 emission rates by method.

Method	$\# \\ {f sampled} \\ {f sites}$	$\begin{array}{c} \mathbf{Arithmetic} \\ \mathbf{mean} \ [\mathbf{kg} \\ \mathbf{h}^{-1}] \end{array}$	${f Median} \ [{f kg} \ {f h}^{-1}]$	$\begin{array}{l} \mathbf{Min} \\ \mathbf{[kg \ h^{-1}]} \end{array}$	$egin{array}{c} \mathbf{Max} \ [\mathbf{kg} \ \mathbf{h}^{-1}] \end{array}$
OTM-33A	54	4.1	1.9	0.1100	73
GPM	70	5.9	1.2	0.0006	46
TDM	27	3.5	0.6	0.0012	27
MBA	33	2.3	1.5	0.0004	18



Fig. 5. Boxplots with the distributions of quantified emission rates from oil wells per method. In each box the red horizontal line signifies the median and the red square box shows the mean. The box extends to the 25th and 75th percentiles. The whiskers extend from the minimum to the maximum value. The data points are overlaid on top of the boxplots. Note the logarithmic y-axis.

A summary of the characteristics from the sampled oil wells and from the total population of oil wells in Romania are shown in Table 6. The distribution for site age shows little variability across the different methods. The mean age for the sites quantified by MBA is 30 years, only slightly higher than that for sites measured by GPM at 29 years. TDM focused on older sites compared to the other methods, with a mean age of 34 years, whereas OTM-33A measured younger sites with a mean age of 28 years.

In terms of production characteristics, the diversity of the sampled oil wells is more prominent. Among all measurement methods, TDM sites had the lowest average oil production of 43 tons per year, followed closely by MBA with 47 tons per year. GPM had the highest production of 77 tons of oil per year, more than double the country average value. OTM-33A also leaned towards high oil producing oil wells with an average of 61 tons per year. For the gas production, around 50% of the sampled oil wells with OTM-33A, GPM and MBA report zero gas production or had no gas production in 2019, whereas for the TDM this value is equal to 60%. These percentages are comparable to the true percentage, 52%, of the total population of oil wells in Romania. For the rest of the sites, which report a non-zero value for their production, their average gas production exhibits a wide range of values, with the TDM sites having the highest production of around 106,000 scm of natural gas per year and GPM having the lowest gas production with an average value of 12,000 scm per year.

As mentioned previously, country average values for age, gas and oil production in 2019 were 37 years, 27,400 scm and 32 tons, respectively. Based on these values, we find that oil wells sampled with all three measurement methods belong to the category of newer and high oil producing type of sites. In terms of gas production, only OTM-33A measurements were more representative to the total population of oil wells. TDM and MBA leaned towards the high, whereas GPM towards the low, gas producing end of the spectrum.

Characteristics	OTM-33A	GPM	TDM	MBA	Total population
Age [years]	28	29	34	30	37
Gas production $[10^3 \text{ scm per year}]$	26	12	106	49	27
Zero gas production [% of sites]	49	51	60	53	52
Oil production [tons per year]	61	77	43	47	32

Tab. 6. Summary of characteristics (production and age) from sampled oil wells based on the measurement method used, and from the total population of oil wells in Romania.

3.2.3 Emissions distributions and emission factors

Fig. 6 shows the pdfs generated from the results of the statistical estimator for each of the measurement methods applied. Similarly, Table 7 summarizes some key parameters and emission factors derived from the statistical estimator. The highest site-level emission factor estimated was 14 kg h⁻¹ of CH₄ (3.4 - 74, 95% confidence interval) for GPM, whereas the lowest one was 3.7 kg h⁻¹ of CH₄ (1.0 - 17, 95% confidence interval) for MBA. The pdf of GPM shows the widest distribution, and the estimated emission factor shows the largest 95% confidence interval, due to the largest range of CH₄ emission rates. The opposite applies for MBA. We get an emission factor of 7.3 kg h⁻¹ of CH₄ per site (2.2 - 30, 95% confidence interval) for OTM-33A, which is comparable to the site-level emission factor of 7.9 kg h⁻¹ of CH₄ (1.2 - 85, 95% confidence interval) for TDM. Compared to the OTM-33A, the effect of the small sample size to the estimations for the TDM is reflected in the large confidence interval. When we combine all the quantifications we get a central estimate of mean site-level emission equal to 8.3 kg h⁻¹ of CH₄ (3.8 - 19, 95% confidence interval). Histograms and fitted pdfs under the statistical estimator for each method used are shown in Fig. S4 of the Supplementary Material.

Method	DL	\mathbf{Sr}	So [% of non- detects]	μ	σ	$rac{\mathbf{EF}}{\mathbf{h}^{-1}}]$	95% CI
OTM-33A	0.11	53	$29 \ [35\%]$	-0.85	2.4	7.3	2.2 - 30
GPM	0.11	57	31 [35%]	-1.00	2.7	14	3.4 - 74
TDM	0.07	21	8 [27%]	-0.97	2.5	7.9	1.2 - 85
MBA	0.11	31	17 [35%]	-1.07	2.2	3.7	1.0 - 17
TOTAL	-	-	-	-0.98	2.5	8.3	3.8 - 19

Tab. 7. Summary of parameters from the statistical estimator.

DL is the detection limit of each measurement method, Sr is the number of measurements above the detection limit, So is the number of measurements at or below the detection limit, CI is the confidence interval and EF is the emission factor. TOTAL presents the results of the statistical estimator considering all four measurement methods.



Fig. 6. Fitted pdfs under the statistical estimator for each measurement method.

