Improving Consistency in Methane Emission Quantification from the Natural Gas Distribution System across Measurement Devices

Judith Tettenborn[1], Daniel Zavala-Araiza[1,2], Daan Stroeken[1], Hossein Maazallahi[1*], Carina van der Veen[1], Arjan Hensen[3], Ilona Velzeboer[3], Pim van den Bulk[3], Felix Vogel[4], Lawson Gillespie[4,5], Sebastien Ars[4], James France[6,7], David Lowry[6], Rebecca Fisher[6], and Thomas Röckmann[1]

[1]Institute for Marine and Atmospheric Research Utrecht (IMAU), Utrecht University, Utrecht, The Netherlands
[2]Environmental Defense Fund, Amsterdam, The Netherlands
[3]Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, The Netherlands
[4]Climate Chemistry Measurements and Research, Climate Research Division, Environment and Climate Change Canada, Toronto, Canada
[5]Department of Physics, University of Toronto, Toronto, Canada
[6]Department of Earth Sciences, Centre of Climate, Ocean and Atmosphere, Royal Holloway, University of London, Egham, United Kingdom
[7]Environmental Defense Fund, London, United Kingdom
[*]Now at: Department of Renewable Energies and Environment, College of Interdisciplinary Science and Technologies, University of Tehran, Tehran, Iran

Correspondence: Thomas Röckmann (t.roeckmann@uu.nl)

Abstract.

Mobile real-time measurements of ambient methane provide a fast and effective method to identify and quantify methane leaks from local gas distribution systems in urban areas. The objectives of these methodologies are to i) identify leak locations for repair and ii) construct measurement-based emission rate estimates, which can improve emissions reporting and contribute to monitoring emission changes over time. Currently, the most common method for emission quantification uses the maximum methane enhancement detected while crossing a methane plume. However, the recorded maximum depends on instrument characteristics, such as measurement cell size, pump speed and measurement frequency. Consequently, the current approach can only be used by instruments with similar characteristics. We suggest that the integrated spatial peak area is a more suitable quantity that can eliminate the bias between different instruments. Based on controlled release experiments conducted with various different devices in four cities (London, Toronto, Rotterdam, and Utrecht), emission estimation methodologies were evaluated. Indeed, when different analyzers were measuring in the same vehicle and from the same air inlet, the integrated spatial peak area was found to be a more robust metric across different methane gas analyzer devices than the maximum methane enhancement. A statistical function based on integrated spatial peak area is proposed for more consistent emission estimations when using different instruments. On top of this systematic relation between actual emission rate and recorded spatial peak area, large variations in methane spatial peak area were observed for the multiple transects across the same release point, in line with previous experiments. This variability is the main contributor of uncertainty in efforts to use mobile measurements to prioritize leak repair. We show that repeated transects can reduce this uncertainty and improve the categorization into dif-

ferent leak categories. We recommend a minimum of three and an optimal range of 5-7 plume transects for effective emission quantification to prioritize repair actions.

1 Introduction

Mitigating methane (CH_4) emissions is an important step to combat climate change. The increase in atmospheric CH_4 has contributed approximately 30% of global warming since pre-industrial times (IPCC, 2023). CH_4 has an 82 times higher global warming potential over 20 years (Forster et al., 2021), and a shorter atmospheric lifetime (9.1 ± 0.9 years (Szopa et al., 2021)) than CO_2. This offers the opportunity for reductions in climate warming by reducing CH_4 emissions, possibly avoiding temperature overshoots (Hu et al., 2013; Collins et al., 2018). General awareness of the importance for timely CH_4 mitigation is rising. More than 150 nations have joined the Global Methane Pledge, established during the UN Climate Change Convention Of Parties (COP 26) in 2021, pledging to reduce human-made CH_4 emissions in 2030 by 30% relative to 2020 (EU Commission, 2021). Following this, regulations specifically targeting CH_4 emissions have been adopted or are underway, for example, in Canada, Colombia, the EU, and the U.S. (Environment and Climate Change Canada, 2022; Ministerio de Minas y Energía, 2022; EU Commission, 2023; U.S. Environmental Protection Agency, 2023).

About a quarter of total global anthropogenic CH_4 emissions can be attributed to the oil and natural gas sector (Saunois et al., 2020). The International Energy Agency (2023) suggests that deploying existing technological solutions could cut 75% of global oil and gas CH_4 emissions, at a cost less than 3% of the net income of the oil and gas industry in 2022. After the production segment, emission intensity is highest in the downstream distribution segment (Cooper et al., 2021; International Energy Agency, 2023). Regionally, this sector can even represent the largest share, especially in the EU, where the vast majority of oil and gas used is imported (EU Commission, 2023).

Leak detection, quantification, and verification surveys are key to mitigate CH_4 emissions from the oil and gas supply chain. CH_4 analyzers on various platforms such as satellites, aircraft or vehicles have been utilized to detect and quantify fossil fuel related CH_4 emissions (Jacob et al., 2016; Cui et al., 2019; Maazallahi et al., 2020; Irakulis-Loitxate et al., 2021; Shen et al., 2021; Zavala-Araiza et al., 2021; Korbeń et al., 2022). Both rapid detection and reliable quantification can help to improve and prioritize capital-intensive leak repair efforts (von Fischer et al., 2017). This is especially important since leak distributions have been found to be skewed, a few large leaks being responsible for a major share of total emissions (Zavala-Araiza et al., 2015; Omara et al., 2016; Robertson et al., 2017; Maazallahi et al., 2020; Weller et al., 2020; Stavropoulou et al., 2023). Beyond the direct mitigation opportunity, reliable quantification can help to evaluate the scale of fugitive emissions and possible emission reductions over time, related to the Global Methane Pledge. Vehicle-based mobile measurements deploying fast CH_4 gas analyzers have proven to be an efficient method for quickly and effectively surveying distribution networks across urban areas. There is currently no universally accepted measurement and data analysis methodology in place. Most widely used is the statistical methodology developed by von Fischer et al. (2017) and later refined by Weller et al. (2019). Based on controlled release experiment data (von Fischer et al., 2017), they derived the following relation between the observed maximum excess

CH$_4$ (in ppm) when crossing a plume and the release rate (r_E, in Lmin^{-1})

$$\ln([\mathrm{CH_4}]_{\max}) = 0.817 \cdot \ln(r_E) - 0.988. \tag{1}$$

In the original multivariate regression model suggested by von Fischer et al. (2017), the release rate was used as the response variable, and maximum enhancement, integrated spatial peak area and an index of the plume kurtosis were used as predictors. Weller et al. (2019) argued that maximum CH$_4$ enhancement proved to be the best predictor of the release rate and that additional predictors did not meaningfully improve emission rate estimates.

This transfer equation has been applied in several measurement campaigns in North America and Europe (Phillips et al., 2013; von Fischer et al., 2017; Weller et al., 2018; Ars et al., 2020; Maazallahi et al., 2020; Defratyka et al., 2021; Fernandez et al., 2022; Wietzel and Schmidt, 2023; Vogel et al., 2024). During each survey, hundreds to thousands of leak indications were detected. The emission estimates were also used to rank leak indications into different categories (high (>40 Lmin^{-1}), medium (6-40 Lmin^{-1}) and small (<6 Lmin^{-1}) emissions), to assist with repair prioritisation decisions (von Fischer et al., 2017; Maazallahi et al., 2020). However, several problems have been identified. Leak indications cannot always be confirmed through re-measurement, and gas providers cannot always identify the source of the emissions (von Fischer et al., 2017; Weller et al., 2018; Maazallahi et al., 2020). Furthermore, environmental factors, such as wind, but also dispersion and turbulence within the emission plume, lead to large uncertainties in measurements with corresponding over- and underestimations. Especially within urban built environments, complex wind flow, re-circulation, and dispersion patterns play an important role (Carpentieri and Robins, 2015; Ražnjević, 2023). In addition, the approach developed in von Fischer et al. (2017) was derived for measurements using one specific device (Picarro G2301). By now, multiple CH$_4$ analyzers are in use, and it has been shown that CH$_4$ enhancements measured with different instruments can differ substantially (Gillespie et al., 2023; Maazallahi et al., 2023). These discrepancies are expected, as different analyzers exhibit distinct instrument characteristics. Parameters such as cell volume, cell temperature, cell pressure, measurement frequency, and flow rate vary among instruments, all of which influence the detected shape of the CH$_4$ peak (Takriti et al., 2021). Thus, the transfer function developed by Weller et al. (2019), where estimated emission rates purely depend on the measured peak maximum, is not transferable to other analyzers.

Maazallahi et al. (2020) noted that the integrated spatial peak area is much more consistent than the peak maximum between two different instruments operated in parallel during mobile measurement campaigns in Hamburg and Utrecht. Gillespie et al. (2023) conducted controlled release experiments and street-level measurements in Toronto, using two CH$_4$ devices. By applying Gaussian plume inversions to both CH$_4$ peak maximum and peak area, they concluded that emission rates should be estimated using either the peak area or, if relying on enhancement height, incorporating explicit modelling of the instrument response function to account for variations in cell residence time. Ars et al. (2020) found the maximum amplitude of CH$_4$ enhancement to be $50-80\%$ lower when sampled with a high inlet at 2.5 m compared to sampling close to the ground, whereas the integrated spatial peak area was better comparable between the measurements at different heights.

This study evaluates results from different controlled CH$_4$ release experiments in four cities. Predictors for statistical emission rate estimation are evaluated. Specifically, we evaluate the consistency between measured peak heights and spatial peak areas across eight different instruments when transecting the same emission plume with the same air inlet. We then propose an

instrument-independent transfer equation that uses the integrated spatial peak area. We evaluate how successful this transfer
equation is in categorizing emissions into different categories for single and multiple passes.

2 Methods

In this section, we first describe the controlled release experiments conducted to generate the dataset (Sect. 2.1). We then
outline the data processing steps, including peak identification and calculation of the Spatial Peak Area (Sect. 2.2). Next, we
present the approach for estimating emission rates from the measurements (Sect. 2.3). Finally, we evaluate the performance of
the quantification method (Sect. 2.4), considering both its categorization ability and the impact of the number of transects on
estimation accuracy.

