Acting on the community comment from Dr. Luise Westphal and the note from the editor, we are providing more insights in the abstract as follows. We thank them for the comments which lead to improvement of the manuscript. The following lines are added to explain observed correlation between gas leak emission rate and subsurface accumulation in the limited and randomly selected population of gas leak locations in this study.

While the number of gas leak locations in this study is small, we observe a correlation between leak emission rate and subsurface accumulation. Wide accumulation places leaks into a safety category that requires immediate repair so that the suction method cannot be applied to these larger leaks in routine operation. This introduced a sampling bias for the suction method in this study towards the low-emission leaks, which do not require immediate repair action. Given that this study is based on random sampling, such a sampling bias may also exist for the suction method outside of this study. While an investigation of the causal relationship between safety category and leak size is beyond the scope of this study, on average higher emission rates were observed from all the three measurement-based quantification methods for leaks with higher safety priority compared to the leaks with lower safety concern.

We thank the referee for the comments which resulted in improving the manuscript. Here, we provide our answers to 'Anonymous Referee #3' for "Intercomparison of detection and quantification methods for methane emissions from the natural gas distribution network in Hamburg, Germany". Please find our answers in normal blue text and changes in the manuscript in ***bold italic blue text***.

Manuscript "Intercomparison of detection and quantification methods for methane emissions from the natural gas distribution network in Hamburg, Germany" presents results of methane emission in Hamburg quantified with three measurement methods, i.e., mobile, tracer release and suction. The measurement campaign was well-organized although there were some practical limitations such as short time period and intensive labor work. The most parts of the manuscript are also well-written despite the part of suction method is relatively weak due to lack of enough results, which I agree with another referee. However, since I am also from measurement community and fully understand the difficulties of organizing and performing such systematical measurement campaign, the results of this study are still valuable. Therefore, I recommend that this manuscript could be published after several minor revisions. The methods in the manuscript are well-established and described in detail, and the results of each method as well as case studies are presented systematically. Therefore, I mainly focus on the Discussion part. The following is my comments, I use the updated version of the manuscript:

----

1. line 762, I think it should be Table 1, not Table 2 since there is no Table 2 in the main text. Please also check throughout.

This has been corrected.

2. line 850-864. Interesting to see that distance of the transect to leak location can influence the quantification of emission rate. I have seen Eq 1 several times in different references, do they also report this problem? If yes, please add the comparison and address the importance to modify this widely-used equation.

The following statement has been added to the manuscript.

Although distance is a parameter in some quantification methods, e.g. gaussian plume dispersion, and not in others, e.g. in mass balance, to the best of our knowledge, this is the first study providing field evidence that distance is a factor that can affect emission quantification using the Weller et al. (2019) method.

3. line 925-961. Tracer method is the standard method to detect and quantify emissions of volatile organic compounds (VOCs) according to the EU standard EN 17628:2022, Fugitive and diffuse emissions of common concern to industry sectors - Standard method to determine diffuse emissions of volatile organic compounds into the atmosphere. In this standard, the typical expanded uncertainty in emission rate of tracer method is 20%-40%. So in your result, is the uncertainty also in this 20%-40% range? While the mobile method in the manuscript is not the standard method and have huge difference from the tracer method, do you think the tracer method is more accurate than mobile method to quantify the emission rate? Please state your evaluation of these two methods if applicable.

The following statement has been added to the manuscript.

Generally, the tracer method has higher precision than the mobile method, but it is more labor intensive. Although the mobile method has lower precision for emission quantification of individual gas leaks, this method can be implemented widely in a shorter time frame at a city scale. The mobile method is an empirical statistical quantification approach based on controlled release experiments, and a large sample size gives a better estimation in total emission (Weller et al., 2020). Acting on parameters in plume dispersion, such as distance and wind speed which are not included in the method, mobile method can overestimate and underestimate individual gas leaks but with a large number of gas leak quantifications these over- and underestimations may cancel each other out. If (i) particularly large concurrent subsurface CH_4 and C_2H_6 accumulations with multiple emission outlets are observed (this has priority - indication of a large leak) or (ii) a few of the leaks are significantly larger with small subsurface accumulation than the other leaks, an optimal approach may be to supplement the mobile method with use of a more precise measurement method such as the tracer method at those selected locations. The divergence to accurate city-wide quantification is dependent on urban planning, e.g. width of streets, location of gas pipelines (under streets or pavements etc.) and emission outlet location (s).