The cumulative distribution functions and Lorenz curves from all measurement methods exhibit highly skewed distributions (Fig. 7). For the total population of quantified oil wells, we find that the top 10% of emitters had emissions greater than 10 kg h⁻¹ and were responsible for over 85-90% of total emissions, while 75% of oil wells with lower emissions (less than 1.2 kg h⁻¹) only account for 5% of total emissions. We get similar results for the OTM-33A and the TDM. MBA shows slightly lower percentages between the methods, having 80% of cumulative CH₄ emissions attributed to 10% of sites. GPM shows a more skewed distribution with the 10% of sites with highest emissions contributing to 90-93% of total emissions.



Fig. 7. a) Cumulative distribution functions, b) Lorenz curve: percent of emissions as a function of percent of sites

3.2.4 Screenings

Table 8 shows the number of maximum CH_4 concentrations detected, the number of normalized CH_4 enhancements, and parameters μ and σ derived from the statistical estimator using the normalized CH_4 enhancements from each vehicle performing the screenings and the combination of their datasets.

We run the statistical estimator for the screening datasets by assuming that their emissions distribution is complete, meaning that there were no measurements below the detection limit. The estimated values for the parameter μ range between 1.4 and 2.3 ppm in logarithmic scale, with an average value of 1.9 ppm for the combined measurements. The estimated values for the width of the distributions, σ , range between 1.8 and 2.3 ppm in logarithmic scale, with an average total value of 2.0 ppm. Based on Eq. 1, the relationship between measured CH₄ concentrations and emission rates is linear, therefore we can qualitatively compare the results of the width of the distributions between 1.2 and 2.7 kg h⁻¹ in logarithmic scale, with an average total value of 2.5 kg h⁻¹. In contrast to what we were expecting, we find that the screening datasets show narrower distributions compared to the quantifications. However, the estimated parameters under the statistical estimator, may not accurately characterize the screening distributions since not all screening datasets passed the statistical tests for lognormality. Another reason for this discrepancy could be the effect of the fraction of non-detects to the width of the distribution. As discussed in Section S2 of the Supplementary Material, depending on the choice of the fraction of non-detects and the detection limit, the log-normal distribution fit might be widened.

Vehicle	$\# \text{ of Maximum CH}_4 \\ \text{Concentrations}$	# of Normalized CH ₄ Enhancements	μ	σ
1	317	181	2.0	1.8
2	138	26	2.3	1.9
3	384	177	2.1	2.3
4	345	169	1.9	1.8
5	171	119	1.4	2.2
Total	1355	672	1.9	2.0

Tab. 8. Overview of the number of maximum CH_4 concentrations detected, the number of normalized CH_4 enhancements, and parameters μ and σ derived from the statistical estimator using the normalized CH_4 enhancements

As mentioned above, we assume that the emissions distributions of the screening datasets are complete. As a consequence, we can use this dataset to obtain information about the sites below the detection limit of our measurement methods. Approximately 217 oil wells had normalized CH₄ enhancements lower than 2.2 ppm, accounting for 32% of the total number of screened oil wells. The value of 2.2 ppm is considered as the limit for OTM-33A (Korben et al., 2022, in preparation). Therefore, this percentage of 32% can be considered as the fraction of non-detects. Since we investigate measurements from a variety of different quantification methods, we calculate this fraction of low enhancements using different background levels. For a limit of 1.9 ppm, we get a fraction of 30%, whereas for a higher limit of 2.5 ppm, we get a fraction of 35%. These percentages are comparable to the fraction of non-detects that we used for the derivation of emission factors, equal to 35% (for OTM-33A, GPM and MBA), based on the results of Korben et al.,

2022 (in preparation), and 27% (for TDM), based on the fraction of BDL values from the TDM team.

Slightly more than half of the maximum CH_4 concentration peaks were rejected by the Gaussian model algorithm. This can be seen in both Table 8 and Fig. 8 depicting the distribution of the log-transformed maximum CH_4 concentrations for the rejected and accepted peaks. We find that more than half (55%) of these rejected peaks are very low values, specifically less than 2.7 ppm (or 1 in log-scale). We refer to this part of the distribution as zero mode. Apart from these sites with low CH_4 values, there are a number of wells with high maximum CH_4 concentration peaks which are rejected by the algorithm. One reason for this pattern is that in some situations, high CH_4 concentrations would be recorded in an area with two wells situated close to each other. In these cases, maximum CH_4 concentrations would be assigned to both of them, but the model for the Gaussian fit will assign normalised CH_4 concentration to one of the wells.



Fig. 8. Emissions distribution of the log-transformed maximum CH_4 concentrations for the rejected and accepted peaks of the Gaussian algorithm.

4 Discussion

On a component scale, the evaluation of the data sampled with the OGI and HFS methods demonstrate that approximately half of the sites visited emit CH_4 . The number of emission points per type of site varies widely. We find that the vast majority (83%) of individual oil wells present only 1 emission point, with a few wells having no more than 2 emission points. On average, larger facilities such as oil parks and especially gas compressor stations show multiple emission sources with up to 58 identified emission points for one gas compressor station.

An important finding of the OGI dataset analysis is the difference of the percentage of non-emitting sites versus emitting sites between cluster C6_03 and the rest of the visited clusters. Oil wells in cluster C6_03 are associated with emissions of Hydrogen Sulfide (H₂S) gas, a by-product that is formed in some fossil fuel reservoirs through natural processes or due to some methods employed in the O&G upstream production [46]. This chemical compound is a colourless, transparent gas with a characteristic of a rotten-egg odour at low concentrations which can be highly toxic to humans and animals, causing serious health problems even at low concentrations [47]. The difference between emitting and non-emitting wells between clusters demonstrate that wells associated with the H₂S component are better maintained to avoid dangerous H₂S emissions. This suggests that it is feasible to reduce emissions by improved practises and better maintenance of facilities.