2.1 Controlled Release Experiments

Controlled release experiments (CRE) were conducted with a total of nine different analyzers in four different cities (London,
Rotterdam, Toronto and Utrecht) situated in three different countries by different research groups (Tab. 1). The release locations
are visualized in the Supplement Information (SI), Fig. S1 and an overview of the specific release rates per release location is
given in the SI, Tab. S2.

Table 1. Controlled release experiments: Overview of cities, release rates, inlet height and location. In all experiments, the release height was set at ground level.

City & Experiment	Date	GHG Analyzer	Release Rate Range [Lmin^{-1}]	Inlet Height [m]	Terrain
London, UK, I Day-1	2019-09-10	G2301-m[a], uMEA[b]	35, 70	1.8	Open field
London, UK, I Day-2	2019-09-11	G2301-m[a], LI-7810[c]	35, 70	1.8	Open field
London, UK, I Day-3	2019-09-13	G2301-m[a]	70	0.3, 1.8	Open field
London, UK, II Day-1	2024-05-13	LI-7810[c]	1–70.5	1.8	Open field
London, UK, II Day-2	2024-05-14	LI-7810[c]	0.2–1	1.8	Open field
Rotterdam, NL	2022-09-06	G2301[a], G4302[a], TILDAS[d], Mira Ultra[e], MGA10[f]	0.15–120	1.7, 3	Sub-urban
Toronto, CA, Day-1	2021-10-20	G2401[a], UGGA[b]	2.5–20	1.6, 2.5	Industrial area
Toronto, CA, Day-2	2021-10-24	G2401[a]	0.1–5	2.5	Parking lot
Utrecht, NL, I	2022-11-25	G2301[a], G4302[a]	2.2–15	0.5	Urban
Utrecht, NL, II	2024-06-11	G2301[a], Mira Ultra[e]	0.15–100	0.5	Urban

[a] Picarro INC, Santa Clara, USA. [b] Los Gatos Research, San Jose, USA. [c] LI-COR Environmental, Lincoln, USA. [d] Aerodyne Research, Billerica, USA. [e] Aeris Technologies, Eden Landing Road Hayward, CA. [f] MIRO Analytical AG, Wallisellen, CH.

The controlled release experiment in Rotterdam was conducted on September 6, 2022. The location was selected to reflect common urban characteristics with houses, parked cars and overhanging trees (see SI, Fig. S2). Methane (purity $> 99.9\%$) was released through a 1/4' O.D. teflon tube from two cylinders placed at a total of three locations along two connected streets at a wide range of flow rates (0.15–120 Lmin^{-1}). However, at two of the three locations, the rotameter used to adjust the release rate was suspected to not work properly. Therefore, only releases at location 1 (flowrate controlled electronically by an Alicat MCP-100SLPM, 5–120 Lmin^{-1}) are considered. The release occurred 1–3 m from the street, with the inlet line moved to 5 m for flow rates above 40 Lmin^{-1}. Atmospheric CH_4 mole fraction was measured while driving along the release locations, using five different instruments distributed over two vehicles. The first vehicle is Utrecht University's Air Quality car (UUAQ, Institute for Risk Assessment Sciences, Utrecht University), an Opel Astra with an air inlet on its roof (inlet height ca. 1.7 m, see SI, Fig. S2). This car contained two cavity ring-down spectroscopy (CRDS) analyzers, model G2301 and G4302 and a mid-infrared laser absorption spectroscopy analyzer MIRA Ultra. Two instruments, a MGA10 analyzer and TILDAS Dual Laser Trace Gas Analyzer, were utilized in the measurement trailer of a semi-trailer truck operated by Netherlands Organisation for Applied Scientific Research (TNO). The inlet is on the side of the trailer around 2.5 m above ground level. In the morning, both vehicles drove separately. During the afternoon session, the G4302 and Mira Ultra analyzer were transferred to the TNO truck to facilitate better comparison between the measuring devices and the UUAQ vehicle ceased its mobile measurements.

Two experiments in Utrecht were conducted at the Utrecht Science Park with multi-storey office and service buildings to the sides of the street (see SI, Fig. S3). On November 25, 2022, CH_4 was released simultaneously at two different locations. Two manual flowmeters (Krohne DK800/PV (25–250 NL/h) at location 1 and Krohne DK800/PV (500-5000 NL/h) at location 2) were used to measure the release rates (spanning from 2.18–15 $Lmin^{-1}$), which were controlled by the pressure reducer of the cylinder. CH_4 mole fractions were measured by the G2301 and G4302 devices, the same devices used during the Rotterdam campaign, on board the UUAQ car. The car was driving in a circle around two buildings, passing each emission point once per circle. On June 11, 2024 CH_4 was released simultaneously at the same site (0.15–100 $Lmin^{-1}$) and measured by the G2301 and Mira Ultra instrument installed in the van of the Institute for Marine and Atmospheric Research Utrecht (Maazallahi et al., 2020). Initially, the two release locations from the previous experiment were used. Due to power issues at location 1, the release point was moved across the street soon after the start. An Alicat device (high rates) and an MKS (PR 4000) controller (low rates) were used. Initially, the Alicat was at location 1 and the MKS at location 2. Midway, they were switched to enable all release rates at both sites.

Nearby London (on an open airfield near Bedford), two measurement campaigns were carried out, the first on September 10-13, 2019 (35 and 70 $Lmin^{-1}$) and the second on May 13 and 14, 2024 (0.2–70.5 $Lmin^{-1}$). A G2301 analyzer was used on all days in the first campaign, on September 10 additionally an Ultraportable Methane:Ethane Analyser and on September 11 a LI-7810 CH_4/CO_2/H_2O Trace Gas Analyzer were used. For the second campaign only data collected by the LI-7810 instrument were evaluated for this study. The driving pattern consisted of parallel legs perpendicular to the estimated wind direction, gradually moving away from the source.

In Toronto, two CREs were carried out. In the Toronto industrial port lands neighbourhood, both a mobile bicycle-trailer-based laboratory equipped with a UGGA analyzer (inlet at 1.6 m above ground), and a vehicle based setup measuring with a G2401 analyzer (inlet at 2.5m above the ground) were deployed on October 20, 2021 (2.5–20 $Lmin^{-1}$). The second experiment was carried out on October 24, 2021 on a parking lot, deploying the same vehicle-based set-up (0.1–5 $Lmin^{-1}$).

The average driving speed during plume transects was 3.8 ± 1.0 ms^{-1}, 5.6 ± 0.9 ms^{-1}, 5.9 ± 0.9 ms^{-1}, 3.0 ± 0.5 to 3.9 ± 0.4 ms^{-1}, 6.0 ± 1.5 to 7.3 ± 1.4 ms^{-1} and 4.0 ± 0.4 to 7.5 ± 0.9 ms^{-1} in Rotterdam, Utrecht I, Utrecht II, London I, London II and Toronto respectively. The median distance between the location where the plumes were detected and release location was 20 m, 20 m, 21 m, 20 to 25 m, 21 to 22 m and 17 to 24 m, which is typical for urban gas distribution networks (see SI, Sect. S6). More detailed descriptions of the measurement devices, experimental set-up and the CH_4 timeseries can be found in the Supplementary Information (SI, Sect. S1–S4), Maazallahi et al. (2020), Gillespie et al. (2023) and Vogel et al. (2024).

2.2 Data Treatment, Peak Identification, Determining Peak Maximum and Spatial Peak Area

The raw measurements were calibrated and corrected for inlet delay and a delay between different instruments (see SI, Sect. S3). A centred 5 minute moving window was applied to determine the atmospheric CH_4 background level at each point in time. The background level was defined as the 10th percentile of the CH_4 mole fraction measurements, which was assessed to represent the background well (comparison is given in SI, Sect. S5). A peak was identified as a CH_4 enhancement reaching above 102% of background level. CH_4 enhancements in the calibrated CH_4 dataset were detected utilizing the python *scipy*

function *find_peaks* (Virtanen et al., 2020). Those individual peaks were then manually quality-checked for overlap of peaks, flawless function of all instruments deployed, validity of the transect and car speed. When multiple instruments recorded measurements in one vehicle, the peak finding algorithm was applied to only one of the instruments. With a manual quality check it was ensured that the peak was valid for all instruments. For the London datasets, peaks obtained at a distance further than 75 m from the source were omitted to keep the distance within the same limits as for the other CREs. The maximum CH_4 enhancement within the time interval of each peak was determined for each instrument. The peak finder algorithm was applied to the G4302 device for the measurements on the UUAQ car in Rotterdam and in Utrecht I, the G2301 instrument for Utrecht II, the MGA10 device for the measurement on the TNO truck in Rotterdam, the uMEA and G2301-m device during the CRE in London I for Day 1 and Day 2 respectively. The peak area was integrated over space rather than time to take different driving speeds into account. To convert the time series to the spatial coordinate, the CH_4 mole fraction enhancement at time t_{i+1} ($c(t_{i+1})$, in ppm, after subtraction of the CH_4 background level) was multiplied with the measurement time step of the individual instrument ($t_{i+1} - t_i = \Delta t$) and the velocity of the vehicle averaged over the whole peak duration ($\bar{v}_{\text{peak}}$), yielding the integrated spatial peak area ($[CH_4]_{\text{area}}$) in $\text{ppm} * \text{m}$.

$$[CH_4]_{\text{area}} = \sum_{i=0}^{n} \Delta t \cdot c(t_{i+1}) \cdot \bar{v}_{\text{peak}} \tag{2}$$

It is important to note that this spatial peak area does not correspond to the integration of CH_4 enhancement of the physical CH_4 plume in the environment across a 2D plane in space. Rather it represents a linear 1D fraction of the plume, that is described by the driving track.

