This document provides our answers to the anonymous referee #1, referee #2, short comment from Luise Westphal, and marked-up version of manuscript of "**Methane mapping, emission quantification, and attribution in two European cities; Utrecht, NL and Hamburg, DE**".

In the following sections, comments from the referees and Luise Westphal are provided in normal black text, our replies are in normal blue, and changes is the manuscript are in *blue italic* format. In the end, we provide the marked-up manuscript with relevant changes in *blue italic*.

List of answers:

Our answers to referee #1:

1 Overview

This study presents an extensive set of vehicle-based measurements of methane and other compounds to investigate methane emissions in Utrecht, NL and Hamburg, DE. The authors used empirical equation developed von Fischer et al. (2017) and updated by Weller et al. (2019) to estimate the emissions from the natural gas distribution network using the methane enhancements they observed during their surveys. They also tested several approaches to determine the origin of these enhancements (biogenic, thermogenic, pyrogenic) such as the isotopic signature of methane, C_2/C_1 ratio or CH_4/CO_2 ratio. Finally, they used a Gaussian dispersion model to estimate methane emissions from larger sources. Overall, despite the large ranges and uncertainties presented here, the study is a valuable contribution to the literature as the state of knowledge of urban methane is not very advanced compared to other pollutants. The measurement campaigns seems carefully done and well designed, I appreciated the authors described their interactions with the local distribution companies and showed how their work helped reducing the emissions. It was also nice to find a list of all the acronyms in the supporting information as there are so many of them in the text and I was a bit lost at first. I recommend publishing it after addressing few minor points (which are also mentioned in the detailed comments):

1 General comments

1) Structure: I would reorganize section 2 a bit and group all the source attribution approaches together in one subsection instead of having two subsections about isotopic analysis and information about the ratios scattered throughout the rest of the subsections. It would also make it easier for the reader if there was a table in this section summarizing these approaches and the limits used to attribute the emissions.

- We have followed this suggestion and section, Section 2 is now rearranged to the following sub-sections: 2.1. Data collection and instrumentation, 2.2. Emission quantification, and 2.3. Emission attribution.
- In the "data collection" section we keep sub-sections of mobile measurement of C_2H_6 and CO_2 separated from sampling for isotopic analysis as the analytical techniques for these two attribution methods are very different. Nevertheless, we now combine data evaluations of C_2H_6, CO_2, and isotopic analysis in sub-section 2.3.

2) I would have liked more discussions about the uncertainty on the emissions estimated with the approach developed by Weller et al. (2019). I understand that this method is the reference for mobile surveys at the moment but the fit of the calibration curve presented in figure 4 of Weller et al. (2019) makes me wonder about the uncertainties associated with these estimates. Also, the authors used this empirical equation to estimate emissions from microbial and combustion sources whereas it was originally designed to estimate emissions from NGDN. While the classification into small, medium and large LIs depending on the maximum amplitude of the enhancement remains correct, I am not very comfortable using the empirical equation to estimate the emissions of these other sources. Biogenic emissions are very different from NGDN emissions, they are way more sensitive to atmospheric conditions (especially temperature) and are likely to vary in time unlike NGDN emissions which should be more constant. They could also potentially be located further away than the usual road side emissions. Figure S16 examples illustrate this: microbial emissions from the water body are likely localized further away and the sewage system seem to emit at a higher level than the road level.

- Yes, the uncertainties of the quantification algorithm introduced by von Fischer et al. (2017) and improved by Weller et al. (2019) for individual LIs are large. It is indeed evident from figure 4 of Weller et al. (2019) that individual LIs can be strongly under- or overestimated. The rationale is that when a complete city is surveyed, the contribution from the different LI categories and the

total emission rate can be estimated more precisely. Following the request from the referee, we have explicitly clarified this better in the revised version (see Sect. 2.2.2, L293-302).

- We acknowledge that the algorithm in the original papers was designed to quantify pipeline leaks. Attempts to exclude emissions from other sources were to restrict the spatial extent of a CH_4 plume to <160 m, to require a minimum enhancement of 10% above background, and to require multiple detections. In our study, we add explicit attribution for many LIs by evaluating the co-emitted (or not) tracers C_2H_6 and CO_2. Rather than simply flagging and neglecting these LIs for the quantification of potential pipeline leaks, we use the same algorithm for quantifying emissions from other categories, namely microbial processes and combustion. We thank the referee for pointing out that this needs to be spelled out specifically in the paper. Whereas biogenic emissions, e.g. from the sewage or wastewater systems, will be released through manholes, waste water drains and other cavities that are also important for the release of CH_4 to the atmosphere, combustion related emissions may come from vehicles or houses and thus have different release pathways. The quantification of emissions from these source categories is thus based on the assumption that the same conversion equation can be used. This is especially the case for microbial emissions from manholes, where the enhancements and distance of the emission release point is very similar to NGDN leaks, hence the emission quantification approach applies for both source categories. In the revised version, we have stated this explicitly. We now suggest that the number of detected enhancements that have been attributed to microbial and combustion sources is more reliable than the emission rates.
- We have considered at an early stage of our research to use the information of co-emitted species (especially C_2H_6) to focus the paper only on NGDN leaks but decided against this. We feel the biogenic emissions are part of the overall anthropogenic urban CH_4 emission and, while being more uncertain, it is relevant to have an approximation of the importance in relation to NGDN. We thank the reviewer for reminding us that we need to be extra careful with this category of emissions (and we agree) but still think it is valuable information.

3) The presentation of the GPDM approach used to quantify the emissions from larger facilities should be reworked and expanded a bit. For example, the authors should talk about the wind data they are using as it is a critical parameter in this approach. Did they adjust the wind direction so that the maxima of the observed plume is aligned with the maxima of the modeled plume, etc. The part about the selection of the sigma y and sigma z is also not very clear. The author should also specify here which observation they are fitting with the model (the measured concentrations? One plume at a time or all the plumes measured during all the surveys?)

- The relevant section has been expanded to provide more information. The reason for using data from the two mentioned towers in Utrecht and Hamburg is that the online data logging setup failed to continuously record all the local wind measurements during the surveys. The distance of the towers to the facilities ranges from 8 to 20 km, and indeed these distances introduce extra uncertainties in emission quantification, mainly related to wind speed (see Sect. 2.1.5). When we compare the data that were recorded on the vehicle with the tower, we derive a difference in wind speed of ± 10 %. After considering the remarks of the reviewer, we increase this to a more conservative error estimate of ± 30 % (see Sect. 2.2.3, L352-359).
- Regarding the wind direction (e.g. the oil wells), for several sources the emission point is relatively certain and confirmed by analysis of Google Earth images. In addition, we passed several sources during different wind conditions and did a "triangulation" based on the observed plumes and wind data (see Sect. 2.2.3, L324-333).
- For explanation on how sigma_y and sigma_z were derived, see the detailed comment below.
- Regarding the last question, we fit each plume individually (since they were often observed during different days at different locations) and average the individual results. This has now been specified (see Sect. 2.2.3, L343-344).

3 Detailed comments

L18-19 and L30: Should be consistent with number notation (whether letters or numbers).

- Done

L22: This should be phrased differently, the largest emission rate in Utrecht is actually coming from the wastewater treatment plant.

- This sentence focuses on the emissions from LIs found in the normal mobile surveys, and doesn't include emissions from larger facilities.

L64: Typo? "high precision" is written twice in a row.

- Done

L76: Typo? A comma is missing after "(Giolo et al., 2012; Helfter et al., 2016)".

- Done

L78-80: This sentence is too vague, most methods quantify emissions methane enhancements! The authors should specify which approach they used and which type of sources they used it for.

- The respective sentence was changed as follows (see Sect. 1, L88-90): *In this study, we quantified LIs emissions using an empirical equation from Weller et al. (2019), which was designed based on controlled release experiments from von Fischer et al. (2017), to quantify ground-level emissions locations in urban area such as leaks from NGDN.*

L83: Typo? Should it be "across the urban areas in these two cities" (rather than "across the urban areas is these two cities")?

- Done

L93: Why specify the time needed to flush the cell for the G2301 but not for the G4302? Could you add a sentence about how the methane enhancements measured by the two instruments compare? This discussion is actually in the SI, the authors could add a sentence to refer it.

- This sentence describes the smoothing effect in the cell of G2301. Information on G4302 is added to the revised manuscript as follows (see Sect. 2.1.1, L110-111): *The flow rate is 2.2 L min^{-1} and the volume of the cell is 35 ml (operated at 600 mb, thus 21 ml STP) so the cell is actually flushed in 0.01 s, which means that mixing is insignificant given the 1 s measurement frequency of the G4302.*

L122: What do the level 2 and 3 roads correspond to?

- Information on the level 2 and 3 are now added to the paragraph and the respective sentence was changed as follows (see Sect. 2.1.2, L151-154): *Level 1 roads are primarily larger roads connecting cities, level 2 roads are the second most important roads and part of a greater network to connect smaller towns, level 3 roads have tertiary importance level and connect smaller settlements and districts.*

L137: The authors should remove "at the following links: Utrecht and Hamburg", the citation "(Maazahalli et al., 2020b)" is enough.

- Done

Section 2.4: I would merge sections 2.4 and 2.6.1 into a source attribution section that details the multiple approached used in this study. I would incorporate in this section a table summarizing the different ratios/isotopic measurements and the ranges used to distinguish between fossil, combustion, microbial and unclassified sources.

- Section two is now rearranged to three main sections: '2.1. Data collection and instrumentation', '2.2 Emission quantification', and '2.3. Emission attribution'. In Sect. 2.1 we present information on the mobile measurement of CO_2 and C_2H_6 and sample collection. In Sect. 2.2, we provide information on emission quantification of LI and larger facilities separately. In Sect. 2.3, we combine the attribution approaches as suggested by the referee.

L168: Did the authors only took samples for isotopic analysis in Hamburg? Why not in Utrecht?

- Due to time and budget limitations, we were only able to take a sufficient number of samples for attribution in Hamburg.

L179: How far are the measurements tower from the studied sites? Wind parameterization are large sources of uncertainty in Gaussian plume dispersion model, especially since wind close to the surface can be very different from the wind measured at 10 meters at these towers.

- The distance of the towers to the facilities ranges from 8 to 20 km. The reason for using the tower data was that wind data were not logged by the 2-D anemometer continuously (instrument failure) and the limited data collected were not sufficient to analyze emissions from these facilities.