Considering the breakdown of CH_4 sources by leaking component type, we find that more than threequarters of emissions from oil wells are from open-ended lines and components from somewhere below the ground, which could be open outlets as well. These emissions are thus part of normal operational practices and can be considered as venting. This finding can also be linked and supported by the fact that approximately 70% of CH_4 emitting oil wells visited with the infrared camera report zero gas production, while this percentage varies between 49% and 60% for the site-level quantifications. In many oil-producing regions, CH_4 is considered a byproduct and is deemed uneconomical to collect and sell. As a result, operators vent or flare CH_4 instead of building infrastructure to capture the produced gas. Since no flaring was observed during the three-week campaign, it is suggested that detected CH_4 emissions from the majority of sampled oil wells are vented emissions.

This is not the case for the other types of sites investigated. For gas compressor stations, more than 90% of the identified leaking components fall into the category of fugitive emissions from malfunctioning equipment such as flanges, valves and compressor seals. Similarly, for oil parks and other larger type of sites, more than 55% of the major sources are attributed to poorly maintained production storage tanks. From the inspection of the recorded infrared videos of the leaking components, we find that the majority of tanks exhibited signs of maintenance issues and aging such as corrosion and rust. Recent studies in North America suggest tanks as one of the largest sources of emissions from production sites, contributing to a very large share of emissions [13, 17, 48]. Due to safety and accessibility restrictions, in this study emissions from tanks were not quantified with the HFS and therefore, we are unable to evaluate the scale of these emissions.

4.1 Comparison with CH₄ emissions reported from other studies

The CH₄ emissions from the O&G sector have recently been the focus of a number of studies. We compare our findings with those from similar ground-based measurements in US and Canada O&G production areas in the US and Canada. Since each of these studies estimated average emission factors using a different approach, we use the reported datasets to derive the emission factors under the statistical estimator used in this paper. In this way we can consistently compare the results between the different regions. Table S3 provides an overview of the studies considered for this comparison and the derived parameters and emission factors using the statistical estimator. Fig 9 shows the derived cumulative distribution functions of each production region. All studies show heavy-tailed distributions; however, Romania presents one of the highest levels of skewness indicating the disproportionate contribution of high-emitting sites to the total emissions. Our results show that 10% of sites are responsible for more than 85% of emissions. By identifying and mitigating these high-emitting sites or "super-emitters", a large share of total emissions reduction can be achieved.

The CH₄ emission factor estimated for this study was 8.3 kg h⁻¹ site⁻¹ (3.8 - 19, 95% confidence interval) including all datasets from the four site-level quantification methods. CH₄ emission factors estimated for the studies used for our comparison range between 1.2 and 8.2 kg h⁻¹ site⁻¹ for O&G production sites, with the majority of the emission factors being below 3 kg h⁻¹ site⁻¹. We find that Romania presents the highest emission factor when compared to US O&G production regions (see Table S3). Specifically, our estimated CH₄ emission factor from Romania is slightly greater than the CH₄ emission factor from the Permian Basin, which is the largest oil-producing region in the US. The differences between production characteristics, age of sites, geologic features and operational procedures in each region could have a significant impact on the various levels of skewness and the emission factors. In Romania, we have identified venting via open-ended lines as one of the main sources for the high emission factor.



Fig. 9. Lorenz curve: cumulative percentage of emissions as a function of cumulative percentage of sites for different North American production regions, including the results from this study. The black dashed line shows the results of the statistical estimator for the ROMEO campaign, considering all four measurement methods. It overlaps with the one from the Fayetteville basin.

4.2 Comparison with CH₄ emissions reported from national inventories

In order to compare our results with the reported emissions from national inventories, we assume that the measured oil wells in this study are representative of oil wells basin-wide and we scale up our emissions to the country level by multiplying the estimated emission factor of 8.3 kg h⁻¹ site⁻¹ with the total number of active oil producing wells reported by the local O&G operator (N=3301). Assuming that these emissions continue year-round, this results in annual emission estimate of 240 ktons CH₄ yr⁻¹ (min = 110 ktons yr⁻¹ and max = 555 ktons yr⁻¹). We find that the annual CH₄ emissions derived from our measurements are significantly larger than the current reported emissions in national inventories. The total 2020 upstream O&G sector CH₄ emissions in the UNFCCC inventory were reported to be 83 ktons of CH₄, which is approximately 3 times less than the 2019 CH₄ emissions from oil wells estimated in this study (Fig. 10) [49]. Total 2020 upstream O&G sector CH₄ emissions estimate of 110 ktons CH₄ yr⁻¹ to published estimates, we still find an underestimation of total emissions.



Fig. 10. Comparison of annual CH₄ emissions estimated in our study for 2019 with estimates reported to the UNFCCC and derived by the IEA for the total 2020 upstream O&G sector. Error bar extends from the minimum value of 110 ktons yr^{-1} to the maximum value of 555 ktons yr^{-1} (not shown in the figure) based on the 95% confidence interval.