2.3 Emission Rate Estimation

An ordinary least-squares regression model (*scipy.stats.linregress* library, Virtanen et al. (2020)) was applied to the spatial peak areas of the combined dataset. The natural logarithm (ln) of the known release rate was used as the independent variable and the ln of integrated spatial peak area of CH_4 enhancements as the dependent variable.

$$\ln([CH_4]_{\text{area}}) = a_1 \ln(r_E) + a_0 \tag{3}$$

The fitting was performed using the entire dataset, without separating it into training and testing data. To assess the conformity of the data with assumptions underlying a linear regression (normality, linearity, independence and homoscedasticity), several analyses were carried out (SI, Sect. S7). A similar fit was applied to the maximum excess CH_4 ($\ln([CH_4]_{\text{max}}) = a_1^{\text{max}} \ln(r_E) + a_0^{\text{max}}$) for comparison to the equation from Weller et al. (2019).

To infer emission rate estimations based on measurements, the linear regression needs to be solved for the release rate. Then the equation can be applied to measurements (the Weller eq. to the peak maximum measurements, the Area eq. to the corresponding area measurements, for both cases background levels subtracted from CH_4 measurements). Sometimes the algorithm produces estimates far outside the calibration range of the method, therefore a cap of $200\ \text{Lmin}^{-1}$ was imposed on emission rate estimations for the following evaluation.

2.4 Evaluating Quantification Performance

2.4.1 Categorization

For each valid plume transect (i) in the dataset used to derive the regression model, emission rates were estimated utilizing both estimation methods, applying the inverse of Eq. 1 and Eq. 3.

$$r^i_{\text{E,Weller eq.}} = \exp\left(\frac{1}{a^W_1}\left(\ln([CH_4]_{\text{max}})^i - a^W_0\right)\right) = \exp\left(\frac{1}{0.817}\left(\ln([CH_4]_{\text{max}})^i + 0.988\right)\right) \tag{4}$$

$$r^i_{\text{E,Area eq.}} = \exp\left(\frac{1}{a^A_1}\left(\ln([CH_4]_{\text{area}})^i - a^A_0\right)\right) \tag{5}$$

The superscript W (Weller eq.) and A (Area eq.) differentiate the regression parameters. Subsequently, an estimated category was assigned to each peak given these inferred emission rates. To remain consistent with previous studies, release rates were classified into four different categories (< 0.5 Lmin^{-1}-Very low, $0.5 - 6$ Lmin^{-1}-Low, $6 - 40$ Lmin^{-1}-Medium and > 40 Lmin^{-1}-High). This approach follows von Fischer et al. (2017), but is extended by a category for very small emissions as used in Vogel et al. (2024). For each group of peaks belonging to the same category, the percentage of correctly classified peaks was calculated, along with the percentages of peaks that were erroneously categorized into other categories.

2.4.2 Percentage Difference in Emission Estimation as Function of Number of Transects

As will be shown below, variability in the plume shape causes large differences in observed peak maxima, spatial peak areas and thus derived emission rates for individual transects at the same actual emission rate. This variability can be reduced by evaluating the average of several transects at the same emission rate. Following the analysis in Luetschwager et al. (2021) the effect of number of detections per CH$_4$ source on variability in estimated emission rate was explored.

Two approaches were followed to evaluate the emission quantification:

1. Comparison against the mean emission rate calculated from the leak indications (following Luetschwager et al. (2021)).

2. Comparison against the true release rate, which is known for our experiments.

To calculate the mean emission rates for the former comparison, the average natural logarithm of the integrated spatial peak area among all observed instances associated with each release rate j was computed.

$$\text{Mean } \ln([CH_4]_{\text{area}})^j = \frac{1}{n}\sum_{i=1}^{n}\ln([CH_4]_{\text{area}})^i \tag{6}$$

For each release rate, one mean emission rate estimation $r^j_{\text{E,mean}}$ was obtained by applying the Area eq. to the calculated mean $\ln([CH_4]_{\text{area}})$. Subsequently, a Monte Carlo simulation was performed where we randomly selected between 2 and 10 emission peaks i at each release rate, and averaged the natural logarithm of the integrated spatial peak areas. Measurements obtained by different instruments during the same transects were treated as separate peaks. The theoretical number of possible subsets N from a certain size of set M with replacement is $\binom{M+N-1}{N} = \frac{(M+N-1)!}{N! \cdot (M-1)!}$. This gives for example between 465 and $6 \cdot 10^7$

combinations for N between 2-10 and $M = 30$. This procedure was repeated 2000 times for each release rate and number of detections. That means for $N = 3$, three peaks were randomly sampled 2000 times from a given release rate, yielding 2000 emission rate estimations $r_{\mathrm{E,sim}}^{k,j,N}$ ($k \in [1, 2000]$) for each release rate. For each of those 2000 emission rate estimations $r_{\mathrm{E,sim}}^{k,j,N}$ the percentage difference to the mean emission rate estimation $r_{\mathrm{E,mean}}^{j}$ and the known release rate $r_{\mathrm{E,true}}^{j}$ were calculated:

$$
210 \quad \text{Percentage Deviation } \Delta\%_{\mathrm{mean}}^{k,j,N} = \frac{r_{\mathrm{E,sim}}^{k,j,N} - r_{\mathrm{E,mean}}^{j}}{r_{\mathrm{E,mean}}^{j}} \cdot 100\% \tag{7}
$$

$$
\text{Percentage Deviation } \Delta\%_{\mathrm{true}}^{k,j,N} = \frac{r_{\mathrm{E,sim}}^{k,j,N} - r_{\mathrm{E,true}}^{j}}{r_{\mathrm{E,true}}^{j}} \cdot 100\% \tag{8}
$$

Then, an average percentage deviation for each release rate j and number of transects N was determined for both cases.

$$
\text{Mean Percentage Deviation } \overline{\Delta\%}^{j,N} = \frac{1}{2000} \sum_{k=1}^{2000} \Delta\%^{k,j,N} \tag{9}
$$

Finally, an average over the different release rates J for each number of transects $N \in [2, 10]$ was calculated for both cases.

$$
215 \quad \text{Overall Mean Per } N = \frac{1}{J} \sum_{j=1}^{J} \overline{\Delta\%}^{j,N} \tag{10}
$$

The analysis was done for each unique release rate and location pair. Experiments with fewer than 10 transects were filtered out, as the Monte Carlo analysis samples up to 10 measurements (N ranging from 2 to 10) from the available transects. This leaves 35 of the original 55 release rates. Lastly, another categorization analysis was conducted with the goal to investigate the classification performance when incorporating multiple transects. For each of the 2000 mean emission rates per release rate

220 and number of transects, a category was assigned and evaluated.

3 Results and Discussion

In this section, we present and analyze the key findings of our study. First, we evaluate instrument performance by comparing peak maximum and spatial peak area as metrics for emission quantification (Sect. 3.1). Next, we introduce a method for converting spatial peak areas into emission rate estimates (Sect. 3.2). We then explore how survey data can be used to support

225 repair prioritization (Sect. 3.3). Additionally, we assess the benefits of multiple transects in improving emission estimates (Sect. 3.4). Finally, we propose a method for the routine implementation of leak surveys based on our findings (Sect. 3.5).

3.1 Instrument Performance: Peak Maximum and Spatial Peak Area

Fig. 1a shows that the comparison of measured peak maxima among different instruments reveals very strong systematic discrepancies. Compared to the G4302, all other instruments (Mira Ultra, G2301, MGA10, and TILDAS) strongly underestimate

230 the peak maximum. Specifically, the maxima recorded by the TILDAS are only half of those measured by the G4302, while the Mira Ultra, G2301, and MGA10 show peak maxima that typically are only between 10% to 30% of the G4302 readings. Fig. 1b

shows that the evaluated spatial peak areas are much more consistent between instruments, with slopes between 0.44 and 1.01. The coefficient of determination R^2 is notably higher for the area fit when compared to the maximum fit for all instruments. R^2 for G2301_U, Mira Ultra, and TILDAS exceed 0.96. The MGA10 device stands out as an exception, as it demonstrates poor R^2 values for both maximum and area fits. Both measured maximum and area values are more scattered and exhibit strong deviations from the other instruments. Interestingly, the G2301 analyzer aligns better with the G4302 for data collected in Utrecht than in Rotterdam (indicated by a higher slope). This reflects the better alignment of instrument measurements for smaller peaks; in Rotterdam, very high release rates were included. A similar pattern arises for the evaluation of the CRE in London I (SI, Fig. S21). Both the uMEA and LI-7810 instrument measurements align more closely with G2301 measurements when assessing spatial peak area rather than peak maximum.