L184: "It has been demonstrated that the algorithm adequately estimates the majority of emissions from a city (Weller et al., 2018)." The authors should specify that this method was specially developed to quantify methane emissions from the natural gas distribution network. In this sentence, the authors seem to imply that they could estimate the emissions from any type of sources from a city.

- The respective sentence was changed as follows (see Sect. 2.2.2, L246-L249): *This algorithm was designed to quantify CH_4 emissions from ground-level emission release locations within 5-40 m from the measurement (von Fischer et al., 2017), such as pipeline leaks and has been demonstrated that the algorithm adequately estimates the majority of those emissions from a city (Weller et al., 2018).* As mentioned in the reply to the general comment 2, in this study we use the same algorithm to provide indicative estimates of emission rates of microbial and combustion emissions as well. We note that emission pathways to the atmosphere are partially different for such emissions. Therefore, the emission rates should be seen as indicative, whereas the LI numbers from the different categories are more reliable.

L192-194: How did the authors know about the mole percent of CH_4 and C_2H_6 in the NGDN in Hamburg and Utrecht? Is it based on measurements or did the NG suppliers give them this information?

- This data was indeed provided by the network operators. This is now indicated as "personal communication" in the manuscript (see Sec. 2.3.1, L362-364) as follows: *During the Utrecht campaign, the overall mole fraction of CH_4 and C_2H_6 in the NGDN was ≈ 80 % and ≈ 3.9 % (STEDIN, personal communication) and in Hamburg the mole fraction of CH_4 and C_2H_6 in the NGDN was about ≈ 95 % and ≈ 3.4 % (GasNetz Hamburg, personal communication) respectively.*

L196: If I understand correctly, this whole part is used to explain how you differentiate car exhaust signals from NG signals. This is not really clear, the authors should introduced it up front to help the reader follow the organization. This could probably also be moved to the source attribution section.

- Done

L204-207: I don't understand why do the authors use different approaches to estimate the CH_4 and CO_2 backgrounds? This should be explained.

- The background determination method for CH_4 from Weller et al. (2019) was used to stay compatible with the quantification algorithm for the urban studies. But this algorithm doesn't include background extraction for CO_2 and here we chose background detection methods commonly used in the literature (see Sect. 2.2.1 and Figure S7 in SI, Sect. S.2.1).

L229: Did the authors really need to convert decimal degrees to Cartesian coordinates in order to cluster enhancements? Doesn't it introduce additional uncertainties than directly estimating the distance between enhancements using decimal degrees?

- The constraint for clustering based on the von Fischer et al. (2017) algorithm is 30 m, thus we need to have the data in metric system. There are many ways to convert the decimal degrees to metric system, we used this way as it gave a very good one-to-one correlation with R^2=1.00 when we compared output of the equation we used for converting to cartesian system to e.g. EPSG:32632 projection. Easy implementation of this equation in the code we wrote to evaluate the data is another advantage of using this method to convert decimal degrees to metric system.

L233: Why did the authors assigned the maximum observed enhancement to the cluster rather than a weighted average just like for the location? Wouldn't that artificially increase the emissions?

- This follows the algorithm from Weller et al. (2019).

L240: The "visited at least twice" criterion in von Fischer et al. (2017) and Weller et al. (2019) was implemented to identify enhancements from the natural gas distribution network that are considered to emit continuously. I would mention that you are using another source attribution method instead.

- The following sentence was added to the paper (see Sect. 2.2.2, L287-288): *Instead, we used explicit source attribution by co-emitted tracers.* This topic was also discussed earlier in (see Sect. 2.3, L189-193): *Due to time and budget restrictions, it was not possible to cover each street at least twice, as done for the US cities. After evaluation of the untargeted first surveys that covered each street at least once, targeted surveys were carried out for verification of observed LIs and for collection of air samples at locations with high CH_4 enhancements. The rationale behind this measurement strategy is that if an enhancement was not recorded during the first survey, it obviously cannot be verified in the second survey.*

Section 2.6.3: I was surprised that the authors did not talk about wind measurements in this section given that this is one of the biggest source of uncertainty of this technique. Maybe they should move part of section 2.5 here.

- We acknowledge that the wind speed is a large source of uncertainty in the GPDM. The section is now revised and more information has been added in Sect. 2.2.3. See our answer to general comment 3.

L252: What do the authors mean by "These data were evaluated using a simple point source GPDM"? What are the authors evaluating?

- We meant that the data collected downwind the larger facilities were analyzed using GPDM. The respective sentence was changed as follows (see Sect. 2.2.3, L306): *We applied a standard point source GPDM (Turner, 1969) to quantify methane emissions from these larger facilities*.

L252: Typo? "()" should be removed.

- Done.

L265-266: The authors should be consistent with the notation: zsource (which is equal to 0 in the text) and h are to the same thing.

- Corrected.

L276-279: This part is not very clear. Do you select sigma y and sigma z separately? Could you end up with a sigma y of a given Pasquill-Gifford stability class and combine it with a sigma z from another stability class?

- We first determined sigma y based on the width of the plume observed during the measurement and the source location. From the distance between the source location and the maximum of the plume location and sigma_y we chose the most suitable Pasquill-Gifford stability class and then we chose the corresponding sigma_z value from the respective Pasquill-Gifford stability tables.

L288: It would be appropriate to at least in a sentence or two explain the isotopic analysis so the reader doesn't need to go back and read these papers (which analyzer, how long were the samples measured...).

Following this comment, we have now added and combined all information related to the isotopic analysis in Sect. 2.3.2 including analyzers, measurement time scales, calibrations, etc.

L314: Typo? "Utrecht and Hamburg correspond to" rather than "Utrecht and Hamburg were correspond to"

- Done

L321: Typo? "Figure 2" looks weird.

- Done

L332: You showed previously that different types of road had very different LI rates per km depending on cities, why didn't the authors use these road-specific emission factors to upscale their emissions?

- The evaluation showed that different types of road have different LI rates per km in these two cities, which means that the smaller or bigger LIs can happen on different road types. In this study we aim to compare cities based on total emissions derived from LIs, so for the upscaling we used total length of road no matter what road types those are.

Figure 5: Typo? "of collected air samples" instead of "of air samples collected". The authors should also show the microbial and pyrogenic clusters on these figures (L342).

- Done
- We tried to add additional boxes to the figure, but this makes the figure quite busy and therefore we prefer to highlight the "gas leaks" category only.

L352: Typo? "combustion-related" instead of "combustion, related".

- Done

L360: Not clear which criteria for CH_4/CO_2 ratios the authors used to classify LIs as combustion-related in the end. $CH_4/CO_2 > 0.2$ ppb/ppm?

- The criteria we used to identify combustion-related signals are based on CH_4 enhancement to CO_2 enhancement (CH_4:CO_2 ratio (ppb:ppm)). If the ratio is between 0.02 and 20 ppb:ppm and linear regression enhancements of these two species has R^2 greater than 0.8, we attributed those LIs as combustion-related sources. This has been specified in the revised version as follows (see Sect. 3.2, L469): *Based on the CH_4:CO_2 ratio (ppb:ppm) criterion defined above (see Sect. 2.3.1),…*

Figure 6: This figure is relatively difficult to interpret, it is difficult to visualize the shape of the observed plumes when they superimposed like this. It would have been interesting to see how and where you triangulated the location of the source for this site. How many sources did you find for this site? In wastewater treatment plant, the main methane source usually correspond to the sludge treatment areas that can be spotted with Google Earth.

- Figure 6 gives the overview of measurements around the WWTP. An example of the shape of a plume is given in Figure 7.

L375-377: The definition of the error estimate is very confusing, what are the 5 sets of measurements if there were only 3 days of measurements at the wastewater treatment plant?

- On some days there were two sets of measurements per day; e.g. one in the morning and one set in the afternoon. We have now defined the definition of measurement set in the paper which described back to back measurement downwind each facility as follows (see Sect. 2.2.3, L344-

345): *A set of plumes is defined as a back to back transects during a period of time downwind each facility on different days.*

L395: Typo? Extra space before "74%".

- Done

L426: Typo? One of the "%" should be removed.

- Done

L413-432: The author should expand the discussion about the different source attribution approaches, is it necessary to use all of them? Which approach would the authors recommend to use in the future?

- The following sentence was added to the manuscript (see Sect. 4.2, L547-549): *Overall, C_2H_6 and CO_2 signals are very useful in eliminating non-fossil LIs in mobile urban measurements and with improvements in instrumentations, analyzing signals of these two species along with evaluation of CH_4 signals can make process of detecting pipeline leaks from NGDN more efficient.*

L479: Shouldn't it be the "annual natural gas leakage rate per capita" rather than the "annual natural gas consumptions per capita"?

- No, this sentence refers to the annual gas consumption provided in the previous mentioned sentence and intends to give a comparison between consumption per capita in Utrecht, Hamburg, and US.

L480: Typo? "per km of pipeline" rather than "per km pipeline"?

- Done

L491: The authors already explained several times that natural gas emissions depends on the age of the pipelines and the type of material used for these pipelines. I am not sure it is useful to repeat it here, especially since it will be discussed again later (L514).

- Here (Sect. 4.3), we mention the pipeline material and age, as these have important influences on the emissions from NGDNs in different cities, and later we give more information on different types and age of pipeline (see Sect. 4.4).

L545-549: The authors should choose one unit for the emissions and use it for all the sources, it would make easier for the reader to compare these emissions (wastewater treatment plant in t/yr, wells in kg/h...).

- Done

L557: Typo? "For emissions from the NGDN, the urban..." rather than "For emissions from the NGDN the urban...".

- Done

L545-557: Did the authors also looked at the ratios of these larger facilities? It could be also be an interesting information.

- Correlations between CH_4:CO_2 for the facilities were not very good, which may be due to the relatively small enhancements of CH_4 downwind the facilities and the expected ratio of CH_4:CO_2.
- For the Utrecht WWTP, a ratio of 0.4 ppb:ppm of CH_4:CO_2 with R^2 of about 0.52 was observed. The sludge treatment part of the WWTP emits both CH_4 and CO_2 while CO_2 is also emitted from other parts of the WWTP, e.g. power generation, anoxic/anaerobic treatment part, which explains why the correlation is not very high.
- Downwind the Compost and Soil Company in Hamburg the CH_4 enhancement was low and no clear correlation between CH_4 and CO_2 was observed.