4. line 962. I suggest to change the heading, since this part actually does not address much implication on the emission inventories, which I agree with another referee, no data of emission inventories is presented and compared to the suction method results. Since the whole manuscript focuses on reporting measurement results, I think this part could only mention the existence of discrepancy and importance of comparing emission inventories and measurements

The title is changed as follows:

Possible sampling bias of suction method toward low gas leak emission locations

Reference:

Weller, Z. D., Yang, D. K., and von Fischer, J. C.: An open source algorithm to detect natural gas leaks from mobile methane survey data, edited by: Mauder, M., PLoS One, 14, e0212287, https://doi.org/10.1371/journal.pone.0212287, 2019.

We thank the referee for the comments which resulted in improving the manuscript. Here, we provide our answers to 'Anonymous Referee #4' for "Intercomparison of detection and quantification methods for methane emissions from the natural gas distribution network in Hamburg, Germany". Please find our answers in normal blue text and changes in the manuscript in ***bold italic blue text***.

Maazallahi et al. summarize the results of field experiments localizing and quantifying ~20 gas leaks in the city of Hamburg, Germany. A key finding is that common methods (the mobile, tracer, suction, and hole methods) for quantifying urban gas leaks from measurements and engineering models are highly uncertain, often differing by an order of magnitude or more. Another is that the suction method suffers from sampling bias that would have major implications for a future emission inventory that may rely on it. The authors describe in detail the different localization/quantification methods and various categories of observed leaks. I believe the paper is clearly written and a good fit for AMT and should be accepted for publication subject to the comments and questions below. A general (minor) criticism is that it feels unnecessarily long. I would recommend the authors trim and condense the text where possible. I also recommend more clearly articulating the weaknesses of the mobile and hole methods and emphasizing the need for improvement in single-leak quantification.

Comments

- L. 130: This is the first reference to the "high-flow sampler" method -- clarify?

Explanation of the method is now added as follows:

In the high-flow sampler method, air was drawn at a flow rate of about 0.2 m^3 min^{-1} through a flexible enclosure which covered a leak from a component completely. In this method, CH_4 mixing ratio was measured with catalytic oxidation and thermal conductivity hydrocarbon sensors and a thermal flow meter was used to determine gas flow.

- L. 139: It's not clear where the discrepancy is. Was the suction method applied to some of the same leaks as the mobile methods in previous studies and was a discrepancy found?

The operators of the suction method reported that they never measure emission rates as high as some of the ones reported from the mobile method. We carried out the campaign to address this discrepancy.

- L. 198: What were these restrictions?

Following explanations have been added:

…regarding safety, i.e. time allowed between detection and repair of leaks for different leak types, and method capacities, e.g. time required, labor intensity, logistics.

Leaks safety types are explained prior to these lines in the introduction, L91 to L97.

- L. 287: What does similar mean here?

'Similar in mole fractions' is now changed to 'same order of magnitude in mole fractions'. Here it means that if CH_4 mixing ratio was from all outlets at a location were in 10s of ppm or 100s of ppm, those reading were considered "similar" and the spatial average of the outlet were considered as the main emission point.

- L. 311-313: How much is mostly? What fraction of the leak locations matched?
- Eq. 1: Can you better motivate this equation? How can it give a meaningful estimate of emission rate from concentration without information on wind speed or distance from the source?

The first version of the equation was developed in von Fischer et al. (2017) which was based on set of release experiment ranging from 0.5 to 50 L min^{-1} at distance of 5, 10, 20 and 40 m. Wind speed was logged using instruments mounted on cars and at a stationary location. In total 276 passes were performed downwind the release and quantification method was developed using maximum CH_4 measurement of each plume, plume area (CH_4 enhancement along driving track) and ratio of the first two.

In the second version of the quantification method, Weller et al. (2019) used the same dataset as above to back calculate CH_4 emission rates using mobile measurements. It was concluded that the best predictor is CH_4 enhancement above background level and inclusion of other parameters didn't meaningfully improve the method.

We have added following lines above Fig. 6 to provide explanations:

Except for three leaks (HH003, HH009 and HH011), leaks were located by the repair team where the carpet method reported. These three leaks were finally found at some distance from the location initially indicated by the carpet method. These three leaks are medium or high emitters and belong to type A1. Although the number of locations in this study is very small, this supports the common sense assumption that bigger leaks can spread out more widely in soil and contaminate larger undersurface area. Therefore, the bigger leaks may be mislocated by the carpet method, and they are also more likely to fill cavities, placing them in a higher safety category.