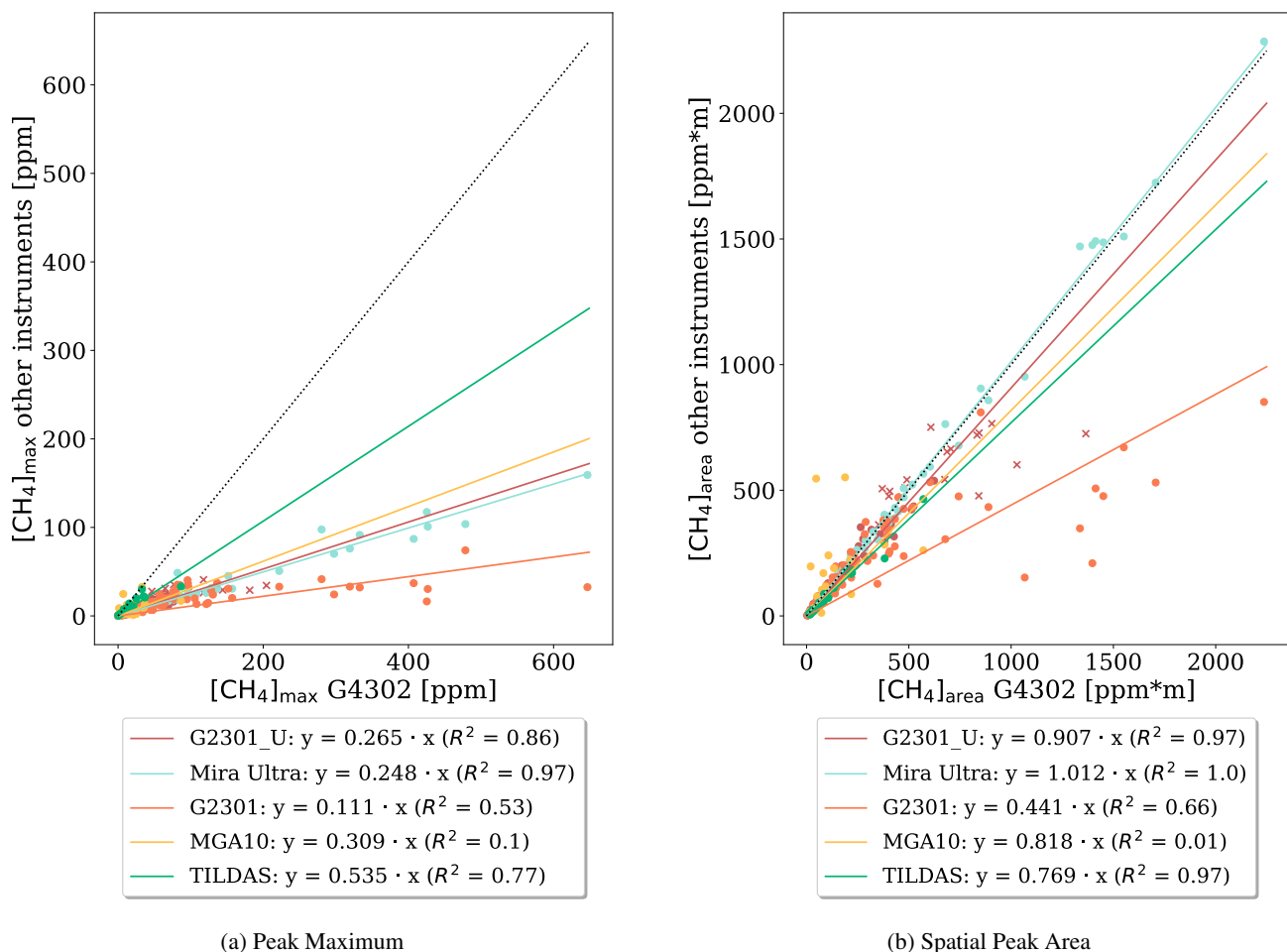
(a) Peak Maximum

(b) Spatial Peak Area

Figure 1. Comparison of peak maximum (a) and spatial peak area (b) from different instruments deployed in Rotterdam and Utrecht I (subscript '_U'), shown are the data points and a linear regression fit with intercept 0 for each instrument. The results from the G2301, Mira Ultra, MGA10 and TILDAS devices are plotted on the y-axis and the results from the G4302 instrument on the x-axis. The black dotted line represents the 1:1 line. (For the G2301 analyzer, peaks exceeding a maximum of 20 ppm are marked with an 'x' and excluded from the fitting process.)

Our analysis demonstrates that different CH_4 analyzers show much better consistency in the integrated CH_4 spatial peak area compared to the maximum CH_4 enhancement. This finding is supported by conceptual considerations: Measured mole fractions are influenced by instrument characteristics, mainly flow rate, cell volume, and measurement frequency (Fig. 2). A higher flow rate or smaller cell volume will generally make measured mole fraction profiles sharper and higher. Lower flow rates and larger cell volumes, on the other hand, lead to smoothing effects, rendering the CH_4 peak smaller and wider (Takriti et al., 2021; Maazallahi et al., 2023). While the peak maximum is strongly affected by these characteristics, the integrated area is not. Each molecule that enters an instrument is measured, either as part of a wide and low or a narrow and high peak. Gillespie et al.

245

(2023) and Maazallahi et al. (2023) made similar observations when comparing different analyzers. In this study, we generalize these results by demonstrating the same effect across eight different instruments, enhancing the robustness of this conclusion.

250 Beyond differences in instrument performance, Ars et al. (2020) noted the area to be more robust against differences in inlet height, further supporting the argument for using the integrated spatial peak area as a more reliable metric for quantifying CH_4 enhancements. In future mobile surveys, several different instruments are expected to be deployed to survey local natural gas distribution networks. It is therefore important to move beyond the maximum enhancement as emission estimation metric and use the spatial peak area instead.

(a) Flow Rate and Cell Volume

(b) Measurement Frequency

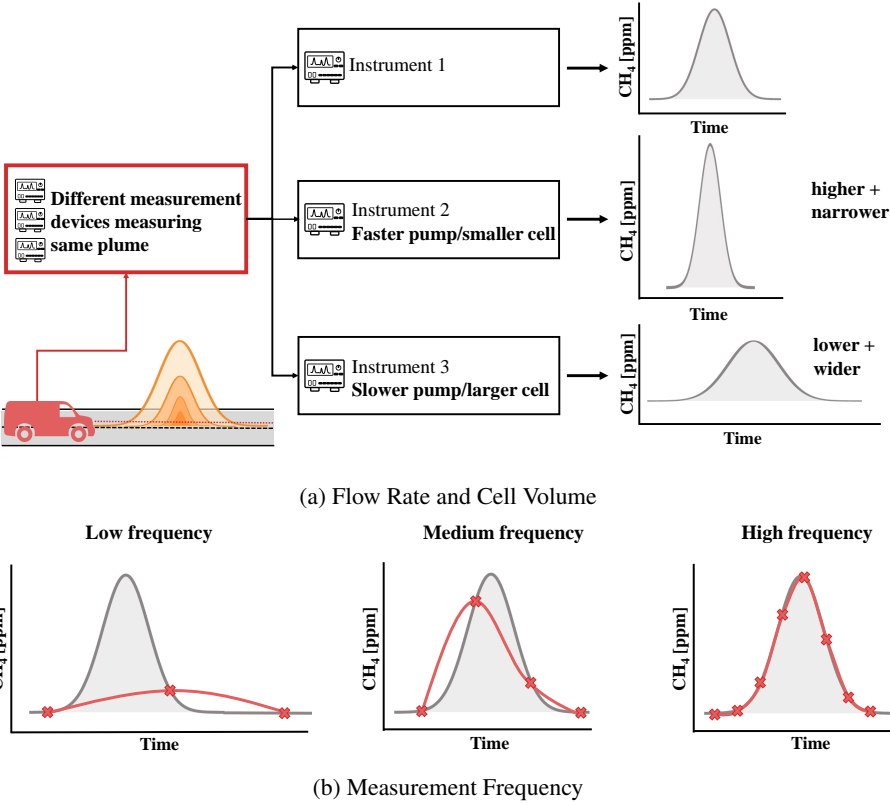



Figure 2. Influence of the instrument characteristics (a) flow rate, cell volume (in dependence of cell temperature and pressure) and (b) measurement frequency on the peak shape detected.

255 ## 3.2 Converting Spatial Peak Areas to Emission Rates

Fig. 3b shows the recorded spatial peak areas as a function of the known release rates (double logarithmic axes) from all controlled release experiments. As anticipated, higher CH_4 release rates in general correspond to higher observed spatial peak area measurements. A linear fit can be applied to these data and the fit equation is proposed as an empirical model relating integrated spatial peak area of CH_4 enhancements ($[CH_4]_{area}$, in $ppm * m$, background levels subtracted from CH_4

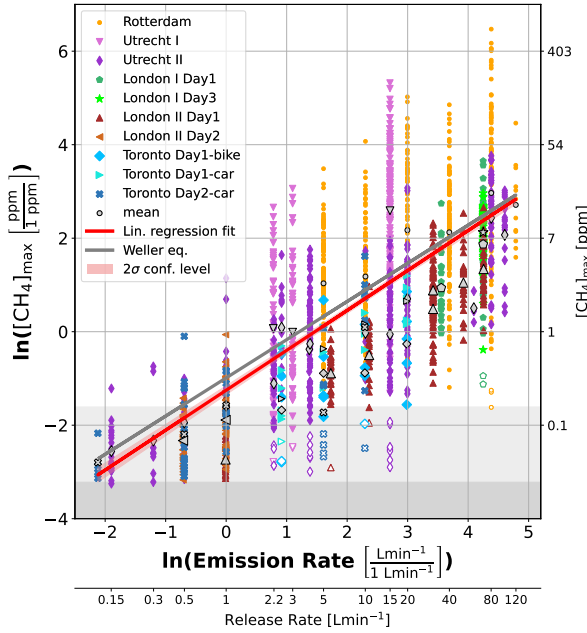
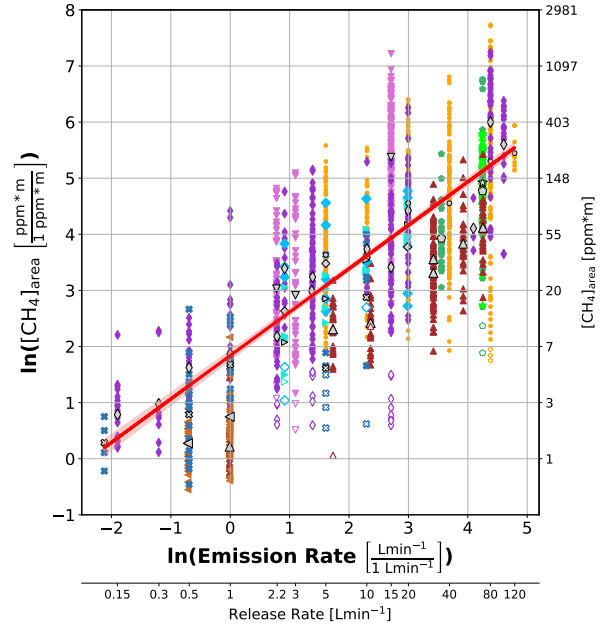



Figure 3. Correlation of the natural logarithm of the peak maximum enhancement (a) and spatial peak area (b) with the natural logarithm of the release rates for all controlled release experiments reported in this manuscript (except London I Day 2). Different cities are indicated by different colours (Rotterdam-orange, Utrecht I-pink, Utrecht II-dark purple, Toronto-blue, London I-green and London II-brown). Black markers indicate mean values per release rate and city, unfilled markers indicate potential outliers. The second x-axis indicates release rates deployed, the red lines are linear regressions to all data. The Weller equation (grey line in (a)) is displayed as a comparison and the light (dark) gray area indicates peaks below 110% (102%) of background level.

measurements) to emission rates (r_E, in $\mathrm{Lmin^{-1}}$):

$$\ln([\mathrm{CH_4}]_{\mathrm{area}}) = 0.774 \cdot \ln(r_E) + 1.84. \tag{11}$$

In practice, emission rate estimations will be derived from area measurements. To achieve this, Eq. 11 can be solved for r_E.

$$r_E = \exp\left(1.292 \cdot \ln([\mathrm{CH_4}]_{\mathrm{area}}) - 2.377\right) \tag{12}$$

The natural logarithm of the spatial peak area associated with a leak indication can then be inserted in the equation. If multiple transects were taken, the average logarithm of the spatial peak area values $\overline{\ln([\mathrm{CH_4}]_{\mathrm{area}})}$ corresponding to the same leak indication should be used.