Supplementary information: Section 1: “Figure S2a and Figure S2b show total length...” rather than “In Figure S2a and Figure S2b total length...are shown”. Same for “In Table S1 and Table S2”.

- Corrected

Section 2.1: Typo? Should it be “CH_4-only mode, which show” (rather than “CH_4-only mode. which show”). It is indeed very strange that the higher inlet measures higher methane enhancements than the bumper inlet. Would it possible that this source was located above the ground (“chimney” emissions or like the sewer pictures showed below)?

- The typo has been corrected, thanks for spotting this.
- Based on the CO_2 and C:C1 analysis this LI can only be attributed to a source of natural gas emission, likely from a pipeline leak in the ground.
- We are presently investigating the influence of intake height and instrument response in more detail for an upcoming publication, where measurements in several cities will be compared. Qualitatively, the relatively slow flush time of the cavity and lower measurement rate in the G2301 relative to the G4302 instrument (see comment above) lead to generally higher maximum enhancements in the G4302 instrument compared to the G2301, which for our measurements in Hamburg and Utrecht counteracts the fact that the inlet of the G2301 is closer to the ground and thus closer to most emission points. For individual plumes, turbulence in the street from driving cars can occasionally lead to higher mole fractions at the top inlet.

Section 2.2: What does “the ratio of the sum of CH_4 enhancements (in ppb) to the sum of CO_2 enhancements (in ppm)” mean? Does it correspond to the area under the plume? There is no mention of Figure S7 in the text.

- The respective sentence was changed as follow (see Sect. S.2.6): *In Figure 12, the ratio of the area under the CH_4 enhancements along the driving track (in ppb*m) to the area of CO_2 enhancements along the driving track (in ppm*m) is 5.5 ppb:ppm which is much higher than reported in previous studies, possibly indicating incomplete combustion.*

Section 2.4: “Errors in wind speed are estimated to be ±10% and for wind direction ± 5°” this seems low to me considering that the wind was not measured on site but at a tower located away from the site. Table 5 caption should be better isolated from Table 4, this is a bit confusing at the moment.

- By comparing some of the recorded measurement from the 2-D anemometer next to the facilities with the data from the towers we noticed that the local wind speed data were within ± 10% of the data from the towers; as described above, following the comment of the referee we now use a more conservative estimate of ± 30%.
- Given that for many sources the emission point is known, either from satellite imagery or from triangulation, wind direction (between emission point and maximum of observed plume is quite well known and here we think that the error estimate of ± 5° is adequate.
- The caption for Table 5 has also been corrected.

Section 2.6: In Figure S10a, shouldn’t the authors constraint delta^{13}C, deltaD, C_2H_6 and CO_2 before clustering? It would avoid clustering enhancements from different types of sources. Figure S11: caption not very precise.

- Based on the Weller et al. (2019) algorithm, it is assumed that LIs which are clustered together should be from the same source. Thus, based on the algorithm if one of the LIs within a cluster belongs to a specific emission class (e.g. microbial or combustion, etc.) then all the others should have fall into that source class. Based on the multi-tracer and isotope data, we have no evidence that in our dataset this is not the case. Therefore, we kept the analysis this way to keep consistency with other studies where no attribution techniques were used to attribute the LIs.

References

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(7), 4091–4099, doi:10.1021/acs.est.6b06095, 2017.

Turner, D. B.: Workbook of Atmospheric Dispersion Estimates, U.S. Environmental Protection Agency. [online] Available from: https://nepis.epa.gov/Exe/ZyPDF.cgi/9101GKEZ.PDF?Dockey=9101GKEZ.PDF, 1969.

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 M. Mauder, PLoS One, 14(2), e0212287, doi:10.1371/journal.pone.0212287, 2019.

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, Environ. Sci. Technol., acs.est.8b03135, doi:10.1021/acs.est.8b03135, 2018.

Our answers to referee #2:

1 Overview

The authors present an extensive study of ground-based mobile measurements of methane and several related tracers (C_2H_6, CH_4/CO_2 ratio, δ^2H-CH_4 and $\delta^{13}C$-CH_4 focused on quantifying and attributing methane emissions in two European cities, namely Utrecht and Hamburg, which both rely on subterranean pipelines as the delivery system for natural gas used in the households and otherwise. Such delivery systems are known to cause leaks that contribute to the anthropogenic global warming, and it has been demonstrated previously that fixing of these gas leaks can be a very cost-effective mean of climate change mitigation. Using a combination of in situ observations (with CRDS), discrete samples collected for identified leak sources as well as Gaussian plume modelling, the authors are able to identify approximately 100 leak sources in both cities over their study period, and thanks to the robust analysis of collected data, are able to differentiate them according to emission source (natural gas distribution system / microbial sources) and the respective source strength (with just several sources responsible for large parts of total emissions). A comparison of the results against previous studies conducted in US point to potential lower specific emissions for the studied cities. The authors also attempt to upscale the measurements performed over these limited campaigns in order to compare them to the publicly available aggregated data, albeit these results should be treated with care as the dataset is limited and much more robust studies are needed to achieve this goal (which the authors accurately point out).

The authors should be commended for the impressive amount of high-quality work that was put into design, execution and data analyses during those campaigns. This is no easy task, as the study encompassed simultaneously using many state-of-the-art techniques from very different scientific fields together with a very large amount of data (both measurement and supplementary) in order to achieved the stated goals. Simultaneous analysis of several tracers and isotopic composition is of particular interest, as it shows great promise in development the methods of precise small-scale emission estimations. I find this study to be a strong contribution to the discussion in the city-scale methane emissions, and the strategy developed here seems to be promising in developing both research-targeted and operational methods for leak detection and its strength estimation.

The article does suffer however, from this wealth of data and methods, and requires multiple improvements before final publication. For example, some sections of the text require more detailed information in order for the study to be considered reproducible. Also, the treatment of uncertainty in the source estimation and Gaussian plume modelling sections should be deepened. In the second case, the sensitivity of the method to the chosen meteorological parameters should be established. In methods section some restructuring is recommended, and the Discussion could benefit from introducing a clearly defined structure in order to appropriately focus attention.
I recommend publishing the article after addressing items listed below.

2 Major comments:

1. The Method description could do with an overhaul. In some places, more details need to be provided (see ‘detailed comments’ section below) for the experiments to be considered reproducible. In others, some information should be combined (2.6.4. and 2.4.).

- We have revised the method section accordingly, see our answers to the detailed suggestions below.

2. Discussion of uncertainties in the urban emissions is very limited, with authors stating that ’We used a Bootstrap method (Nelson, 2008) to estimate 247 emission uncertainties similar to Weller et al. (2018) for the US city studies by resampling from all recorded LIs randomly 30,000 times.’ No further comment is given, and in the discussion section the authors quickly skim over this and analyze the statistics of LI, without providing information on how precise those classifications might be. The method

described by Weller et al. (and earlier by von Fisher et al.) relies heavily on assumptions regarding the distance to the source and (calibration using control releases) it stands to reason that this simplistic approach must produce very large uncertainties if not supported by multiple measurement repetition. This is critically important here as the data from limited detections is interpreted and up-scaled. As it is now, it is not possible to get a realistic impression about the numbers given, and puts the resulting data analysis in question. The discussion of uncertainty and potential biases should therefore be expanded.

- As written in the manuscript, we follow the algorithm that has been described in detail in von Fischer et al. (2017) and Weller et al (2019). These studies include information on the precision, and we do not think that it is necessary to repeat all of the work that was performed there in our manuscript. We follow their approach so closely because we feel the inter-comparison with the US studies is important for a better overall understanding. The algorithm was applied to LIs that were surveyed at least twice, which we did not have as a strict criterion in our study. This difference is discussed in detail in our manuscript. Revisiting is not expected to have a large influence on the classification, since the Weller et al. (2019) algorithm only considers the maximum reading of the (at least two) identifications, so a revisit can only result in a higher emission rate estimate.
- Like other studies, we do not have information on the distance of the leaks from the point of observation, since we do not have access to the full grid of pipelines in both cities. We do mention, that individual leak rate estimates can have large errors, and have included this statement explicitly regarding the highest derived leak rates in the revised manuscript. Since we covered a very large fraction of the city, the error of the total emissions is relatively small, i.e., about 30 % and 15 % in Utrecht and Hamburg, respectively. The bootstrapping is a standard technique and it has been demonstrated that it performs well in emission quantification and the determination of emission factors. The bootstrapping is a standard technique and it has been proven that it performs well in emission quantification and the determination of emission factors.
- The following paragraph is now expanded in Sect. 2.2.2, L292-301:

To account for the emission uncertainty, similar to Weller et al. (2018) for the US city studies, we used a bootstrap technique which was initially introduced in Efron (1979, 1982), as this technique is adequate in resampling of both parametric and non-parametric problems with even non-normal distribution of observed data. Tong et al. (2012) indicated that bootstrap resampling technique is sufficiently capable in estimating uncertainty of emissions with sample size of equal or larger than 9. Efron and Tibshirani (1993) suggested that minimum of 1,000 iterations are adequate in bootstrap technique. In this study, we used non-parametric bootstrap technique to account for the uncertainty of total CH_4 emissions from all LIs in each city with 30,000 replications. As mentioned above the algorithm is based on CH_4 enhancements of measurement with 5-40 m distance from controlled release location, and can produce large uncertainty for emission quantification of individual LI (Figure 4 in Weller et al. (2019)), but with sufficient number of sample size, the uncertainty associated with total emission quantified in an urban area is more precise.

3. The authors state that 'Emissions from facilities show significant contributions to the total emissions in both cities.' (supplement), but many details on the method used for estimating them are missing. This section needs to be expanded and more info should be given about the analysis as well as the uncertainty estimation and sensitivity of the method to the stated assumptions. For example, the use of measurement data from distant towers in order to drive the transport model raises an eyebrow, as these are critical for calculating the emission rate. What was the average distance between the tower and the measurement location and what was the elevation of that measurement? Are the wind speed and direction uncertainties reasonable? What about the elevation of the source, which is only very briefly discussed in the supplement? In the end, the reader should have a comprehensive view on whether the method is able to provide good emission estimation in a given setting, and at the moment the result with error bars (on the order of ±50 % 1-σ) suggest it is not. This should also be discussed in more detail.

- The focus of our paper is on the detection of methane elevations from unknown sources with the mobile vehicle approach and their attribution to natural gas, microbial and combustion sources.