Our 250 ktons of $CH_4 \text{ yr}^{-1}$ estimate is likely to be conservative since it only includes oil wells, in particular a fraction of the total population of oil wells in Romania, and emissions from other types of sites (e.g., oil parks with our documented emissions from leaking tanks, and the entire gas production infrastructure) were not included. Non-producing oil wells were also neglected from the derivation of country-level annual emissions. However, we find that emissions were still detected from 9 oil wells that were characterised as non-operating at the time of measurement in the data provided by the operators. Therefore, the actual total CH_4 emissions from upstream O&G sector are likely more than a factor of 3 higher than the current inventory emissions reported.

Discrepancies between available inventory estimates and direct measured CH₄ emissions have been indicated by numerous studies in other areas [10, 14, 15, 16, 17, 50], and we now confirm this for the Romanian oil production basin. As mentioned in the introduction of this study, one main reason for these discrepancies is the use of outdated and highly uncertain emission factors for the derivation of inventory estimates. This is especially relevant for Romania since their published estimates are based on a tiered methodological approach following the IPCC 2006 guidelines [51]. Specifically, Romania uses the basic Tier 1 method which relies on multiplying default emission factors applicable for all countries by country-specific activity data. Thus, these reported emissions do not consider the characteristics of the actual O&G infrastructure of Romania, such as its age and state of maintenance, or current operational practises. For example, while gas flaring is almost eliminated as a practise in Romania, unlit or partly dismantled flare systems can be a significant source of unreported vented emissions into the atmosphere [52]. Additionally, alternatives to flaring, such as measures to recover natural gas, should be implemented when trying to reduce flaring. However, infrastructure for the collection and economical utilization of the natural gas that would otherwise be flared or vented is inadequate or non-existing in the surveyed areas. This combination of lack of gas flaring and gas recovery infrastructure raises questions about the reported low emissions estimates from national inventory.

As independent line of evidence, 70% of the screened oil wells and more than 50% of measured oil wells report zero gas production, but high CH₄ emissions were detected. Linking this finding to the attribution of emissions to specific equipment sources reveals that unreported extensive venting is the main driver of the majority of the emissions from oil wells in Romania. These results have great implications not only for the accuracy of current national inventories, but also for the feasibility of reaching EU emissions reductions targets. The total CH_4 emissions from the O&G sector in Romania reported to the UNFCCC decreased by 92.5% between 1989 and 2020 [53]. However, this significant reduction in emissions is mainly a consequence of decrease in production and changes in reporting methodology and not indicative of changes in operations that would result in lower emissions. The lack of gas flaring and gas collection infrastructure across oil production sites in Romania is evidence of the relatively high emissions. This might be the case for other countries as well. A large number of countries rely on the Tier 1 method, rather than direct site-level measurements, for the derivation of their national emissions estimates from the energy sector. However, since technological and operating conditions vary significantly between countries, it is likely that these estimates will be characterised by relatively large uncertainties and might not reflect actual emissions. This lack of accurate estimations and precise knowledge on the magnitude and the sources of CH_4 emissions makes it difficult to assess and ensure if emissions reductions are actually accomplished.

5 Conclusions

In this work, we provide a thorough characterization of CH_4 emissions from the O&G production sector of Romania by using a variety of ground-based quantification methods on a component and facility scale. The main findings are summarized as follows:

- 1. Major drivers of CH_4 emissions in Romania are the venting of gas through open-ended lines for oil wells, technical malfunctioning equipment for gas compressor stations and poorly maintained storage tanks for larger facilities, such as oil parks. These results highlight the need for operational improvement, better maintenance of sites, and replacement and upgrade of equipment.
- 2. Oil wells associated with emissions of H_2S are more well-maintained and had a lower number of detected emission points compared to oil wells without this component. Thus, effective mitigation of emissions can be achieved by improved practices and frequent leak detection and repair programs.
- 3. Quantified emission rates were represented by a mean emission factor equal to 8.3 kg h⁻¹ site⁻¹ (3.8 19, 95% confidence interval). The derived emission factor of Romania is among the highest ones compared to other production regions in North America.
- 4. CH_4 emission rates quantified from oil wells using a variety of different quantification methods were characterised by very skewed distributions, with 10% of sites contributing to more than 85% of the total CH_4 emissions. The combination of identifying these high-emitting sites and implementing regulations for unreported vented emissions serve as the key mitigation opportunities for reducing CH_4 emissions in Romania.
- 5. When scaling up our results to a national level, we find that national inventories might underestimate CH_4 emissions by at least a factor of 3. Given the importance of mitigating CH_4 emissions in the near-term future, there is an urgent need for emission mitigation and a reconsideration of published estimates from national inventories.