- L. 358-360: In windy conditions the plume will be long and narrow, in calm conditions short and wide. Why should the plume area be the same in both cases? Is this the area under the curve of the methane plume transect?

The area refers to the CH_4 enhancement along mobile measurement track. Real plumes show many different physical representations and the integrated enhancement along the measurements track (referred to as peak area) is more likely to represent the total number of molecules present at a certain distance downwind from a source compared to the peak maximum, because the plume can under similar conditions be narrower and higher or wider and lower.

- L. 847-849: If this is the area under the curve, then yes, that does seem a better metric. But wind speed and distance to the source would still need to be accounted for.

In the routine application of the mobile method, distance is not known, and therefore also not included in the equation. It can only be added once the main outlet has been identified. In the earlier development of the mobile gas leak quantification method, it was stated that variation in wind speed has little or no impact on the leak quantification (Von Fischer et al. (2017).

- L. 855-857: This is a basic property of atmospheric plume dispersion and it's surprising to me that it isn't accounted for (along with wind speed) in your and e.g., Weller et al.'s emission quantifications.

We agree that this is surprising. As mentioned above, distance is usually not known unless you actually get out and identify outlet (s). In some occasions there are several outlets spread spatially, thus it may not be clear from which outlet an emission is observed during a mobile measurement pass. Von Fischer et al. (2017) binned emission rates into small (<6 L min^{-1}), medium (6-40 L min^{-1}) and high (>40 L min^{-1}) following communication with gas distribution operators to plan repair actions. Von Fischer et al. (2017) explicitly state that including wind speed as parameter did not improve the quantification. Weller et al. (2019) also stated that inclusion of other parameters other than CH_4 enhancement did not improve the quantification method. The reason is possibly the turbulent nature of plume dispersion on streets, where traffic additionally affects the air flow considerably. Therefore, the Weller et al. (2019) quantification method evaluates gas leak locations statistically. This method has lower precision for individual gas leak quantifications which improves with performing several transects at a leak location (Luetschwager et al., 2019).

- L. 862-864: Yes, and wind speed. It is not surprising that methods ignoring these basic parameters of the problem produce highly imprecise emission estimates.

See explanation to the previous comment.

- L. 869-871: Should we be able to see this in Table 1? The mobile and tracer estimates in that table don't seem to match very well.

This comparison between mobile quantification and the known release rates for C_2H_2 is provided in Table S8 (in Sect. S.10 in SI) and it is not provided in Table 1.

Corrections

- L. 18: "emission" → "emissions"

Corrected.

- L. 68: "needs" → "need"

Corrected.

- L. 84: "sewage system" → "sewage systems" or "the sewage system"

Corrected.

- L. 102: "on annual basis" → "on an annual basis"

Corrected.

- L. 103: "platform" → "platforms"

Corrected.

- L. 107-109: references appear to be in random order

The order is now adjusted chronologically.

- L. 253: "which take about couple of minutes per outlet location" – grammar and please be more specific.

The sentence is modified as follows:

The instrument can detect C_2H_6 by gas chromatography in batch mode, which means that after taking air samples from a suspected outlet, the instrument operator needs to wait for couple of minutes to test possible detection of C_2H_6. This is substantially slower than instrument with 1 Hz frequency used in the mobile method.

- L. 519: HH006 is specified in the next sentence.

The sentence is modified as follows:

The suction method was applied at 8 gas leak locations (see Table 1). At only one location the quantification could be completed according to protocol where an equilibrium mixing ratio has to be reached.

- L. 528: “supposedly” → “likely”?
This is changed.

- L. 557: “far apart” → “apart”
This is corrected.

- L. 837: dangling “then”?
This is removed now.

Reference:

Luetschwager, E., von Fischer, J. C., Weller, Z. D., Characterizing detection probabilities of advanced mobile leak surveys: Implications for sampling effort and leak size estimation in natural gas distribution systems. Elementa: Science of the Anthropocene; 9 (1): 00143. https://doi.org/10.1525/elementa.2020.00143, 2021.

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, Environ. Sci. Technol., 51, 4091–4099, https://doi.org/10.1021/acs.est.6b06095, 2017.

Weller, Z. D., Yang, D. K., and von Fischer, J. C.: An open source algorithm to detect natural gas leaks from mobile methane survey data, edited by: Mauder, M., PLoS One, 14, e0212287, https://doi.org/10.1371/journal.pone.0212287, 2019.