In addition to this new attribution component, we also decided to report rough emission rates derived for larger known facilities, in contrast to previous publications that excluded emissions from such facilities. The reason was to demonstrate that such elevations are also picked up by the mobile measurements, to estimate whether estimates on emission rates can be derived, even if the approach was not targeted to facilities, and to put the order of magnitude of emissions from the unknown sources into perspective of the known sources (e.g. waste water treatment plants).
We kept this section short on purpose to not distract too much from the main objective but realize from the referee reports that a more detailed is warranted. Some more description has been incorporated in the revised version of the manuscript, and the replies to the specific questions are included in our answers to the detailed comments below.
In addition to the information which had been already provided in supplementary section we now have expanded the paragraphs related to the emission height (L333-341) and wind speed (L351-358) in Sect. 2.2.3 as follows:
Neumann and Halbritter (1980) showed that the main parameters in sensitivity analysis of GPDM are the wind speed and source emission height in close distance and the influence of emission height become less further downwind compared to the mixing layer height. In this study, the heights of emission sources were low (<10m) and/or estimated during surveys and Google Earth imageries, and considering that such a larger measurement distance from the facilities, the main sources of uncertainty of the emission estimates for the WWTP and Compost and Soil company are most likely the mean wind speed and for the upstream facilities in Hamburg the major sources of uncertainties can be the mean wind speed and emission height. We considered 0-4 m source height for the WWTP in Utrecht, and for the upstream facilities in Hamburg we considered 0-5 m emission height for the Compost and Soil site, 0-2 m for the separator, 0-10 m for the storage tank, and 0-1 m for the oil extraction well-head. We used 1 m interval for each of these height ranges to quantify emissions in GPDM.
Due to technical issues, local wind data were not logged continuously and thus we used wind data from two towers which are 8 to 20 km away from the facilities we focused for emission quantifications. These distances introduce extra uncertainties in analyzing the emissions using GPDM mainly on the wind speed. By comparing some of the local high-quality wind data to data from the towers, we estimated that the local wind speed is within the range of ± 30 % of the collected tower data. This range was adopted to estimate the wind speed for emission quantifications for the set of plumes measured downwind of the facilities. The wind directions were aligned at local scale of each facility based on the locations of sources and locations of maxima of average CH_4 enhancements from a set of transects in each day's survey and we considered ± 5° uncertainty in wind direction for the GPDM quantification.

4. The discussion section would benefit from introducing subsections to provide focus for specific items under discussion.

- This has been updated, and we agree that this provides a clearer structure.

5. I find the overall quality of language very high, yet there are multiple minor deficiencies that still need to be addressed. Below I have listed some of them. I believe that this is mostly due to heavy editing during manuscript preparation, and I ask the authors to take special care of that issue before resubmitting.

- Thanks for pointing out several detailed issues, which have all been incorporated (see detailed replies below). The manuscript has also been carefully proof-read again.

6. The supplement is large and – I'm sorry to say that – poorly edited (tables are too large - the font can safely be made much smaller; order of figures and sections does not correspond to the manuscript reference order). In some cases, itis a source of important information that is also in some places missing from the main text (already mentioned section 2.4. and corresponding S.2.4, figure S16). If the authors want to keep some technical details apart from the main text (understandable with that much material),

then I would ask to consider putting the more important parts in the Annex, in order to a) maintain the high editing standard and thus make reading easier, b) keep the important information together with the text. At the very minimum the editing of the supplement needs to be improved.

- The editing of the supplement has been updated according to the suggestions of the referee. We have moved some of the text from the supplement to the main text, as suggested.

3 Detailed comments:

Line numbers are given for identification. Comments for figures are given at the end of the list.

L25, also later in the text: ACP requires exponential notation of units, consult the 'manuscript preparation' on the ACP website for details (www.atmospheric-chemistry-and-physics.net/for_authors/manuscript_preparation.html)

- This has been adjusted in the revised version of the manuscript

L35: I'd suggest putting ppm outside of parentheses and mole fraction inside, as the ppm/ppb notation is the dominant one in the manuscript.

- The sentence was changed accordingly (L35-36): *The increase of CH_4 mole fraction from about 0.7 parts per million (ppm) or 700 parts per billion (ppb) …*

L46-49: Sentence needs rephrasing

- The sentence was rephrased as follows (see Sect. 1, L47-50): *CH_4 emissions originate from a wide variety of natural and anthropogenic sources, for example emissions from natural wetlands, agriculture (e.g. ruminants or rice agriculture), waste decomposition, or emissions (intended and non-intended) from oil and gas activities that are associated with production, transport, processing, distribution, and end-use of fossil fuel sector (Heilig, 1994).*

L60-62: This paragraph does not fit well here, would be better if info given as part of previous or next.

- The information has been added to the end of previous paragraph (L60-63).

L64: 'high precision' used twice

- Corrected

L75-77: What were the main findings from these studies? Specifically, it would be good to comment on whether these methods can be useful for up-scaling.

- The following information has been added to the manuscript (see Sect. 1, L78-86): *Gioli et al. (2012) showed that about 85 % of methane emissions in Florence, Italy originated from natural gas leaks. Helfter et al. (2016) estimated CH_4 emissions of 72 ± 3 t km^{-2} yr^{-1} in London, UK mainly from sewer sesytem and NGDNs leaks, which is twice as much as reported in the London Atmospheric Emissions Inventory. O'Shea et al. (2014) also showed that CH_4 emissions in greater London is about 3.4 times larger than the report from UK National Atmospheric Emission Inventory. Zimnoch et al. (2019) estimated CH_4 emissions of $(6.2 \pm 0.4) \times 10^6$ m^3 $year^{-1}$ for Krawko, Poland, based on data for the period of 2005 to 2008 and concluded that leaks from NGDNs are the main emission source in Krawko, based on carbon isotopic signature of CH_4. Chen et al. (2020) also showed that incomplete combustion or loss from temporarily installed natural gas appliances during big festivals can be the major source of CH_4 emissions from such events, while these emissions have not been included in inventory reports for urban emissions.*

L78: 'We quantified emissions in this study using measured CH_4 enhancements above background, which were detected' - This needs revision; also, it feels like Weller et al. 2019 should already be quoted here,

perhaps something like: 'In this study, we have quantified the CH_4 emissions using the method described by Weller et al., who demonstrated...'

- The following sentence has been added to the manuscript (see Sect. 1, L88-90): *In this study, we quantified LIs emissions using an empirical equation from Weller et al. (2019), which was designed based on controlled release experiments from von Fischer et al. (2017), to quantify ground-level emissions locations in urban area such as leaks from NGDN.*

L91: Was the reproducibility tested by the authors? Picarro currently gives 0.5 ppb for 5 s raw data. If the reproducibility was tested by the authors, please provide some details on the testing (either reference, or brief description of the experiment). Was the water correction modified or the factory settings were used? Please state that explicitly and also provide information if necessary.

- We now added more info about the G2301 to the manuscript (see Sect. 2.1.1, L100-108).
- The numbers represent the approximate range of the instrument noise when measuring background air. This information is provided to indicate to which order of magnitude we can identify elevations of CH_4 and C_2H_6.
- The factory settings were used for the water correction.

L99: Discard 'about' or the approximation sign

- Done

L100-101: Similar to comment for L91, please provide more info.

- Similar to the details provided for G230, information about the Picarro G4302 is now expanded. We also used factory setting to consider water correction (see Sect. 2.1.1, L109-117).

L104: Info on how the delay was calculated should be given here, but can be found later in L202. Please combine both (see major comment no. 1)

- Done

L111-112: Please spell out the main findings of the discussed comparison. Also, the reference to annex section number (S.2.) where it is discussed should be present (next to table S3 ref.). In general, sections of supplement should be referenced and not only tables or figures from it.

- The comparison sentence is now edited as follows with the main findings from the comparison (see Sect. 2.1.1, L135-142): *A comparison between the two instruments during simultaneous measurements showed that all LIs were detected by both instruments despite difference in instrument characteristics and inlet height. In the majority of cases CH_4 enhancements for each LI from both instruments were similar to each other. We note that there is likely a compensation of differences from two opposing effects between the two measurement systems. The inlet of the G2301 was at the bumper, thus closer to the surface sources, but the rather low flow rate and measurement rate of the instrument lead to some smoothing of the signal in the cavity. Because of the high gas flow rate, signal smoothing is much reduced for the G4302, but the inlet was on top of the car, thus further away from the surface sources (see Table S3 in SI, Sect. S.1.3).*

L129: When reading the sentence for the first time I have understood that the gas pipeline network corresponds to the street network 1:1. Is that correct, or the general coverage of municipal areas is meant? Please clarify.

- That is correct. The sentence was changed as follows (see Sect. 2.1.2, L162-165): *The local distribution companies (LDCs) in Utrecht (STEDIN (https://www.stedin.net/)) and Hamburg (GasNetz Hamburg (https://www.gasnetz-hamburg.de)) confirmed that full pipeline coverages are available beneath all streets. Therefore, the length of roads in the study area of Utrecht and Hamburg are representatives of NGDNs length.*

L137: discard 'at the following links: Utrecht and Hamburg'

- Done

L147-148: Please briefly explain how the vehicles can be methane sources (with reference for subsequent discussion further in the text).

- The sentence was changed as follows (see Sect. 2.1.3, L183-185): *For example, they could be related to emissions from vehicles which run on compressed natural gas, or vehicles operated with traditional fuels but with faulty catalytic converter systems. Later we will discuss how to exclude or categorize these unintended signals (see Sect. 2.2.2 and Sect. 2.3.1).*

L148: Please state clearly how many revisits were usually made.

- The respective sentence was changed as follows (see Sect. 2.1.3, L185-187): *Therefore, we revisited a large number of locations (65 in Utrecht (≈80 %) and 100 in Hamburg (≈70 %)) where enhanced CH_4 had been observed in during the first survey in order to confirm the LIs.*

L159-161: Have any cases where new leaks have occurred in-between surveys been observed?

- There have been some locations where we observed new LIs during revisits but not in the earlier visits or all the way round. This can be mainly due to the fact that not all LIs (mostly small LIs) are observable in all visits.

L166-167: Parentheses missing? 3 L bag for a price of 2 L bag is too good to be true.

- Corrected

L168-169: More details on sampling are needed.