References

- V. Naik B. Adhikary P. Artaxo T. Berntsen W.D. Collins S. Fuzzi L. Gallardo A. Kiendler-Scharr Z. Klimont H. Liao N. Unger Szopa S. and P. Zanis. "2021: Short-Lived Climate Forcers". In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by P. Zhai A. Pirani S.L. Connors C. Péan S. Berger N. Caud Y. Chen L. Goldfarb M.I. Gomis M. Huang K. Leitzell E. Lonnoy J.B.R. Matthews T.K. Maycock T. Waterfield O. Yelekçi R. Yu Masson-Delmotte V. and B. Zhou. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2021. Chap. 6, pp. 817–922. URL: https://www.ipcc.ch/report/ar6/wg1/.
- [2] "Paris agreement". In: Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change (21st Session, 2015: Paris). Retrived December. Vol. 4. HeinOnline. 2015, p. 2017.
- [3] Van Dingenen R., Crippa M., Anssens-Maenhout G., Guizzardi D., and Dentener F. "Global trends of methane emissions and their impacts on ozone concentrations". In: JRC Science for Policy Report (2018). DOI: doi.org/10.2760/657568.
- [4] Saunois M., Stavert A. R., Poulter B., Bousquet P., Canadell J. G., Jackson R. B., Raymond P. A., Dlugokencky E. J., Houweling S., Patra P. K., and Ciais P. "The global methane budget 2000–2017". In: *Earth system science data* 12.3 (2020), pp. 1561–1623. DOI: 10.5194/essd-12-1561-2020.
- [5] Methane Emissions from Oil and Gas. https://www.iea.org/reports/methane-emissions -from-oil-and-gas. Accessed: 2022-08-28.
- [6] Tackling methane emissions from fossil fuel operations is essential to combat near-term global warming. https://www.iea.org/reports/methane-emissions-from-oil-and-gas-operations. Accessed: 2022-09-28.
- [7] Saunier S., Haugland T., and Pederstad A. "Quantifying cost-effectiveness of systematic leak detection and repair programs using infrared cameras". In: *Carbon Limits: Oslo, Norway* (2014).
- [8] Brandt A. R., Heath G. A., Kort E. A., O'Sullivan F., Pétron G., Jordaan S. M., Tans P., Wilcox J., Gopstein A. M., Arent D., Wofsy S., Brown N. J., Bradley R., Stucky G. D., Eardley D., and Harriss R. "Methane leaks from North American natural gas systems". In: *Science* 343.6172 (2014), pp. 733–735. DOI: 10.1126/science.1247045.
- Zavala-Araiza D., Lyon D, Alvarez R. A., Palacios V., Harriss R., Lan X., Talbot R., and Hamburg S. P. "Toward a functional definition of methane super-emitters: Application to natural gas production sites". In: *Environmental science & technology* 49.13 (2015), pp. 8167–8174. DOI: 10.1021/acs.est.5b00133.
- [10] Robertson A. M., Edie R., Field R. A., Lyon D., McVay R., Omara M., Zavala-Araiza D., and Murphy S. M. "New Mexico Permian Basin measured well pad methane emissions are a factor of 5–9 times higher than US EPA estimates". In: *Environmental Science & Technology* 54.21 (2020), pp. 13926–13934. DOI: 10.1021/acs.est.0c02927.
- [11] Zavala-Araiza D., Herndon S. C., Roscioli J. R., Yacovitch T. I., Johnson M. R., Tyner D. R., Omara M., and Knighton B. "Methane emissions from oil and gas production sites in Alberta, Canada". In: *Elementa: Science of the Anthropocene* 6 (2018). DOI: 10.1525/elementa.284.

- [12] Omara M., Sullivan M. R., Li X., Subramanian R., Robinson A. L., and Presto A. A. "Methane emissions from conventional and unconventional natural gas production sites in the Marcellus Shale Basin". In: *Environmental science & technology* 50.4 (2016), pp. 2099–2107. DOI: 10.1021/acs.est .5b05503.
- [13] Lyon D. R., Alvarez R. A., Zavala-Araiza D., Brandt A. R., Jackson R. B., and Hamburg S. P.
 "Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites". In: *Environmental science & technology* 50.9 (2016), pp. 4877–4886. DOI: 10.1021/acs.est.6b00705.
- [14] MacKay K., Lavoie M., Bourlon E., Atherton E., O'Connell E., Baillie J., Fougère C., and Risk D. "Methane emissions from upstream oil and gas production in Canada are underestimated". In: *Scientific reports* 11.1 (2021), pp. 1–8. DOI: 10.1038/s41598-021-87610-3.
- [15] Alvarez R. A., Zavala-Araiza D., Lyon D. R., Allen D. T., Barkley A. R. Brandt Z. R., Davis K. J., Herndon S. C., Jacob D. J., Karion A., Kort E. A., Lamb B. K., Lauvaux T., Maasakkers J. D., Marchese A. J., Omara M., Pacala S. W., Peischl J., Robinson A. L., Shepson P. B., Sweeney C., Townsend-Small A., Wofsy S. C., and Hamburg S. "Assessment of methane emissions from the US oil and gas supply chain". In: *Science* 361.6398 (2018), pp. 186–188. DOI: 10.1126/science.aar7204.
- [16] Zavala-Araiza D., Lyon D. R., Alvarez R. A., Davis K. J., Harriss R., Herndon S. C., Karion A., Kort E. A., Lamb B. K., Lan X., Marchese A. J., Pacala S. W., Robinson A. L., Shepson P. B., Sweeney C., Talbot R., Townsend-Small A., Yacovitch T. I., Zimmerle D. J., and Hamburg S. P. "Reconciling divergent estimates of oil and gas methane emissions". In: *Proceedings of the National Academy of Sciences* 112.51 (2015), pp. 15597–15602. DOI: doi.org/10.1073/pnas.15221261.
- [17] Tyner D.R. and Johnson M.R. "Where the Methane Is—Insights from Novel Airborne LiDAR Measurements Combined with Ground Survey Data". In: *Environmental Science & Technology* 55.14 (2021), pp. 9773–9783. DOI: 10.1021/acs.est.1c01572.
- [18] bp Statistical Review of World Energy 2022, 71st edition. https://www.bp.com/en/global/ corporate/energy-economics/statistical-review-of-world-energy.html. Accessed: 2022-09-08.
- [19] EEA greenhouse gases. https://www.eea.europa.eu/data-and-maps/data/data-viewers/ greenhouse-gases-viewer. Accessed: 2022-09-28.
- [20] European Commission. "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on an EU Strategy to Reduce Methane Emissions". In: (2020).
- [21] Thomas Röckmann. "ROMEO-ROmanian Methane Emissions from Oil and Gas". In: EGU General Assembly Conference Abstracts. 2020, p. 18801.
- [22] Oil and Gas Methane Science Studies. https://www.ccacoalition.org/en/activity/oil-and -gas-methane-science-studies. Accessed: 2022-10-20.
- [23] Delre A., Hensen A., Velzeboer I., van den Bulk P., Edjabou M. E., and Scheutz C. "Methane and ethane emission quantifications from onshore oil and gas sites in Romania, using a tracer gas dispersion method". In: *Elementa: Science of the Anthropocene* 10.1 (2022). DOI: 10.1525/elementa .2021.000111.
- [24] Lyman S. N., Tran T., Mansfield M. L., and Ravikumar A. P. "Aerial and ground-based optical gas imaging survey of Uinta Basin oil and gas wells". In: *Elementa: Science of the Anthropocene* 7 (2019). DOI: 10.1525/elementa.381.