A similar linear regression fit was applied to peak maxima data: $\ln([\mathrm{CH_4}]_{\mathrm{max}}) = 0.854 \cdot \ln(r_E) - 1.25$ (Fig. 3a). The inferred regression equation for the maximum shows good agreement with the model proposed by Weller et al. (2019).

3.3 Using Survey Data for Repair Prioritisation

270 A wide range of peak maxima and areas is observed for each release rate. This illustrates the shortcomings of the quantification of emissions for individual leak indications using this simple statistical method. This spread reflects the nature of turbulent dispersion of plumes. Environmental factors, in particular built environment and meteorology, e.g. convection, complex wind patterns within urban areas (back circulation, wind channelling, blockages), turbulence and diffusion, play a major role in determining the location, shape and CH_4 mole fraction of the meandering CH_4 plume (Carpentieri and Robins, 2015; Cassiani

275 et al., 2020; Ražnjević et al., 2022). Still, a greater CH_4 release increases the likelihood of detecting higher CH_4 mole fractions. This is the relation that is captured in the proposed transfer equation. Yet, changing wind conditions or turbulence can still cause the main plume to become highly diluted or be transported away from the street, leading to smaller peak maxima and spatial peak areas. At the extremes, high CH_4 mole fractions are measured in one transect, and no peak at all in another one (Luetschwager et al., 2021). Ideally, measurements should be conducted downwind of the emission source and perpendicular to

280 the wind. This approach is commonly used when targeting individual emitters, such as oil and gas processing stations or farms (Hensen et al., 2006; Korbeń et al., 2022). However, when measuring emissions from urban gas distribution networks, the built environment imposes constraints on positioning. Incorporating wind measurements from the survey vehicle could potentially improve emission estimates and should be explored in future research.

Additionally, some assumptions underlying the application of a linear regression are partially violated (homoscedasticity,

285 normal distribution of errors). However, the premises were deemed sufficiently met to justify the use of linear regression, allowing for a straightforward and practical statistical inference model. However, as all experiments included in this study were executed during daytimes (measurements were conducted between 09:00-18:00 local time), the method might not be suitable for the evaluation of nighttime measurements. During the night, the atmospheric boundary layer becomes more stable, which suppresses updrafts and leads to an accumulation of CH_4 emissions at the surface. This causes higher measured mole fractions,

290 increasing both the peak maximum and spatial peak area measurements; hence most likely considerable overestimation of the emission rate.

Similarly, though to a lesser extent, variations in solar irradiation during the day can influence updraft behavior. Around noon, stronger solar irradiation is expected to enhance updrafts, which could reduce ground-level methane concentrations. Only a few release rates were deployed twice on the same day, preventing a conclusive analysis of measurement comparisons

295 at different times. We recommend that future experiments investigate diurnal effects more systematically.

(a) Weller eq., based on Peak Maximum (b) Area eq., based on Spatial Peak Area

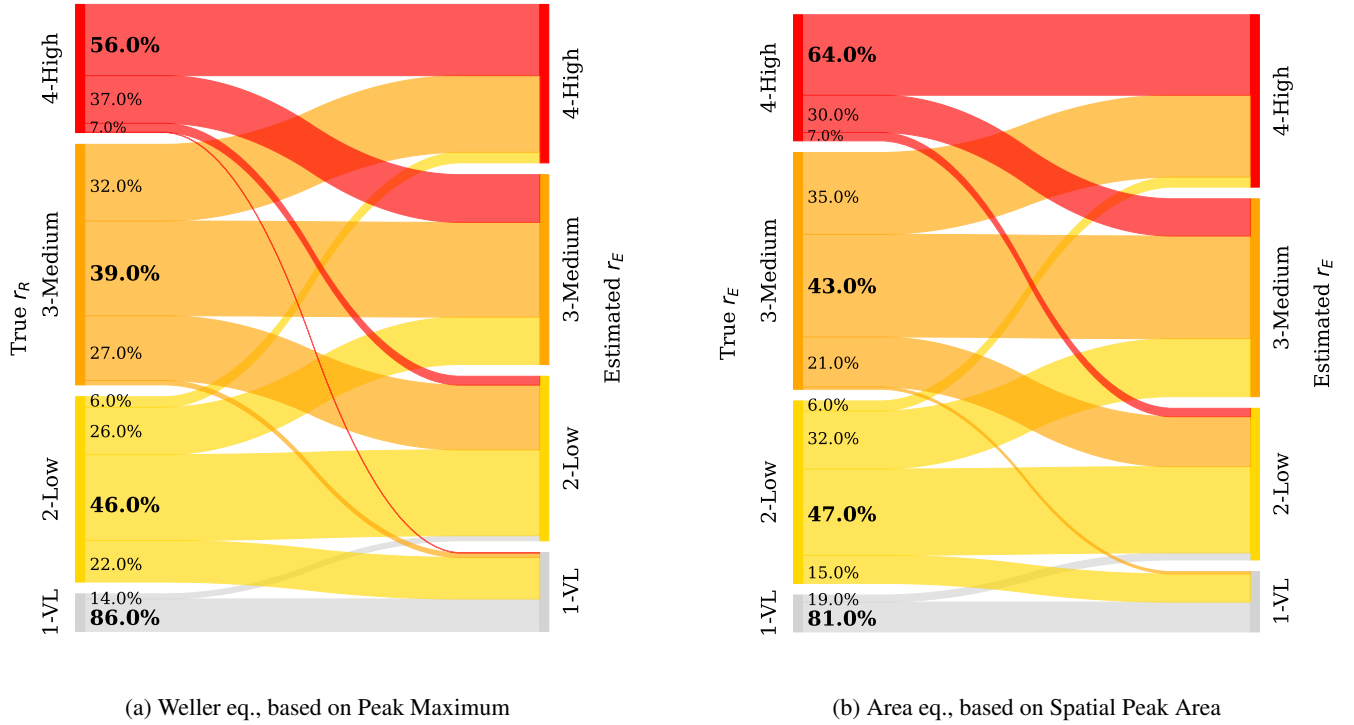



Figure 4. Categorization performance of (a) the Weller and (b) Area method (including data from all six controlled release experiments). The left y axis represents the true emission rate r_E, where the width of the bars indicate the amount of plumes belonging to each emission category (categories: 1-Very Low, 2-Low, 3-Medium and 4-High). The right y axis represents the categories estimated by the statistical model and the connecting lines visualize the amount of plumes (in %) from each category pool which the algorithm classifies into another (or the same) category.

Deviation of emission rate estimation from true emission rate can be very high for an individual gas leak, especially for single transects. For our set of experiments under- and overestimation range from -100% to +2700%. To facilitate a pragmatic repair priorization, categorizing leaks can be an important tool. We quantified the emission rate of all leak indications and subsequently assigned them to one of the four categories described above. Fig. 4 illustrates the categorization performance based on quantification using the spatial peak area (a) and peak maximum (b). It displays the percentage of peaks correctly classified (bold number) and also shows in which categories the rest is falsely overestimated or underestimated. A suitable benchmark for categorization is a 25% accuracy rate, reflecting the expected success when peaks are randomly assigned to one of the four categories.

The majority of very low and high peak indications are correctly attributed. Peaks in the medium category are less well classified, with a large portion being either over- or underestimated. Overall, the proportion of peaks that are misclassified by more than one category is very low. Around 7% (22%) of category 4 (3) peaks are underestimated into category 1 or 2, when applying the area eq. This means that the vast majority (78% to 94%) of higher emission peaks from categories 3 and 4 are

correctly identified as medium–high, effectively differentiating them from low emission peaks. At the same time, 38% (0%) of category 2 (1) emission rates are falsely identified as high emission rates. The same analysis using the Weller eq. yields comparable results, with a slightly lower categorization performance except for the lowest emission category (Fig. 4a).

It is important to note that these categorization performance rates vary across our set of release locations and are therefore not directly transferable to any leak populations in urban areas. The quantification and categorization performance differs for different locations, for example, the categorization precision for the London I data is 62% for category 4, but 88% for category 3, much higher than for the average (SI, Fig. S22 and S23). Given the local circumstances regarding built environment and meteorological conditions, peak measurements can be systematically lower or higher compared to the mean of our dataset. However, the numbers can still be considered as an indication of typical categorization performance.

In von Fischer et al. (2017), a categorization success rate of over 80% for category 2 and 3 was reported, which is higher than for our dataset. For the highest category, their performance was lower (38%) than ours (64%). The differences might be due to the specific local conditions where their controlled release took place. In practice, $> 80\%$ of urban leaks have a low emission rate (von Fischer et al., 2017; Maazallahi et al., 2020), category 1 was not assessed in those publications).

Overall, our results show that the categorization approach for leak repair prioritization would perform far better than making repair decision by chance. Therefore, this approach offers a potential to reduce emissions since larger leaks can be targeted first. The Area eq. does not consistently outperform the Weller model in terms of categorization performance, but it yields more consistent results across different instruments, which would facilitate comparison of measurement surveys performed with different instruments.