- Was the data collection stationary or also mobile?
  The following sentence has been added to the manuscript (Sect. 2.1.4, L206-209):
  All the samples taken in the North Elbe study area and from most of the facilities were collected when the car was parked, but the samples inside the New Elbe tunnel and close to some facilities where there was no possibility to park were taken in motion while we were within the plume.

- How was the plume / non-plume location determined?
  The following sentences have been added to the manuscript (Sect. 2.1.4, L209-211):
  The sampling locations across the North Elbe study area of Hamburg were determined based the untargeted surveys, and the confirmation during revisits. The C_2H_6 information was not used in the selection of sampling locations in order to avoid biased sampling.

- What was the flushing time?
  The samples were taken using a pump with flow rate of 0.25 L min^{-1}. [info added to the sampling section (Sect. 2.1.4, L214)].

- Was the sample dried? How?
  Samples were dried in the lab followed by the CH_4 extraction. [info added to the lab analysis section (Sect. 2.3.2, L383-384 and L391-392)].

L173: See major comment no. 3.

- The relevant section has been expanded to provide more information (see Sect. 2.2.3). The reason for using data from the two mentioned towers in Utrecht and Hamburg is that the online data logging setup failed to continuously record all the local wind measurements during the surveys. The distance of the towers to the facilities ranges from 8 to 20 km, and indeed these distances introduce extra uncertainties in emission quantification mainly wind speed.
- When we compare the data that were recorded on the vehicle with the tower, we derive a difference in wind speed of ± 10 %. After considering the remarks of the referee, we increase this to a more conservative error estimate of ± 30 %. Regarding the wind direction (e.g. the oil wells),

for several sources the emission point is relatively certain and can by analysis of Google Earth images. In addition, we passed several sources during different wind conditions and did a "triangulation" based on the observed plumes and wind data (see Sect. 2.2.3, L352-359).

L191-192: Case shown in S5 is special and I strongly recommend to remove it from here and discuss later. As it is now, the text does not explain it, and thus may imply that all the cars are potential sources of CH_4, which is certainly not the case.

- Section 2 has been rearranged and this part has been moved to the sub-section of '2.3. Emission attributions' (Sect. 2.3.1, L367-370). The text has been revised as follows to mention that not all vehicles emit CH_4 but vehicles running on compressed natural gas.
  Compressed natural gas vehicles can be mobile CH_4 emission sources (E. K. Nam et al., 2004; Curran et al., 2014; Naus et al., 2018; Popa et al., 2014) and in this study we also observed CH_4 signals from vehicles. For example, the point to point C_2H_6:CH_4 ratio (C_2:C_1) calculated from road measurements of a car exhaust shown in Figure S12 is 14.2 ± 7.1 %.

L194: Consider providing these standards in the parentheses or the supplement section S.7.

- Done

L196: How many such cases were observed? Could they be important for the overall budget?

- In the study are of Hamburg (north Elbe), there were 34 cases and in the study area of Utrecht (inside the ring) there were 7 cases which we excluded based on this constraint (C2:C1 >10). Note that the methane/ethane instrument (Picarro G4302) was used for all surveys in Hamburg, while this instrument was not available for all surveys in Utrecht.
- The CH_4 enhancements measured by the G4302 for these cases were 0.37 ± 0.24 ppm in Hamburg and 0.26 ± 0.03 ppm in Utrecht. Based on the quantification from Weller et al. (2019), these LIs are not important to the overall budget but they will of course affect the total number of leak indications (LIs). Therefore, it is important to exclude them from the evaluation.

L198: Please state the reasons for this exclusion, briefly.

- The speed constraints are now reformulated as follows (see Sect. 2.2.2, L253-256): *Following the algorithm from von Fischer et al. (2017), measurements at speeds above 70 km h^{-1} were excluded, as the data from the controlled release experiments (von Fischer et al., 2017) were not reliable at high speed (Weller et al., 2019). We also excluded measurements during periods of zero speed (stationary vehicle) to avoid unintended signals coming from other cars when the measurement car was stopped in traffic.*

L199-200: Just a small comment, no action needed: I don't see the benefit of this artificial increase of the data frequency. This brings no new information at the cost of tripling of the data that needs to be processed.

L200-201: Wording. If the time was just converted to UTC, then calling it 'a correction' is not warranted. Consider changing to: 'Following the interpolation step, the data was converted to UTC, and subsequently corrected for ...'

- We used the word 'correction' as the clock on the Picarro instruments had a drift over the period of the campaign and we needed to correct this drift to set all the measurements to the correct UTC time. The text has been adjusted as follows:
  The clocks on the Picarro instruments were set to UTC but showed drift over the period of the campaigns. We recorded the drifts for each day's survey and corrected to UTC time. The data were also corrected for the delay between air at the inlet and the signal in the CH_4 analyzers. This delay was determined by exposing the inlet to three small CH_4 pulses from exhaled breath, ranging from 5-30 seconds, depending on the instrument and tubing length. We averaged the three attempts to determine the delay for each instrument and used the delays for each instrument.

Individual attempts were 1 to 2 s different from each other. For the G4302 the delay was generally about 5 s and for the G2301 it was about 30 s; the difference is mainly due to the different flow rates. The recorded CH_4 mole fractions were projected back along the driving track according to this delay.

L202: About the delay time estimation: 5–30 seconds is a very broad range. Were the ranges so variable for both instruments, or was it 5 for one and 30 for the other? Also, how was the pulse generated? Can you estimate precision of that delay estimation (even grossly)?

- The answers to this comment are merged with the answers to the previous comment (see Sect. 2.1.1, L122-129).

L204: Reference order needs correction. Previous reference supplement figure was S5 (L190).

- Corrected

L207: In CO_2 signal (Fig. S8), it can be clearly seen that the background line is some-times higher than the observed signal. Since this plot is about the background, it would be good to change the limits of y axis to make the calculated background visible clearer, especially for methane. Please give some comment about the possible negative enhancements after subtracting such background (can it affect the estimation of emissions?).

- For CO_2, individual plumes, e.g. from vehicles, can overlap and create a locally enhanced background according to our background extraction procedure (5^{th} percentile of CO_2 measurements in a ± 2.5 min time window). Negative deviations are not considered, but the enhanced background may result in a potential small underestimate of the CO_2 enhancement. Such small changes will not affect the categorization of the CH_4 enhancement based on the CH_4:CO_2 ratio, very wide ranges were assumed for this source attribution. We have updated the figure in the revised version as suggested (see Figure S7 in SI, Sec. S.2.1).

L217: Please add 'peak' after enhancement, to make it clear that it's not about the release height.

- We changed this to: *…to demonstrate that the magnitude of the observed methane enhancement…* (see Sect. 2.2.2, L264)

L237: Wording. Why should results from different cities be comparable? The authors clearly mean that the analysis software used on a given dataset should be comparable. Please clarify. Actually, this whole paragraph can be limited to information that 'Our software was compared to analysis tools developed by CSU (von Fisher et al.2017, Weller et al. 2019) and no significant differences were observed (see SI, section S.2.7)'.

- Thank you for the suggestion, we updated the paragraph as follows (see Sect. 2.2.2, L284-286): *We compared the outputs of our software to the one developed by Colorado State University (CSU) for the surveys in US cities (von Fischer et al., 2017; Weller et al., 2019). 30 LIs were detected and no significant differences were observed (linear fit equation y = 1.00 * x - 0.00, R^2 = 0.99) (see SI, Sect. S.2.4, Figure S10).*

L251: Erase 'areas'.

- Done

L252: Erase empty parentheses.

- Done

L253 and L384: 'drive-by'- I propose 'mobile'. I was surprised to find it used in Fisher et al. (albeit only once), as in U.S. this word is sometimes used to describe something much more nefarious then GHG observations.

- Done

L254: 'We report (...)' – Unclear what is meant in this sentence, please rephrase it.

- The sentence was rephrased as follows (Sect. 2.2.3, L309-310): *In this study, we also report the data obtained from larger facilities, since rough emission estimates from facilities can be obtained in the city surveys.*

L256-257: Erase 'both' and 'each day's'

- Done

L269-274: How was the release height determined? How is the uncertainty of this determination included in the uncertainty of emission?

- Information on the release height and uncertainty associated with each facility has been added to this section (see answer to the major comment 3).

L283: This section should be combined with 2.4.

- Section 2 has been rearranged and information regarding air sampling collection is now provided in Sect. 2.1.4 and details on analysis of samples are provided in Sect. 2.3.2.

L286: Info on the isotopic scales used in this study needs to be given.

- The information on the scale and instruments for isotopic measurement is now added (see Sect. 2.3.2, L380-383 and L393-395).

L293-294: Please provide explanation on why these particular ranges were selected. For signatures, specifically also provide references supporting the choice of isotopic signatures. Please keep in mind that for fossil fuel signatures, figure 7 from Rockmann et al. 2016 doesn't give a full picture – see e.g. Sherwood et al. 2017 for a broad overview of isotopic signatures for fossil methane.

- We acknowledge that the full range of isotopic signatures of natural gas from different reservoirs is wide, in particular for $\delta^{13}C$, as documented in Sherwood et al. (2017), and we have indeed adjusted this range in a recent publication of Menoud et al., (2020). Nevertheless, this does not mean that the full range of signatures would be encountered in one region in a single campaign. No data are available for Hamburg, but Levin et al., (1999) reported a mean $\delta^{13}C$ value of -40 ‰ for the NGDN in Heidelberg, with a seasonal cycle of ± 10‰. Hoheisel et al. (2019) reported $\delta^{13}C$ = -43.3 ± 0.8‰ for the NGDN in Heidelberg the period 2016-2018, but without the strong seasonality, and -46.1 ± 0.8‰ for natural gas storage tanks and compressor stations. These values are compatible with the selected range that we have chosen based on Röckmann et al. (2016).

L314-315: 'were correspond' – corresponded

- Done

L315: What is the uncertainty here? This relates to major comment 2.

- See the answer to major comment 2. We have also added the following sentence (Sect. 3, L422-424): *Noted that estimates for individual leaks with the Weller et al. (2019) algorithm can have large error, thus these results are indicative of large leaks, but the precise emission strength is very uncertain.*

L378: Why uncertainty given only for wind direction?

- Corrected

L394: von Fisher – V should be capitalised at the start of the sentence (Von Fisher).

- Done

L404: 'About 50 %' - please give the specific number that was used in the calculation.

- We used precisely 50 % for both cities for the calculation. The wording has been adjusted.