- [25] INC Bacharach. "Hi-Flow Sampler[™] for natural gas leak rate measurement, Instruction Manual, Rev. 7 – July 2015". In: (2015).
- [26] Lamb B. K., McManus J. B., Shorter J. H., Kolb C. E., Mosher B., Harriss and E. Allwine R. C., Blaha D., Howard T., Guenther A., and Lott R. A. "Development of atmospheric tracer methods to measure methane emissions from natural gas facilities and urban areas". In: *Environmental science* & technology 29.6 (1995), pp. 1468–1479.
- [27] Kumar P., Broquet G., Yver-Kwok C., Laurent O., Gichuki S., Caldow C., Cropley F., Lauvaux T., Ramonet M., Berthe G., and Martin F. "Mobile atmospheric measurements and local-scale inverse estimation of the location and rates of brief CH₄ and CO₂ releases from point sources". In: *Atmospheric Measurement Techniques* 14.9 (2021), pp. 5987–6003. DOI: 10.5194/amt-14-5987-2021.
- [28] Riddick S. N., Connors S., Robinson A. D., Manning A. J., Jones P. S., Lowry D., Nisbet E., Skelton R. L., Allen G., Pitt J., and Harris N. R. "Estimating the size of a methane emission point source at different scales: from local to landscape". In: *Atmospheric Chemistry and Physics* 17.12 (2017), pp. 7839–7851. DOI: 10.5194/acp-17-7839-2017.
- [29] Yacovitch T. I., Herndon S. C., Pétron G., Kofler J., Lyon D., Zahniser M. S., and Kolb C. E. "Mobile laboratory observations of methane emissions in the Barnett Shale region". In: *Environmental* science & technology 49.13 (2015), pp. 7889–7895. DOI: 10.1021/es506352j.
- [30] Thoma E. and Squier B. "OTM 33 geospatial measurement of air pollution, remote emissions quantification (gmap-req) and OTM33A geospatial measurement of air pollution-remote emissions quantification-direct assessment (GMAP-REQ-DA)". In: US Environmental Protection Agency, Cincinnati, OH (2014).
- [31] Morales R., Ravelid J., Vinkovic K., Korbeń P., Tuzson B., Emmenegger L., Chen H., Schmidt M., Humbel S., and Brunner D. "Controlled-release experiment to investigate uncertainties in UAVbased emission quantification for methane point sources". In: Atmospheric Measurement Techniques 15.7 (2022), pp. 2177–2198. DOI: 10.5194/amt-15-2177-2022.
- [32] Allen G., Hollingsworth P., Kabbabe K., Pitt J. R., Mead M. I., Illingworth S., Roberts G., Bourn M., Shallcross D. E., and Percival C. J. "The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots". In: Waste Management 87 (2019), pp. 883–892. DOI: 10.1016/j.wasman.2017.12.024.
- [33] Vinković K., Andersen T., de Vries M., Kers B., van Heuven S., Peters W., Hensen A., van den Bulk P., and Chen H. "Evaluating the use of an Unmanned Aerial Vehicle (UAV)-based active AirCore system to quantify methane emissions from dairy cows". In: *Science of the Total Environment* 831 (2022), p. 154898. DOI: 10.1016/j.scitotenv.2022.154898.
- [34] Nathan B. J., Golston L. M., O'Brien A. S., Ross K., Harrison W. A., Tao Tao L., Laryt D. J., Johnson D. R., Covington A. N., Clark N. N., and Zondlo M. A. "Near-field characterization of methane emission variability from a compressor station using a model aircraft". In: *Environmental* science & technology 49.13 (2015), pp. 7896–7903. DOI: 10.1021/acs.est.5b00705.
- [35] Andersen T., Scheeren B., Peters W., and Chen H. "A UAV-based active AirCore system for measurements of greenhouse gases". In: Atmospheric Measurement Techniques 11.5 (2018), pp. 2683– 2699. DOI: 10.5194/amt-11-2683-2018.
- [36] Shapiro-Wilk test. https://en.wikipedia.org/wiki/Shapiro-Wilk_test. Accessed: 2022-10-10.