3.4 Benefit of Multiple Transects

The discussed high uncertainty of individual estimations of the emission rate suggests that the performance may be improved by carrying out several transects. Including more measurements will reduce the random error associated with measurements and increase the confidence to capture the mean of this particular distribution (Cochran, 1977). This effect has been demonstrated in Luetschwager et al. (2021) and is illustrated in Fig. 5a. There, the Monte Carlo mean of the absolute percentage deviation of emission estimation, based on N transects ($N \in [2, 10]$), from the calculated mean emission rate, based on all M transects measured for this release rate, is shown. The percentage deviation from the mean decreases drastically with an increasing number of transects, and converges to 0 for $N = M$ (not shown). Note that here the absolute percentage error is displayed, no difference is made between a negative or positive deviation.

The observation that many of the mean values for individual release rates are still considerably different from the value expected from our conversion equation (Fig. 3) suggests that in addition to random error, there is also a systematic error. This is likely due to two factors; the offset of the sample mean from the population mean (determined by longer time scale weather phenomena) and the offset of the population mean from our linear regression (determined by (built) environment).

Therefore, it follows that the error with respect to the true release rate will not necessarily decrease with more measurements for mean values that are far apart from the regression line. This is the case for the data from London II experiment, which exhibit a large negative offset as well as for the Utrecht I 15 Lmin^{-1} release, which exhibits a very high positive offset (see

SI, Sect. S10 for a more detailed discussion). Still, if the offset is not too large, the error in emission estimation decreases with more transects. This is visualized in Fig. 5b, where the Monte Carlo mean of the absolute percentage deviation of emission estimation in respect to the true emission rate is shown.

345   The largest decrease in emission estimation error can be achieved from 1 to 3 transects, reducing the estimation error by one third. A further decrease can be reached until around 5-7 transects are included, after which the gained variation reduction levels off, including more transects then reduces the uncertainty by less than 3%.

(a) Calculated mean emission rate   (b) True release rate

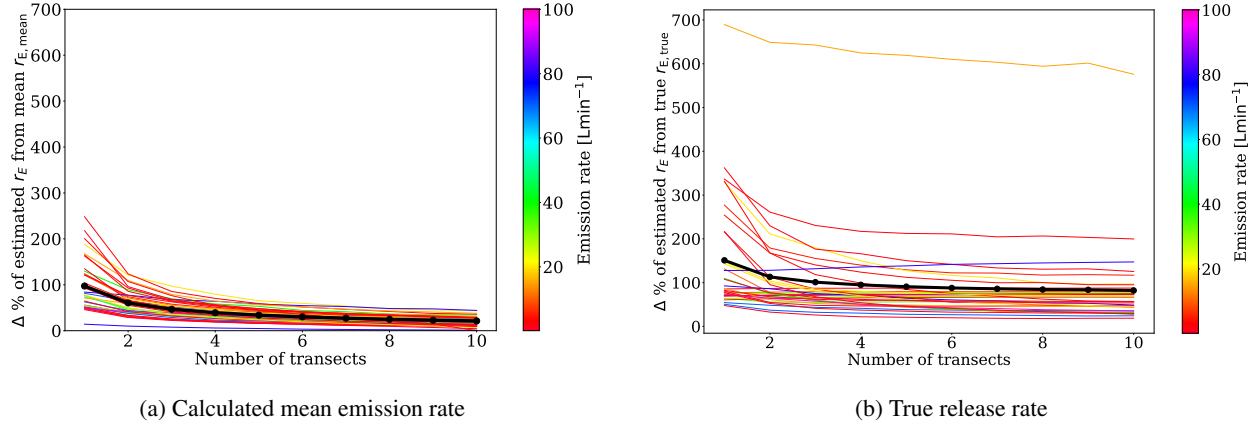



Figure 5. Mean absolute percentage deviation of emission estimation from the (a) calculated mean emission rate and (b) true release rate as a function of the number of transects. Each line represents one release rate at one location, and was created by a Monte-Carlo analysis with 2000 repetitions. The total mean over all release rates considered in this analysis is shown in black.

Luetschwager et al. (2021) carried out a similar analysis on multiple passages of real-world leak locations, comparing individual emission rate estimate to the emission rate estimate based on the distribution mean. They reported the sharpest
350   decline in variation with increased detections from 2 to 4, aligning with our analysis. They also noted that, beyond 5-6 transects, the reduction in variability diminishes. Additionally, Luetschwager et al. (2021) reported errors relative to the mean of 50% deviation based on 2 transects. This compares well with the variability in our corresponding analysis, which amounts to an average of 62% for $N = 2$ comparing to the mean (Fig. 5a). Nevertheless, in comparison to true release rates from real measurements, the overall deviations are higher than in the aforementioned studies. We find the mean percentage deviation
355   from the true release rate to be 115% for $N = 2$ (Fig. 5b). Further, the deviations are skewed, with higher overestimations than underestimations, a result of the right-skewed distribution of measured spatial peak area values.

Ražnjević (2023) stated that at least 10 transects are necessary to estimate the source strength within 40% of the true emission, simulating a mobile measurement set-up in an Large Eddy Simulation model. We found for half of the cases an estimation error within 58% when including 10 transects, for the other half errors ranged between 65-200% with one outlier at
360   over 500%. Thus, our findings are in closer agreement with Ražnjević (2023) than with the smaller estimation errors reported in Luetschwager et al. (2021). The generally higher estimation errors demonstrate the complexity of real world environmental influences.

The persistent estimation error, even with the inclusion of more transects, likely reflects the influence of the built-environment and meteorological conditions. If there is a systematic bias for that specific location or time frame of measurements, combining several transects will still include that high or low bias. Additional measurements can only balance out random errors or fluctuations. Taking measurements under different meteorological conditions could potentially improve estimations, enabling sampling of the full distribution of possible realizations of the emission plume.

(a) $N = 1$ (b) $N = 3$ (c) $N = 6$

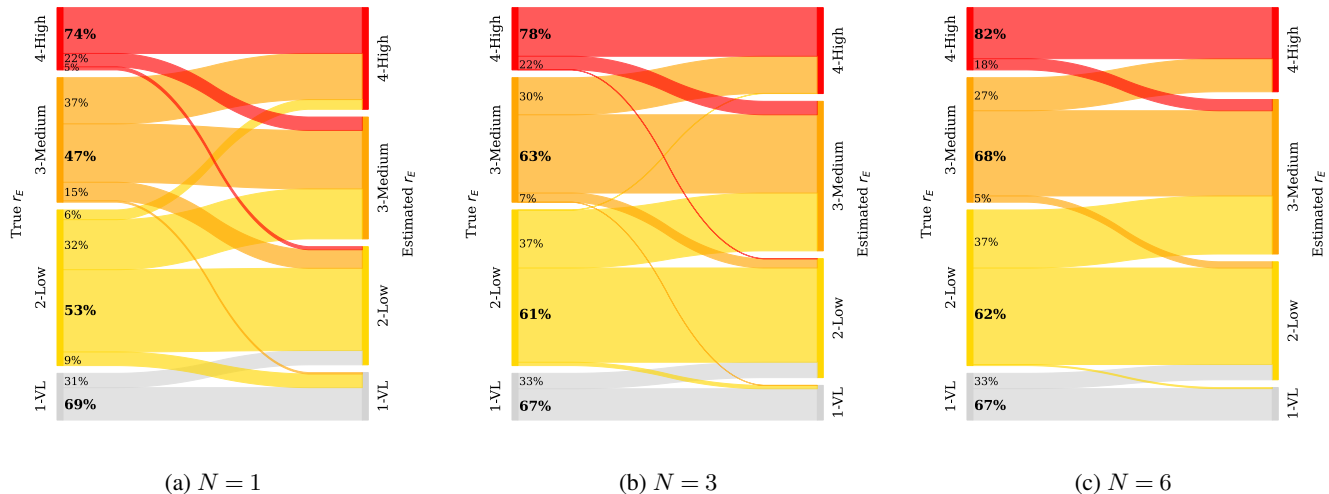



Figure 6. Categorization performance of the area equation with (a) $N = 1$, (b) $N = 3$ and (c) $N = 6$ transects for distributions with a small to medium offset (25 of the 35 release rates). The left axis represents the true r_E, where the width of the bars indicate the amount of plumes belonging to each emission category (categories: 1-Very Low, 2-Low, 3-Medium and 4-High). The right axis represent the categories estimated by the statistical model and the connecting lines visualize the amount of plumes from each category pool which the algorithm classifies into another (or the same) category. The underlying dataset is based on the Monte-Carlo simulation resulting in 2000 estimates per release rate.

Fig. 6 illustrates that the accuracy of categorization on average improves when taking into account one (Fig. 6a), three (Fig. 6b) or six (Fig. 6c) transects. The London II and Utrecht I 15 Lmin^{-1} releases are not included in this analysis because they include strong systematic biases. For emission rate estimations based on one single transect ($N = 1$), the categorization performance ranges between 47% and 74%, averaging 58%. When three transects are used ($N = 3$), the categorization performance improves to an average of 65%. Including six transects ($N = 6$) further enhances performance to an average of 69%, ranging from 62% to 82%. More importantly, the number of strong over or underestimations decreases. For example, whereas for only one transect 5% of category 4 peaks and 15% of category 3 peaks are categorized as category 2, this reduces to 0% and 5% respectively for 6 transects. Also the false overestimation of 6% of category 2 peaks into category 4 vanishes when increasing sampling effort to 6 transects.

3.5   Suggested Method for Routine Implementation of Leak Surveys

We suggest the following approach for routine mobile leak detection surveys in urban areas:

1. Survey: Perform surveys with an instrument that falls under the '1-1-10 principle' - at least 1 Hz measurement frequency, 1 s residence time and 10 ppb precision. Record GPS data together with the mole fraction measurements, ideally with the same measurement frequency as the CH_4 analyzer. Ensure that the GPS and CH_4 recordings are time-synchronized. Measure and record the inlet delay time for the CH_4 measurements at the beginning of the survey. Optimally, gather wind measurements as well, such as with an anemometer mounted on the vehicle's roof. If possible, maintain a sufficient distance from previously passed cars while conducting the measurements. During the survey, if an enhancement above the CH_4 background exceeds 0.2 ppm, attempt to conduct six transects downwind of the leak indication.