L420-422: These numbers are in fact very similar. The variability of $\delta^{13}C$ in the natural gas can be quite substantial. See e.g. Fig 4 in Sherwood et al, 2017.

- We have updated this sentence as follows (see Sect. 4.2, L533-537): *These numbers do not agree within combined errors, but are also not very different. $\delta^{13}C$ values of CH_4 from the NGDN can vary regionally and temporally, e.g. due to differences in the mixture of natural gas from various suppliers for different regions in Germany (DVGW, 2013). It is also shown that how $\delta^{13}C$ values of fossil fuel CH_4 have significant variabilities in different regions at global scale (Figure 4 in Sherwood et al. (2017)).*

L426: % used twice

- Corrected

L433: Reference to Figure 1?

- Corrected

L436-438: a) Fig S16 also points to the local sewage system as potentially important source, but this is not mentioned here. b) Please be more precise in the argument here- i.e. explain why measurements around the lake point to anaerobic methanogenesis specifically. Linking to a), please include info on the potential role of the sewage system if needed. Is it possible that the sewage is seeping into the lake?

a) We changed this sentence as follows (see Sect. 4.2, L553-559): *Many of the microbial LIs encountered in Hamburg are around the Binnenalster lake (Figure S15), which suggests that anaerobic methanogenesis (Stephenson and Stickland, 1933; Thauer, 1998) can cause these microbial emission in this lake, as seen in other studies focused on emissions from other lakes (DelSontro et al., 2018; Townsend-Small et al., 2016). Microbial CH_4 emissions from sewage system (Guisasola et al., 2008) can also be an important source of in this area, as seen in US urban cities (Fries et al., 2018). Fries et al. (2018) performed direct measurement of CH_4 and nitrous oxide (N_2O) from a total of 104 sites, and analyzed $\delta^{13}C$ and δD signatures of samples from 27 of these locations, and attributed 47 % of these locations to microbial emissions in Cincinnati, Ohio, USA.*
b) We described in the text that the isotope and multi-tracer observations point to microbial sources. Unfortunately, we do not know if and/or to what degree the sewage system seeps into the lake.

L443: 'because there is no publicly available activity data for associated activity data'– please rephrase.

- The sentence was rephrased as follows (see Sect. 4.3, L564-566): *Also, it is not possible to calculate a robust city-level estimate using the nationally reported emission factors because there is no publicly available associated activity data, i.e., pipeline materials and lengths for each material, at the level of individual cities.*

L452: Too many parentheses. I suggest '(...) 40 kg km^{-1} yr^{-1} (for other material, p <200 mbar; see p. 130 in Peek et al., 2019, for details)

- The sentence was split in several parts and rewritten as follows (see Sect. 4.3, L571-577): *The Netherlands National Institute for Public Health and the Environment (RIVM) inventory report derived an average NGDN emission factor of ≈ 110 kg km^{-1} yr^{-1} using 65 leak measurements from different pipeline materials and pressures in 2013. This weighted average ranged from a maximum of 230 kg km^{-1} yr^{-1} for grey cast iron pipelines to a minimum of 40 kg km^{-1} yr^{-1} for pipelines of other materials with overpressures <= 200 mbar (for details, see P. 130 in Peek et al. (2019)). This results in an average CH_4 emissions of ≈ 70 t yr^{-1} (min = 30 t yr^{-1} and max = 150 t yr^{-1}) for the study area of Utrecht, assuming ≈ 650 km of pipelines inside the ring, and further assuming that Utrecht's NGDN is representative of the national reported average (see qualifiers above).*

L473: ‘credibility interval’ – confidence interval

- Done

L491: ’(...) factors can be gas pipeline age and material, sewer system.’ Part of sentence missing? Please rephrase.

- The sentence was changed as follows (see Sect. 4.3, L611-613): *CH_4 emissions can vary among different cities, depending on the age, management and material of NGDNs, and/or the management of local sewer systems.*

L531: ‘were’ – where

- Corrected

L533: ‘as shown’ used twice.

- Corrected

L542-544: The scheme from S18 cannot be treated as a ‘protocol’ without a proper description of the method. In reality, it describes the main components of the method applied in the study, so in fact the manuscript itself is more of such a protocol. As it stands now, consider either expanding the description in the supplement (so gas companies might actually use it as a protocol) or discard it altogether.

- We agree with this comment and removed the term “protocol”. We reformulated the sentence as follows (see Sect. 4.4, L668-669): *Figure S19 (see SI, Sect. S.5) illustrates how the individual measurement components can be efficiently combined in a city leak survey program.*

L561: ‘corresponding to emissions of about 107 tCH_4/ yr’ – exactly 110 t yr^{-1} is given in L332.

- Corrected

L562: Please state the method, e.g. ’These estimates, based only on the studied area, were then up-scaled for the total municipal area, using the road network map as a proxy to (...)’

- The up-scaled emissions are related to the emissions across the study area; inside the ring in Utrecht and north Elbe in Hamburg and not the total municipal area of these two cities. The text has been adjusted as follows (see Sect. 5, L688-690): *These estimates, based on the streets covered, were then up-scaled to the total study area, using the road network map as a proxy for the length of the pipeline network which then yielded total emissions of 150 t yr^{-1} and 440 t yr^{-1} across the study area of Utrecht and Hamburg respectively.*

L567: ‘were from’ - I suggest ‘originated from’

- ‘were from’ has been changed to ‘*originated from*’ (L695)

Figure 2: a) Please fix the x axis description - extra arrow unnecessary. b) the plot is cropped in the lower part, by several pixels. c) extra grid dashed lines (green) are unnecessary, make the labels difficult to read.

a) We prefer to keep the arrow. It indicates that the total emissions are attributed to four different emission sources and how much these sources contribute to the total emissions individually
b) Done, this probably happened during conversion to pdf format.
c) The extra green dashed lines on the x-axis have been removed

The caption has been changed as follows: *Total CH_4 emission rates from different sources in (a) Utrecht and (b) Hamburg; the arrow shows how the emissions are attributed to different sources*

Figure 3: a) again, the arrow from total emissions to ’Road classes’ seems unnecessary. b) please explain the arrows from the plot in the caption.

a) The arrow is used to indicate that the total emissions are categorized into six different road classes and how much these classes contribute to the total emissions individually

b) the caption has been changed as follows: *Total CH_4 emissions in Utrecht and Hamburg; the arrow shows how the total emissions are distributed over different road classes*

Figure 4: I recommend plotting all US cities in a single colour (grey?) and simply label the line as ('15 US Cities', Weller et al.) or similar. The colours are indistinguishable anyway.

- Done

Figure 5: Excellent plot! a) Please add comment about the uncertainty of $\delta^{13}C$ and $\delta^{2}H$ signatures in the caption (they were only plotted for C2/C1. I also softly suggest to label the plots with a-b-c-d and move the labels into figure caption. b) On previous plot the units were placed in parentheses, consider keeping notation consistent.

- Both suggestions have been included

Figure 6: a) Plumes of what? b) Please also provide information on whether the peaks are on the same scale; if yes, then what is the plotted range of mole fractions (if those are mole fractions)?

- Done

References

DelSontro, T., Beaulieu, J. J. and Downing, J. A.: Greenhouse gas emissions from lakes and impoundments: Upscaling in the face of global change, Limnol. Oceanogr. Lett., 3(3), 64–75, doi:10.1002/lol2.10073, 2018.

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(7), 4091–4099, doi:10.1021/acs.est.6b06095, 2017.

Hoheisel, A., Yeman, C., Dinger, F., Eckhardt, H. and Schmidt, M.: An improved method for mobile characterisation of δ 13 CH 4 source signatures and its application in Germany, Atmos. Meas. Tech., 12(2), 1123–1139, doi:10.5194/amt-12-1123-2019, 2019.

Levin, I., Glatzel-Mattheier, H., Marik, T., Cuntz, M., Schmidt, M. and Worthy, D. E.: Verification of German methane emission inventories and their recent changes based on atmospheric observations, J. Geophys. Res. Atmos., 104(D3), 3447–3456, doi:10.1029/1998JD100064, 1999.

Menoud, M., van der Veen, C., Scheeren, B., Chen, H., Szénási, B., Morales, R. P., Pison, I., Bousquet, P., Brunner, D. and Röckmann, T.: Characterisation of methane sources in Lutjewad, The Netherlands, using quasi-continuous isotopic composition measurements, Tellus B Chem. Phys. Meteorol., 72(1), 1–19, doi:10.1080/16000889.2020.1823733, 2020.

Peek, C. J., Montfoort, J. A., Dröge, R., Guis, B., Baas, K., Huet, B. van, Hunnik, O. R. van and Berghe, A. C. W. M. van den: Methodology report on the calculation of emissions to air from the sectors Energy, Industry and Waste, as used by the Dutch Pollutant Release and Transfer Register., 2019.

Sherwood, O. A., Schwietzke, S., Arling, V. A. and Etiope, G.: Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Burning Sources, version 2017, Earth Syst. Sci. Data, 9(2), 639–656, doi:10.5194/essd-9-639-2017, 2017.

Townsend-Small, A., Disbennett, D., Fernandez, J. M., Ransohoff, R. W., Mackay, R. and Bourbonniere, R. A.: Quantifying emissions of methane derived from anaerobic organic matter respiration and natural gas extraction in Lake Erie, Limnol. Oceanogr., 61(S1), S356–S366, doi:10.1002/lno.10273, 2016.

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 M. Mauder, PLoS One, 14(2), e0212287, doi:10.1371/journal.pone.0212287, 2019.

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, Environ. Sci. Technol., acs.est.8b03135, doi:10.1021/acs.est.8b03135, 2018.

Our answers to Luise Westphal:

Gasnetz Hamburg could not confirm 80 % of the LI as pipeline leaks. This issue requires further investigation. Therefore, we set up a joint project together with IMAU. The field test campaign in Hamburg is ongoing. The objective of the project is to compare leak rate estimates from mobile methods with ground measurements applying the suction method for a small sample of leaks in a real-life situation. For this reason, we request to give more explanation on that statement, e.g. by the conditional "[…] once the LIs were shared. Further, it must be considered that the leak detection of the gas utility and University of Utrecht did not take place at the same time (several weeks in between). It might be possible that changing weather and soil conditions prevented finding leaks on different events. Furthermore, a "fossil leak" does not necessarily originate from a pipeline. It could also come from natural gas vehicles, thus, it is only presented for a very short time. We are highly confident, that regular LDAR (Leak Detection and Repair) is capable of finding the vast majority of leaks. Accordingly, we suggest rewording the sentence for example to "Gasnetz Hamburg could not confirm 80 % of the LI as pipeline leaks. This issue requires further Investigation."