- [37] Lilliefors test. https://en.wikipedia.org/wiki/Lilliefors_test. Accessed: 2022-10-10.
- [38] Razali N. M. and Wah Y. B. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests". In: *Journal of statistical modeling and analytics* 2.1 (2011), pp. 21–33.
- [39] SciPy. https://scipy.org/. Accessed: 2022-10-10.
- [40] Statsmodels. https://www.statsmodels.org/stable/index.html. Accessed: 2022-10-10.
- [41] Jason Brownlee. A Gentle Introduction to Maximum Likelihood Estimation for Machine Learning. https://machinelearningmastery.com/what-is-maximum-likelihood-estimation-in -machine-learning/. Accessed: 2022-09-01.
- [42] Clarke Patrone. The Likelihood-Ratio Test. https://towardsdatascience.com/the-likelihood -ratio-test-463455b34de9. Accessed: 2022-09-10.
- [43] Delre A., Mønster J., and Scheutz C. "Greenhouse gas emission quantification from wastewater treatment plants, using a tracer gas dispersion method". In: Science of the Total Environment 605 (2017), pp. 258–268. DOI: 10.1016/j.scitotenv.2017.06.177.
- [44] Brantley H. L., Thoma E. D., Squier W. C., Guven B. B., and Lyon D. "Assessment of methane emissions from oil and gas production pads using mobile measurements". In: *Environmental science* & technology 48.24 (2014), pp. 14508–14515. DOI: 10.1021/es503070q.
- [45] Roscioli J. R., Yacovitch T. I., Floerchinger C., Mitchell A. L., Tkacik D. S., Subramanian R., Martinez D. M., Vaughn T. L., Williams L., Zimmerle D., and Robinson A. L. "Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods". In: Atmospheric Measurement Techniques 8.5 (2015), pp. 2017–2035. DOI: 10.5194/amt -8-2017-2015.
- [46] Marriott R. A., Pirzadeh P., Marrugo-Hernandez J. J., and Raval S. "Hydrogen sulfide formation in oil and gas". In: *Canadian Journal of Chemistry* 94.4 (2016), pp. 406–413. DOI: 10.1139/cjc -2015-0425.
- [47] Doujaiji B. and Al-Tawfiq J. A. "Hydrogen sulfide exposure in an adult male". In: Annals of Saudi medicine 30.1 (2010), pp. 76–80.
- [48] Wang J., Barlow B., Funk W., Robinson C., Brandt A.R., and Ravikumar A. "Large-Scale Controlled Experiment Demonstrates Effectiveness of Methane Leak Detection and Repair Programs at Oil and Gas Facilities". In: (2021).
- [49] IEA. Methane Tracker Database. https://www.iea.org/data-and-statistics/data-tools/ methane-tracker-data-explorer. License: Creative Commons Attribution CC BY-SA 4.0. IEA Paris, France, 2022.
- [50] Rutherford J. S., Sherwin E. D., Ravikumar A. P., Heath G. A., Englander J., Cooley D., Lyon D., Omara M., Langfitt Q., and Brandt A. R. "Closing the methane gap in US oil and natural gas production emissions inventories". In: *Nature communications* 12.1 (2021), pp. 1–12. DOI: 10.1038/ s41467-021-25017-4.
- [51] Eggleston H. S., Buendia L., Miwa K., Ngara T., and Tanabe K. "2006 IPCC guidelines for national greenhouse gas inventories". In: (2006).
- [52] Plant G., Kort E. A, Brandt A. R., Chen Y., Fordice G., Gorchov Negron A. M., Schwietzke S., Smith M., and Zavala-Araiza D. "Inefficient and unlit natural gas flares both emit large quantities of methane". In: Science 377.6614 (2022), pp. 1566–1571. DOI: 10.1126/science.abq0385.

- [53] Time Series Annex I. https://di.unfccc.int/time_series. Accessed: 2022-09-08.
- [54] Gini coefficient. https://en.wikipedia.org/wiki/Gini_coefficient. Accessed: 2022-10-10.

Supplementary Material

OIL WELLS Legend GPM MBA OTM-33A TDM 8 Sampled oil wells Total number of Region oil wells per отм TDM MBA GPM region 33A C2 59 778 14 C4 3 597 C5A 20 14 8 523 C6 4 12 10 35 C7 14 26 283 3 2 C8 30 1 2 2 boogle Earth CO 1584 27 54 33 Total 70 3854

S1 Location of quantified oil wells

Fig. S1. Map illustrating the location of quantified oil wells grouped by region. The table shows the number of measured sites and the total population of sites in each region.

S2 Sensitivity analysis of the statistical estimator

The results of the statistical estimator depend heavily on two parameters, the detection limit of the measurement method and the number of sites below this detection limit, in other words, the non-detects. We tested the sensitivity of the log-normal fits by running the statistical estimator for three different parameter values for both the detection limit and the fraction of non-detects. We use the subset of oil wells from the OTM-33A method for the sensitivity analysis. Table S1 provides the summary of the parameters and Fig. S2 presents the fitted pdfs derived from the statistical estimator. By decreasing the value of the detection limit or by increasing the fraction of non-detects, the estimated EFs increase, due to the widening of the distribution towards the lower end. This behaviour is more prominent and results in very large EF estimates when the detection limit is very low. The choice of the detection limit does not affect the high end of the distribution substantially, and the choice of the percentage of non-detects has an even smaller impact. These findings underscore the sensitivity of the statistical estimator to the low end of the distribution and highlight the need for caution when choosing the values of these two parameters.

Tab. S1. Summary of parameters from the statistical estimator calculated using different values for the detection limit and for the fraction of non-detects.

Parameter	DL	\mathbf{Sr}	So [% of non- detects]	μ	σ	${f EF} \ [kg \ h^{-1}]$	95% CI
	0.036	54	$29 \; [35\%]$	-1.39	3.06	27.2	4.43 - 234.58
Detection limit	0.11	53	29 [35%]	-0.85	2.38	7.30	2.21 - 29.50
	0.2	52	28 [35%]	-0.52	2.01	4.52	1.83 - 12.80
	0.11	53	18 [25%]	-0.31	2.03	5.71	2.33 - 16.11
% of non-detects	0.11	53	29 [35%]	-0.85	2.38	7.30	2.21 - 29.50
	0.11	53	43 [45%]	-1.47	2.73	9.65	2.13 - 57.41

DL is the detection limit of each measurement method, Sr is the number of measurements above the detection limit, So is the number of measurements at or below the detection limit, CI is the confidence interval and EF is the emission factor.