2. Evaluation: After the survey campaign, determine CH_4 background levels and subtract them from the CH_4 measurements (CH_4 background at a certain time = 10th percentile of CH_4 mole fractions within a ± 2.5 min time window). Use a peak detection algorithm to identify peaks in the time series, including start and end time of each peak. If available, use additional information such as ethane measurements, methane : carbon dioxide ratio or CH_4 isotopic composition to differentiate between biogenic, combustion and fossil sources. Use GPS data to infer the driving speed to convert the time series to space coordinates and calculate spatial peak areas (using the CH_4 elevation above background levels). In case several transects were done for the same leak indication, take the average of the $\ln([CH_4]_{area})$ of the individual peaks detected during the different transects. Calculate the emission rate by inserting the $\ln([CH_4]_{area})$ value into the inverse Area eq (Eq. 12).

3. Categorization: Rank the leak into one of the four emission categories (< 0.5 Lmin^{-1}-Very low, $0.5 - 6$ Lmin^{-1}-Low, $6 - 40$ Lmin^{-1}-Medium and > 40 Lmin^{-1}-High) and use this categorization in addition to safety concerns to prioritize leak repair efforts.

4. Report: Report the number of leaks per category and number of leaks per pipeline kilometre. Report the leak size distribution and total emissions for the surveyed area. Ideally, provide these numbers separately for different pipeline material (e.g. cast iron, steel, polyethylene and P.V.C.).

4   Conclusions

Our results clearly highlight systematic differences in maximum CH_4 enhancement recorded by different commercial CH_4 gas analyzers. Maximum enhancements differ by up to a factor ten between instruments. When the spatial peak areas are evaluated, differences between instruments reduce to less than a factor two, confirming our hypothesis that the spatial peak area is a more suitable quantity than the recorded peak maximum. We suggest moving beyond the peak maximum as emission estimation metric to avoid biased emission predictions.

We propose an empirical equation for the emission rate (r_E, in Lmin^{-1}) based on integrated spatial peak area of CH_4 enhancements ($[CH_4]_{\mathrm{area}}$, in $\mathrm{ppm} * \mathrm{m}$, background levels subtracted from CH_4 measurements):

$$r_E = \exp\left(1.292 \cdot \overline{\ln([CH_4]_{\mathrm{area}})} - 2.377\right) \tag{13}$$

410   Here, $\overline{\ln([CH_4]_{\mathrm{area}})}$ is the mean taken over several transects associated with the same leak indication. This formula is very simple in deployment, it does not require any additional information than the actual CH_4 mole fraction and GPS measurements. However, it therefore ignores other important local influences. The most important parameters are likely the built environment, wind speed and direction, atmospheric stability and turbulent processes. It is expected that incorporating parameters reflecting those factors could improve emission rate predictions. However, especially when it comes to incorporating turbulent processes,

415   the associated measurement effort would become extensive. Therefore, the strength of our simple statistical model is the simplicity of its implementation, still providing reasonable emission estimates. Higher sampling effort can improve quantification accuracy significantly. Including three transects instead of one improved correct categorization into four emission categories from 47%-74% to 61%-78% and to 62%-82% when incorporating six transects instead of one transect. Thus, applying this method to identify, quantify and prioritize leak repairs in urban natural gas distribution systems can support greenhouse gas

420   emission mitigation during the transition to a fossil-fuel-free energy system.

Code and data availability. The python code and a sub-sample of the data used to produce the results in this article are available on GitHub: https://github.com/judith-tettenborn/CRE_CH4Quantification.git

Author contributions. Contributed to conception and design: T.R., D.Z.-A., H.M.

Contributed to acquisition of data: J.T., D.S, H.M, C. vd V., A.H., I.V., P. vd B., F.V., L.G., S.A., J.F., D.L., R.F., T.R.

425   Contributed to analysis and interpretation of data: J.T., D.Z.-A., D.S., H.M., A.H., J.F., T.R.

Drafted and/or revised the article: J.T., D.Z.-A., H.M, A.H., I.V., F.V., L.G., S.A., J.F., D.L., R.F., T.R.

Approved the submitted version for publication: J.T., D.Z.-A., D.S, H.M, C. vd V., A.H., I.V., P. vd B., F.V., L.G., S.A., J.F., D.L., R.F., T.R.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques.

430   *Disclaimer.* Judith Tettenborn was supported through a grant from Environmental Defense Fund.

Acknowledgements. We thank all who contributed to data acquisition during the measurement campaigns across various research groups. Special thanks to Roberto Paglini from Politecnico di Torino and Ceres Woolley-Maisch from Utrecht University for their dedicated efforts during the campaign, assistance with data analysis, and valuable discussions.

435   In the drafting and programming of this publication the AI tool ChatGPT (https://chat.openai.com/) has been utilized as aid based upon initial drafts. Every response stemming from the AI has been checked, evaluated and only implemented with care.

References

Ars, S., Vogel, F., Arrowsmith, C., Heerah, S., Knuckey, E., Lavoie, J., Lee, C., Pak, N. M., Phillips, J. L., and Wunch, D.: Investigation of the Spatial Distribution of Methane Sources in the Greater Toronto Area Using Mobile Gas Monitoring Systems, Environmental Science & Technology, 54, 15 671–15 679, https://doi.org/10.1021/acs.est.0c05386, 2020.

Carpentieri, M. and Robins, A. G.: Influence of Urban Morphology on Air Flow over Building Arrays, Journal of Wind Engineering and Industrial Aerodynamics, 145, 61–74, https://doi.org/10.1016/j.jweia.2015.06.001, 2015.

Cassiani, M., Bertagni, M. B., Marro, M., and Salizzoni, P.: Concentration Fluctuations from Localized Atmospheric Releases, Boundary-Layer Meteorology, 177, 461–510, https://doi.org/10.1007/s10546-020-00547-4, 2020.

Cochran, W. G.: Sampling Techniques, Johan Wiley & Sons Inc., ISBN 0-471-16240-X, 1977.

Collins, W. J., Webber, C. P., Cox, P. M., Huntingford, C., Lowe, J., Sitch, S., Chadburn, S. E., Comyn-Platt, E., Harper, A. B., Hayman, G., and Powell, T.: Increased Importance of Methane Reduction for a 1.5 Degree Target, Environmental Research Letters, 13, 054 003, https://doi.org/10.1088/1748-9326/aab89c, 2018.

Cooper, J., Balcombe, P., and Hawkes, A.: The Quantification of Methane Emissions and Assessment of Emissions Data for the Largest Natural Gas Supply Chains, Journal of Cleaner Production, 320, 128 856, https://doi.org/10.1016/j.jclepro.2021.128856, 2021.

Cui, Y. Y., Henze, D. K., Brioude, J., Angevine, W. M., Liu, Z., Bousserez, N., Guerrette, J., McKeen, S. A., Peischl, J., Yuan, B., Ryerson, T., Frost, G., and Trainer, M.: Inversion Estimates of Lognormally Distributed Methane Emission Rates From the Haynesville-Bossier Oil and Gas Production Region Using Airborne Measurements, Journal of Geophysical Research: Atmospheres, 124, 3520–3531, https://doi.org/10.1029/2018JD029489, 2019.

Defratyka, S. M., Paris, J.-D., Yver-Kwok, C., Fernandez, J. M., Korben, P., and Bousquet, P.: Mapping Urban Methane Sources in Paris, France, Environmental Science & Technology, 55, 8583–8591, https://doi.org/10.1021/acs.est.1c00859, 2021.

Environment and Climate Change Canada: Faster and Further: Canada's Methane Strategy, https://www.canada.ca/en/services/environment/weather/climatechange/climate-plan/reducing-methane-emissions/faster-further-strategy.html, 2022.

EU Commission: Launch by United States, the European Union, and Partners of the Global Methane Pledge to Keep 1.5C within Reach, https://ec.europa.eu/commission/presscorner/detail/en/statement_21_5766, 2021.

EU Commission: Commission Welcomes Deal on First-Ever EU Law to Curb Methane Emissions in the EU and Globally, https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5776, 2023.

Fernandez, J. M., Maazallahi, H., France, J. L., Menoud, M., Corbu, M., Ardelean, M., Calcan, A., Townsend-Small, A., van der Veen, C., Fisher, R. E., Lowry, D., Nisbet, E. G., and Röckmann, T.: Street-Level Methane Emissions of Bucharest, Romania and the Dominance of Urban Wastewater., Atmospheric Environment: X, 13, 100 153, https://doi.org/10.1016/j.aeaoa.2022.100153, 2022.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. 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 [Masson-Delmotte, V., 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, and B. Zhou (Eds.)], pp. 923–1054, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.009, 2021.

Gillespie, L. D., Ars, S., Williams, J. P., Klotz, L., Feng, T., Gu, S., Kandapath, M., Mann, A., Raczkowski, M., Kang, M., Vogel, F., and Wunch, D.: A Modified Gaussian Plume Model for Mobile in Situ Greenhouse Gas Measurements, Atmospheric Measurement Techniques Discussions, pp. 1–26, https://doi.org/10.5194/amt-2023-193, 2023.

475 Hensen, A., Groot, T. T., van den Bulk, W. C. M., Vermeulen, A. T., Olesen, J. E., and Schelde, K.: Dairy Farm CH4 and N2O Emissions, from One Square Metre to the Full Farm Scale, Agriculture, Ecosystems & Environment, 112, 146–152, https://doi.org/10.1016/j.agee.2005.08.014, 2006.

Hu, A., Xu, Y., Tebaldi, C., Washington, W. M., and Ramanathan, V.: Mitigation of Short-Lived Climate Pollutants Slows Sea-Level Rise, Nature Climate Change, 3, 730–734, https://doi.org/10.1038/nclimate1869, 2013.

480 International Energy Agency: Global Methane Tracker 2023, https://www.iea.org/reports/global-methane-tracker-2023, 2023.

IPCC: Summary For Policymakers, in: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Team, C. W., Lee, H., and Romero, J., pp. 1–34, IPCC, Geneva, Switzerland, https://doi.org/10.59327/IPCC/AR6-9789291691647.001, 2023.