We acknowledge the fact that no leaks were found at a large number of locations could have several reasons. Therefore, we changed the respective sentence to (see Abstract, L30-32): *The largest leaks were located and fixed quickly by GasNetz Hamburg once the LIs were shared, but 80 % of the (smaller) LIs attributed to the fossil category could not be detected/confirmed as pipeline leaks. This issue requires further investigation.*

We want to specify that in our algorithm emissions from vehicles are identified (when attribution is possible) by using the co-emitted species C_2H_6 and CO_2. The 80% of LIs where no leaks were detected by GNH refers to the LIs that we have attributed to the category "fossil", which should be specific for leaks from the NGDN.

Methane mapping, emission quantification, and attribution in two European cities; Utrecht, NL and Hamburg, DE

Hossein Maazallahi[1,2], Julianne M. Fernandez[3], Malika Menoud[1], Daniel Zavala-Araiza[1,4], Zachary D. Weller[5], Stefan Schwietzke[6], Joseph C. von Fischer[7], Hugo Denier van der Gon[2], and Thomas Röckmann[1]

[1]Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University (UU), Utrecht, The Netherlands
[2]Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, The Netherlands
[3]Department of Earth Sciences, Royal Holloway University of London (RHUL), Egham, United Kingdom
[4]Environmental Defense Fund (EDF), Utrecht, The Netherlands
[5]Department of Statistics, Colorado State University (CSU), United States of America
[6]Environmental Defense Fund (EDF), Berlin, Germany
[7]Department of Biology, Colorado State University (CSU), United States of America

Correspondence to: Hossein Maazallahi (h.maazallahi@uu.nl)

Abstract. Characterizing and attributing methane (CH_4) emissions across varying scales is important from environmental, safety, and economic perspectives, and is essential for designing and evaluating effective mitigation strategies. Mobile real-time measurements of CH_4 in ambient air offer a fast and effective method to identify and quantify local CH_4 emissions in urban areas. We carried out extensive campaigns to measure CH_4 mole fractions at the street level in Utrecht, The Netherlands (2018 and 2019) and Hamburg, Germany (2018). We detected 145 leak indications (LIs, i.e., CH_4 enhancements of more than 10% above background levels) in Hamburg and 81 LIs in Utrecht. Measurements of the ethane-to-methane ratio (C_2:C_1), methane-to-carbon dioxide ratio (CH_4:CO_2), and CH_4 isotope composition ($\delta^{13}C$ and δD) show that in Hamburg about 1/3 of the LIs, and in Utrecht 2/3 of the LIs (based on a limited set of C_2:C_1 measurements), were of fossil fuel origin. We find that in both cities the largest emission rates in the identified LI distribution are from fossil fuel sources. In Hamburg, the lower emission rates in the identified LI distribution are often associated with biogenic characteristics, or partly combustion. Extrapolation of detected LI rates along the roads driven to the gas distribution pipes in the entire road network yields total emissions from sources that can be quantified in the street-level surveys of 440 ± 70 t yr^{-1} from all sources in Hamburg, and 150 ± 50 t yr^{-1} for Utrecht. In Hamburg, C_2:C_1, CH_4:CO_2, and isotope-based source attributions shows that 50 - 80 % of all emissions originate from the natural gas distribution network, in Utrecht more limited attribution indicates that 70 - 90 % of the emissions are of fossil origin. Our results confirm previous observations that a few large LIs, creating a heavy tail, are responsible for a significant proportion of fossil CH_4 emissions. In Utrecht, 1/3 of total emissions originated from one LI and in Hamburg >1/4 from 2 LIs. *The largest leaks were located and fixed quickly by GasNetz Hamburg once the LIs were shared, but 80 % of the (smaller) LIs attributed to the fossil category could not be detected/confirmed as pipeline leaks. This issue requires further investigation.*

1 Introduction

Methane (CH_4) is the second most important anthropogenic greenhouse gas (GHG) after carbon dioxide (CO_2) with a global warming potential of 84 compared to CO_2 over a 20-year time horizon (Myhre et al., 2013). The increase of CH_4 mole fraction from about *0.7 parts per million (ppm) or 700 parts per billion (ppb)* in pre-industrial times (Etheridge et al., 1998; MacFarling Meure et al., 2006) to almost 1.8 ppm at present (Turner et al., 2019) is responsible for about 0.5 W m^{-2} of the total 2.4 W m^{-2} radiative forcing since 1750 (Etminan et al., 2016; Myhre et al., 2013). In addition to its direct radiative effect, CH_4 plays an important role in tropospheric chemistry and affects the mixing ratio of other atmospheric compounds, including direct and indirect greenhouse gases, via reaction with the hydroxyl radical (OH), the main loss process of CH_4 (Schmidt and Shindell, 2003). In the stratosphere CH_4 is the main source of water vapor (H_2O) (Noël et al., 2018), which adds another aspect

to its radiative forcing. Via these interactions the radiative impact of CH_4 is actually higher than what can be ascribed to its mixing ratio increase alone, and the total radiative forcing ascribed to emissions of CH_4 is estimated to be almost 1 W m^{-2}, ≈ 60 % of that of CO_2 (Fig 8.17 in Myhre et al., 2013). Given this strong radiative effect, and its relatively short atmospheric lifetime of about 9.1 ± 0.9 yr (Prather et al., 2012), CH_4 is an attractive target for short- and medium-term mitigation of global climate change as mitigation will yield rapid reduction in warming rates.

CH_4 emissions originate from a wide variety of natural and anthropogenic sources, for example emissions from natural wetlands, agriculture (e.g. ruminants or rice agriculture), waste decomposition, or emissions (intended and non-intended) from oil and gas activities that are associated with production, transport, processing, distribution, and end-use of fossil fuel sector (Heilig, 1994). Fugitive unintended and operation-related emissions occur across the entire oil and natural gas supply chain. In the past decade, numerous large studies have provided better estimates of the emissions from extended oil and gas production basins (Allen et al., 2013; Karion et al., 2013; Omara et al., 2016; Zavala-Araiza et al., 2015; Lyon et al., 2015), the gathering and processing phase (Mitchell et al., 2015), and transmission and storage (Zimmerle et al., 2015; Lyon et al., 2016) in the United States (US). A recent synthesis concludes that the national emission inventory of the US Environmental Protection Agency (EPA) underestimated supply chain emissions by as much as 60 % (Alvarez et al., 2018). McKain et al. (2015) discussed how inventories may underestimate the total CH_4 emission for cities. Also, an analysis of global isotopic composition data suggests that fossil related emissions may be 60 % higher than what has been previously estimated (Schwietzke et al., 2016). A strong underestimate of fossil fuel related emissions of CH_4 was also implied by analysis of $\delta^{14}C$-CH_4 in pre-industrial air (Hmiel et al., 2020). These emissions do not only have adverse effects on climate, but also represent an economic loss (Xu and Jiang, 2017) and a potential safety hazard (West et al., 2006). *While CH_4 is the main component in natural gas distribution networks (NGDNs), composition of natural gas varies from one country or region to another. In Europe the national authorities provide specifications on components of natural gas in the distribution network (Table 8 in UNI MISKOLC and ETE, 2008).*

Regarding CH_4 emissions from NGDNs, a number of intensive CH_4 surveys with novel mobile high precision laser-based gas analyzers in US cities have recently revealed the widespread presence of leak indications (LIs: CH_4 enhancements of more than 10 % above background level) with a wide range of magnitudes (Weller et al., 2020; Weller et al., 2018; von Fischer et al., 2017; Chamberlain et al., 2016; Hopkins et al., 2016; Jackson et al., 2014; Phillips et al., 2013). The number and severity of natural gas leaks appears to depend on pipeline material and age, local environmental conditions, pipeline maintenance and replacement programs (von Fischer et al., 2017; Gallagher et al., 2015; Hendrick et al., 2016). For example, NGDNs in older cities with a larger fraction of cast iron or bare steel pipes showed more frequent leaks than NGDNs that use the newer plastic pipes. The data on CH_4 leak indications from distribution systems in cities have provided valuable data for emission reduction in the US cities which allows local distribution companies (LDCs) who are in charge of NGDN to quickly fix leaks and allocate resources efficiently (Weller et al., 2018, von Fischer et al., 2017, Lamb et al., 2016; McKain et al., 2015).

Urban European cities CH_4 emissions are not well known, which requires carrying out extensive campaigns to collect required observation data. Few studies have estimated urban CH_4 fluxes using eddy covariance measurements (Gioli et al., 2012; Helfter et al., 2016), airborne mass balance approaches (O'Shea et al., 2014) and the Radon-222 flux and mixing layer height techniques (Zimnoch et al., 2019). *Gioli et al. (2012) showed that about 85 % of methane emissions in Florence, Italy originated from natural gas leaks. Helfter et al. (2016) estimated CH_4 emissions of 72 ± 3 t km^{-2} yr^{-1} in London, UK mainly from sewer sesytem and NGDNs leaks, which is twice as much as reported in the London Atmospheric Emissions Inventory. O'Shea et al. (2014) also showed that CH_4 emissions in greater London is about 3.4 times larger than the report from UK National Atmospheric Emission Inventory. Zimnoch et al. (2019) estimated CH_4 emissions of (6.2 ± 0.4) × 10^6 m^3 $year^{-1}$ for Krawko, Poland, based on data for the period of 2005 to 2008 and concluded that leaks from NGDNs are the main emission*

source in Krawko, based on carbon isotopic signature of CH_4. Chen et al. (2020) also showed that incomplete combustion or loss from temporarily installed natural gas appliances during big festivals can be the major source of CH_4 emissions from such events, while these emissions have not been included in inventory reports for urban emissions.

Here we present the result of mobile in-situ measurements at street level for whole-city surveys in two European cities, Utrecht in the Netherlands (NL) and Hamburg in Germany (DE). *In this study, we quantified LIs emissions using an empirical equation from Weller et al. (2019), which was designed based on controlled release experiments from von Fischer et al. (2017), to quantify ground-level emissions locations in urban area such as leaks from NGDN.* In addition to finding and categorizing the CH_4 enhancements (in a similar manner as done for the US cities in order to facilitate comparability), we made three additional measurements to better facilitate source attribution: the concomitant emission of ethane (C_2H_6) and CO_2, and the carbon and hydrogen isotopic composition of the CH_4. These tracers allow an empirically based source attribution for LIs. In addition to emission quantifications of LIs across the urban areas in these two cities, we also quantified CH_4 emissions from some of facilities within the municipal boundary of Utrecht and Hamburg using Gaussian plume dispersion model (GPDM).