Fig. S2. Pdfs derived from the statistical estimator calculated using different values for the detection limit and for the fraction of non-detects.

S3 Sensitivity analysis of the Gaussian fit

We investigate how different parameters in the model for the Gaussian fit affect the results. Specifically, we want to further explore what causes certain plumes, especially the ones in the zero mode of the maximum CH_4 distribution (Fig. 8), to be rejected or accepted for the Gaussian fit. The team, that developed this algorithm, performed an advanced sensitivity analysis and found two possible parameters, the first derivative of the concentration and the background level. The effect of each parameter was evaluated by changing the values for one parameter and keeping the value of the other parameter constant. The analysis was carried out for two of the instruments used during the screening phase. The analysis of only one of the instruments is reported here since both of them had similar results.

The first parameter is related to the sensitivity of the first derivative of the concentration as a function of space. This parameter is important for locating the beginning and the end of peaks as the derivative increases or decreases, and the top of peaks where the derivative is close to zero. The team tested the effect of this parameter to the number and distribution of the peaks by setting the threshold of the increase of the derivative between values of 10^{-7} and 10^{-2} , and keeping the background level constant and equal to 1.1. Fig. S3a shows the tested values of the threshold parameter as a function of the number of peaks identified from the algorithm. We find that by lowering the threshold, more peaks are identified, and vice versa. This is something that we expected since the lower the threshold, the larger the peaks, or in other words, the steeper the slope of the peak has to be in order to be detected. For values below 10^{-4} , the number of peaks remains constant.

The second parameter is the background value required for peak to end. The background level is determined as the factor by which the concentration before the peak is multiplied. For example, background level equal to 1.1 means that if the value before the beginning of the peak was 2ppm then the end of the peak will be set when concentration falls below 2.2ppm. The team tested the effect of this parameter to the number and distribution of the peaks by setting the background level between values of 1.001 and 1.285, and keeping the threshold constant and equal to 10^{-6} . Fig. S3b shows the tested values of background level as a function of the number of peaks identified from the algorithm. We find that by increasing the background level, a higher number of peaks are detected, and vice versa. A lower background value will not distinguish between individual smaller peaks and will group them together treating them as one large peak. For values below 10^{-4} , the number of peaks remains constant.

The final screening dataset analysed in our study used the values 0.001 and 1.1 as the threshold and the background level factor, respectively. These values might not be optimal and might be causing the algorithm to reject a number of peaks, however, the results do not change substantially.



Fig. S3. a) Background level, and b) First derivative of the concentration (or threshold). Both are a function of the number of identified peaks.

S4 Overview of the sampled sites per type of emission rate assessment used by the TDM team.

Tab. S2. Overview of the sampled sites per type of evaluation for the measurements of the TDM team.

Site Description	Number of sites				
	BDL	GPM	TDM	Estimates	
Oil wells	34	15	27	46	
Gas wells	16	-	6	11	
Facilities	5	8	19	13	
Total	55	23	52	70	



S5 Histograms and fitted pdfs under the statistical estimator for each measurement method used.

Fig. S4. Histograms and fitted pdfs under the statistical estimator for each measurement method used.



S6 Screening data distributions

Fig. S5. Emissions distributions from a) different screening vehicles, b) the total number of the screenings from every vehicle used. The black dashed vertical line in the lower graph indicates the detection limit of 2.2 ppm used for the OTM-33A dataset.

S7 Comparison with CH_4 emissions reported from other studies

Tab. S3. Summary of estimated parameters derived from the statistical estimator for each of the production regions used in our comparison.

Dataset	μ	σ	$\mathbf{EF} \ [\mathbf{kg} \ \mathbf{h}^{-1}]$	$\begin{array}{c} {\bf Gini} \\ {\bf coefficient}^1 \end{array}$	Reference
Denver - Julesburg (Colorado, US)	-0.62	1.3	1.2	0.63	Robertson et al., 2017
Barnett Shale (Texas, US)	-1.8	2.2	1.8	0.88	Zavala-Araiza et al., 2015
Red Deer (Alberta, Canada)	-0.31	1.5	2.2	0.70	Zavala-Araiza et al., 2018
Upper Green River (US)	0.32	1.0	2.4	0.53	Robertson et al., 2017
Fayetteville (Arkansas, US)	-2.1	2.5	2.5	0.92	Robertson et al., 2017
Uintah (Wyoming, US)	0.17	1.3	2.7	0.63	Robertson et al., 2017
Marcellus (US)	0.39	1.8	7.3	0.79	Omara et al., 2015
Permian (Texas, New Mexico, US)	1.5	1.1	8.2	0.56	Robertson et al., 2020
Romania (Europe)	-0.98	2.5	8.3	0.92	This study
Compressor stations (national, US)	3.1	1.5	64	0.71	Zavala-Araiza et al., 2015
Processing plants (national US)	4.4	1.3	190	0.64	Zavala-Araiza et al., 2015

¹ The Gini coefficient is a measure of statistical dispersion used to estimate the inequality among values of a frequency distribution. A Gini coefficient of 0 represents complete equality, whereas a Gini coefficient close to one expresses the maximum inequality among values where one component emits strongly while all the other components emit nothing [54].