Irakulis-Loitxate, I., Guanter, L., Liu, Y.-N., Varon, D. J., Maasakkers, J. D., Zhang, Y., Chulakadabba, A., Wofsy, S. C., Thorpe, A. K.,
485 Duren, R. M., Frankenberg, C., Lyon, D. R., Hmiel, B., Cusworth, D. H., Zhang, Y., Segl, K., Gorroño, J., Sánchez-García, E., Sulprizio, M. P., Cao, K., Zhu, H., Liang, J., Li, X., Aben, I., and Jacob, D. J.: Satellite-Based Survey of Extreme Methane Emissions in the Permian Basin, Science Advances, 7, eabf4507, https://doi.org/10.1126/sciadv.abf4507, 2021.

Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C.: Satellite Observations of Atmospheric Methane and Their Value for Quantifying Methane Emissions, Atmospheric Chemistry and Physics, 16,
490 14 371–14 396, https://doi.org/10.5194/acp-16-14371-2016, 2016.

Korbeń, P., Jagoda, P., Maazallahi, H., Kammerer, J., Nęcki, J. M., Wietzel, J. B., Bartyzel, J., Radovici, A., Zavala-Araiza, D., and Röckmann, T.: Quantification of Methane Emission Rate from Oil and Gas Wells in Romania Using Ground-Based Measurement Techniques, Elementa: Science of the Anthropocene, 10, https://doi.org/10.1525/elementa.2022.00070, 2022.

Luetschwager, E., von Fischer, J. C., and Weller, Z. D.: Characterizing Detection Probabilities of Advanced Mobile Leak Surveys: Impli-
495 cations for Sampling Effort and Leak Size Estimation in Natural Gas Distribution Systems, Elementa: Science of the Anthropocene, 9, 00 143, https://doi.org/10.1525/elementa.2020.00143, 2021.

Maazallahi, H., Fernandez, J. M., Menoud, M., Zavala-Araiza, D., Weller, Z. D., Schwietzke, S., Von Fischer, J. C., Denier Van Der Gon, H., and Röckmann, T.: Methane Mapping, Emission Quantification, and Attribution in Two European Cities: Utrecht (NL) and Hamburg (DE), Atmospheric Chemistry and Physics, 20, 14 717–14 740, https://doi.org/10.5194/acp-20-14717-2020, 2020.

500 Maazallahi, H., Delre, A., Scheutz, C., Fredenslund, A. M., Schwietzke, S., Denier van der Gon, H., and Röckmann, T.: Intercomparison of Detection and Quantification Methods for Methane Emissions from the Natural Gas Distribution Network in Hamburg, Germany, Atmospheric Measurement Techniques, 16, 5051–5073, https://doi.org/10.5194/amt-16-5051-2023, 2023.

Ministerio de Minas y Energía: Resolution 40066/2022. Technical Requirements for the Detection and Repair of Leaks, the Utilisation, Flare and Venting of Natural Gas During Exploration and Production of Hydrocarbons' Activities., https://www.andi.com.co/Uploads/40066%
505 20de%202022.pdf, 2022.

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, Environmental Science & Technology, 50, 2099–2107, https://doi.org/10.1021/acs.est.5b05503, 2016.

Phillips, N. G., Ackley, R., Crosson, E. R., Down, A., Hutyra, L. R., Brondfield, M., Karr, J. D., Zhao, K., and Jackson, R. B.: Mapping Urban Pipeline Leaks: Methane Leaks across Boston, Environmental Pollution, 173, 1–4, https://doi.org/10.1016/j.envpol.2012.11.003, 2013.

Ražnjević, A.: High-Resolution Modelling of Plume Dispersion, Ph.D. thesis, Wageningen University, https://doi.org/10.18174/634327, 2023.

Ražnjević, A., van Heerwaarden, C., and Krol, M.: Evaluation of Two Common Source Estimation Measurement Strategies Using Large-Eddy Simulation of Plume Dispersion under Neutral Atmospheric Conditions, Atmospheric Measurement Techniques, 15, 3611–3628, https://doi.org/10.5194/amt-15-3611-2022, 2022.

Robertson, A. M., Edie, R., Snare, D., Soltis, J., Field, R. A., Burkhart, M. D., Bell, C. S., Zimmerle, D., and Murphy, S. M.: Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions, Environmental Science & Technology, 51, 8832–8840, https://doi.org/10.1021/acs.est.7b00571, 2017.

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., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth System Science Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

Shen, L., Zavala-Araiza, D., Gautam, R., Omara, M., Scarpelli, T., Sheng, J., Sulprizio, M. P., Zhuang, J., Zhang, Y., Qu, Z., Lu, X., Hamburg, S. P., and Jacob, D. J.: Unravelling a Large Methane Emission Discrepancy in Mexico Using Satellite Observations, Remote Sensing of Environment, 260, 112 461, https://doi.org/10.1016/j.rse.2021.112461, 2021.

Stavropoulou, F., Vinković, K., Kers, B., de Vries, M., van Heuven, S., Korbeń, P., Schmidt, M., Wietzel, J., Jagoda, P., Necki, J. M., Bartyzel, J., Maazallahi, H., Menoud, M., van der Veen, C., Walter, S., Tuzson, B., Ravelid, J., Morales, R. P., Emmenegger, L., Brunner, D., Steiner, M., Hensen, A., Velzeboer, I., van den Bulk, P., Denier van der Gon, H., Delre, A., Edjabou, M. E., Scheutz, C., Corbu, M., Iancu, S., Moaca, D., Scarlat, A., Tudor, A., Vizireanu, I., Calcan, A., Ardelean, M., Ghemulet, S., Pana, A., Constantinescu, A., Cusa, L., Nica, A., Baciu, C., Pop, C., Radovici, A., Mereuta, A., Stefanie, H., Hermans, B., Schwietzke, S., Zavala-Araiza, D., Chen, H., and Röckmann, T.: High Potential for CH_4 Emission Mitigation from Oil Infrastructure in One of EU's Major Production Regions, EGUsphere, pp. 1–20, https://doi.org/10.5194/egusphere-2023-247, 2023.

Szopa, S., Naik, V., Adhikary, B., Artaxo, P., Berntsen, T., Collins, W. D., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A., Klimont, Z., Liao, H., Unger, N., and Zanis, P.: 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, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., pp. 817–922, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.008, 2021.

Takriti, M., Wynn, P. M., Elias, D. M. O., Ward, S. E., Oakley, S., and McNamara, N. P.: Mobile Methane Measurements: Effects of Instrument Specifications on Data Interpretation, Reproducibility, and Isotopic Precision, Atmospheric Environment, 246, 118 067, https://doi.org/10.1016/j.atmosenv.2020.118067, 2021.

550 U.S. Environmental Protection Agency: Methane Emissions Reduction Program, https://www.epa.gov/inflation-reduction-act/methane-emissions-reduction-program, 2023.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald,
555 A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Vogel, F., Ars, S., Wunch, D., Lavoie, J., Gillespie, L., Maazallahi, H., Röckmann, T., Nęcki, J., Bartyzel, J., Jagoda, P., Lowry, D., France, J., Fernandez, J., Bakkaloglu, S., Fisher, R., Lanoiselle, M., Chen, H., Oudshoorn, M., Yver-Kwok, C., Defratyka, S., Morgui, J. A., Estruch, C., Curcoll, R., Grossi, C., Chen, J., Dietrich, F., Forstmaier, A., Denier van der Gon, H. A. C., Dellaert, S. N. C., Salo, J., Corbu, M., Iancu,
560 S. S., Tudor, A. S., Scarlat, A. I., and Calcan, A.: Ground-Based Mobile Measurements to Track Urban Methane Emissions from Natural Gas in 12 Cities across Eight Countries, Environmental Science & Technology, 58, 2271–2281, https://doi.org/10.1021/acs.est.3c03160, 2024.

von Fischer, J. C., Cooley, D., Chamberlain, S., Gaylord, A., Griebenow, C. J., Hamburg, S. P., Salo, J., Schumacher, R., Theobald, D., and Ham, J.: Rapid, Vehicle-Based Identification of Location and Magnitude of Urban Natural Gas Pipeline Leaks, Environmental Science &
565 Technology, 51, 4091–4099, https://doi.org/10.1021/acs.est.6b06095, 2017.

Weller, Z. D., Roscioli, J. R., Daube, W. C., Lamb, B. K., Ferrara, T. W., Brewer, P. E., and von Fischer, J. C.: Vehicle-Based Methane Surveys for Finding Natural Gas Leaks and Estimating Their Size: Validation and Uncertainty, Environmental Science & Technology, 52, 11 922–11 930, https://doi.org/10.1021/acs.est.8b03135, 2018.

Weller, Z. D., Yang, D. K., and von Fischer, J. C.: An Open Source Algorithm to Detect Natural Gas Leaks from Mobile Methane Survey
570 Data, PLOS ONE, 14, e0212 287, https://doi.org/10.1371/journal.pone.0212287, 2019.

Weller, Z. D., Hamburg, S. P., and von Fischer, J. C.: A National Estimate of Methane Leakage from Pipeline Mains in Natural Gas Local Distribution Systems, Environmental Science & Technology, 54, 8958–8967, https://doi.org/10.1021/acs.est.0c00437, 2020.

Wietzel, J. B. and Schmidt, M.: Methane Emission Mapping and Quantification in Two Medium-Sized Cities in Germany: Heidelberg and Schwetzingen, Atmospheric Environment: X, 20, 100 228, https://doi.org/10.1016/j.aeaoa.2023.100228, 2023.

575 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, Proceedings of the National Academy of Sciences, 112, 15 597–15 602, https://doi.org/10.1073/pnas.1522126112, 2015.

Zavala-Araiza, D., Omara, M., Gautam, R., Smith, M. L., Pandey, S., Aben, I., Almanza-Veloz, V., Conley, S., Houweling, S., Kort, E. A.,
580 Maasakkers, J. D., Molina, L. T., Pusuluri, A., Scarpelli, T., Schwietzke, S., Shen, L., Zavala, M., and Hamburg, S. P.: A Tale of Two Regions: Methane Emissions from Oil and Gas Production in Offshore/Onshore Mexico, Environmental Research Letters, 16, 024 019, https://doi.org/10.1088/1748-9326/abceeb, 2021.