2 Materials and methods

2.1 Data collection and instrumentation

2.1.1 Mobile measurements for attribution and quantification

Mobile atmospheric measurements at street level were conducted using two Cavity Ring-Down Spectroscopy (CRDS) analyzers (Picarro Inc. model G2301 and G4302) which were installed on the back seat of a 2012 Volkswagen Transporter, (see supplementary information (SI), Sect. S.1.1, Figure S1). The model G2301 instrument provides atmospheric mole fraction measurements of CO_2, CH_4 and H_2O, each of them with an integration time of about 1 s., which results in a data frequency of ≈ 0.3 Hz for each species. The reproducibility for CH_4 measurements was ≈ 1 ppb for 1 s integration time. The G2301 instrument was powered by a 12 V car battery via a DC-to-AC converter. The flow rate was ≈ 187 ml min^{-1}. Given the volume and pressure of the measurement cell (volume = 50 ml and pressure ≈ 190 mbar) the cell is flushed approximately every 3 s, so observed enhancements are considerably smoothed out. *The factory settings for CH_4 and CO_2 were used for the water correction.*

The G4302 instrument is a mobile analyzer that provides atmospheric mole fraction measurements of C_2H_6, CH_4, and H_2O. *The flow rate is 2.2 L min^{-1} and the volume of the cell is 35 ml (operated at 600 mb, thus 21 ml STP) so the cell is flushed in 0.01 s, which means that mixing is insignificant given the 1 s measurement frequency of the G4302.* The additional measurement of C_2H_6 is useful for source attribution since natural gas almost always contains a significant fraction of C_2H_6, whereas microbial sources generally do not emit C_2H_6 (Yacovitch et al., 2014). The G4302 runs on a built-in battery which lasts for ≈ 6 h. The instrument can be operated in two modes at ≈ 1 Hz frequency for each species: the CH_4-only mode and the CH_4 - C_2H_6 mode. In the CH_4-only mode the instrument has a reproducibility of ≈10 ppb for CH_4. *The factory settings for CH_4 and C_2H_6 were used for the water correction.* In the CH_4 - C_2H_6 mode the reproducibility is about 100 ppb for CH_4 and 15 ppb for C_2H_6. For Utrecht surveys (see SI, Sect. S.1.2, Figure S2a), the G4302 was not yet available for the initial surveys in 2018, but it was added for the later re-visits (see SI, Sect. S.1.2, Table S1). For Hamburg (see SI, Sect. S.1.2, Figure S2b), both instruments operated during the entire intensive 3-week measurement campaign in Oct/Nov 2018 (see SI, Sect. S.1.2, Table S2). The time delay from the inlet to the instruments was measured and accounted for in the data processing procedure. The Coordinated Universal Time (UTC) time shifts between the Global Positioning System (GPS) and the two Picarro instruments were corrected for each instrument in addition to the inlet delay (see SI, Sect. S.1.2, Table S1 and Table S2). The clocks on

the Picarro instruments were set to UTC but showed drift over the period of the campaigns. We recorded the drifts for each day's survey and corrected to UTC time. The data were also corrected for the delay between air at the inlet and the signal in the CH_4 analyzers. This delay was determined by exposing the inlet to three small CH_4 pulses from exhaled breath, ranging from 5-30 seconds, depending on the instrument and tubing length. We averaged the three attempts to determine the delay for each instrument and used the delays for each instrument. Individual attempts were 1 to 2 s different from each other. For the G4302 the delay was generally about 5 s and for the G2301 it was about 30 s; the difference is mainly due to the different flow rates. The recorded CH_4 mole fractions were projected back along the driving track according to this delay.

One-quarter inch Teflon tubing was used to pull in air either from the front bumper (0.5 m above ground level) to the G2301 or from the rooftop (2 m above ground level) to the G4302. To avoid dust into the inlets for both instruments, Acrodisc® syringe filter, 0.2 μm was used for G2301 and Parker Balston 9933-05-DQ was used for G4302. The G2301 was used for quantification and attribution purposes and the G4302 mainly for attribution. After data quality check, a comparison between the two instruments during simultaneous measurements showed that all LIs were detectable by both instruments despite difference in inlet height (see SI, Sect. S.1.3, Figure S3). *A comparison between the two instruments during simultaneous measurements showed that all LIs were detected by both instruments despite difference in instrument characteristics and inlet height. In the majority of cases CH_4 enhancements for each LI from both instruments were similar to each other. We note that there is likely a compensation of differences from two opposing effects between the two measurement systems. The inlet of the G2301 was at the bumper, thus closer to the surface sources, but the rather low flow rate and measurement rate of the instrument lead to some smoothing of the signal in the cavity. Because of the high gas flow rate, signal smoothing is much reduced for the G4302, but the inlet was on top of the car, thus further away from the surface sources (see Table S3 in SI, Sect. S.1.3).* The vehicle locations were registered using a GPS system that recorded the precise driving track during each survey.

2.1.2 Target cities: Utrecht and Hamburg

Utrecht is the 4th largest city in the Netherlands with population of approximately 0.35 million inhabitants within an area of roughly 100 km^2. It is located close to the center of the Netherlands and is an important infrastructural hub in the country. The Utrecht city area that we target in this study is well constrained by a ring of highways around the city (A27, A12, A2, and N230) with inhabitants of approximately 0.28 million living within this ring on roughly 45 km^2 of land. Figure S2a (see SI, Sect. S.1.2) shows the streets that were driven in Utrecht and Figure 1a shows the street coverage over four street categories (level 1, 2, 3, residential, and unclassified) obtained from the Open Street Map (OSM; www.openstreetmap.org). Table S4 (see SI, Sect. S.1.5) provides information on road coverage based on different street categories. The hierarchy of OSM road classes is based on the importance of roads in connecting parts of the national infrastructure. *Level 1 roads are primarily larger roads connecting cities, level 2 roads are the second most important roads and part of a greater network to connect smaller towns, level 3 roads have tertiary importance level and connect smaller settlements and districts.* Residential roads are roads which connect houses and unclassified roads have the lowest importance of interconnecting infrastructure. Moreover, several transects were also made to measure the atmospheric mole fraction of CH_4 from the road next to the waste water treatment plant (WWTP) in Utrecht – a potentially larger single source of CH_4 emissions in the city (see SI, Sect. S.1.6, Table S5).

Hamburg is the 2nd largest city in Germany (about 1.9 million inhabitants, 760 km^2 area) and hosts one of the largest harbors in Europe. The study area in Hamburg is North of the Elbe river (Figure 1b) with ≈1.4 million inhabitants on about 400 km^2 land. Figure S2b (see SI, Sect. S.1.2) shows the streets that were covered in Hamburg and Figure 1b shows the street coverage categorized in the four categories of OSM. More information on road coverage based on OSM street categories are provided in Table S4 (see SI, Sect. S.1.5). *The local distribution companies (LDCs) in Utrecht (STEDIN (https://www.stedin.net/)) and Hamburg (GasNetz Hamburg (https://www.gasnetz-hamburg.de)) confirmed that full pipeline*

coverages are available beneath all streets. Therefore, the length of roads in the study area of Utrecht and Hamburg are representatives of NGDNs length. The Hamburg harbor area hosts several large industrial facilities that are related to the midstream / downstream oil and gas sector including refineries and storage tanks. An oil production site (oil well, separator and storage tanks) at Allermöhe (in Hamburg-Bergedorf) was also visited. Information from the State Authority for Mining, Energy and Geology (LBEG, 2018) was used to locate facilities. Precise locations of the facilities surveyed are given in the Table S6 (see SI, Sect. S.1.6). In order to separate these industrial activities from the NGDNs emissions in this study, CH_4 emissions from these locations were estimated, but evaluated apart from the emissions found in each city. The reported in-situ measurement, GPS data, and boundary of study areas reported here are available on the Integrated Carbon Observation System (ICOS) portal (Maazallahi et al., 2020b).

2.1.3 Driving strategy

The start/end point for each day's measurement surveys across Utrecht and Hamburg were the Institute for Marine and Atmospheric research Utrecht (IMAU; Utrecht University) and the Meteorological Institute (MI; Hamburg University), respectively. From these starting locations, each day's surveys targeted the different districts and neighborhoods of the cities (see SI, Sect. S.1.2, Table S1 and Table S2). Measurement time periods and survey areas were chosen to select favorable traffic and weather conditions and to avoid large events (e.g., construction; see SI, Sect. S.1.5, Figure S4), which normally took place between 10 - 18 LT. Average driving speeds on city streets were in the range of 17 ± 7 km h^{-1} in Utrecht and 20 ± 6 km h^{-1} in Hamburg.

As part of our driving strategy, we revisited locations where we had observed enhanced CH_4 readings (see SI, Sect. S.1.7, Figure S5). Not all recorded CH_4 mole fraction enhancements are necessarily the result of a stationary CH_4 source. For example, they could be related to emissions from vehicles which run on compressed natural gas, or vehicles operated with traditional fuels but with faulty catalytic converter systems. Later we will discuss how to exclude or categorize these unintended signals (see Sect. 2.2.2 and Sect. 2.3.1). *Therefore, we revisited a large number of locations (65 in Utrecht (≈80 %) and 100 in Hamburg (≈70 %)) where enhanced CH_4 had been observed in during the first survey in order to confirm the LIs.* In contrast to the measurements carried out in many cities in the United States (US) (von Fischer et al., 2017), our measurements were not carried out using Google Street View cars, but with a vehicle from the Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University (see SI, Sect. S.1.1, Figure S1). Due to time and budget restrictions, it was not possible to cover each street at least twice, as done for the US cities. After evaluation of the untargeted first surveys that covered each street at least once, targeted surveys were carried out for verification of observed LIs and for collection of air samples at locations with high CH_4 enhancements. The rationale behind this measurement strategy is that if an enhancement was not recorded during the first survey, it obviously cannot be verified in the second survey. The implications of the difference in the measurement strategy will be discussed in the Results and Discussion sections below.

In total, approximately 1,300 km of roads were driven during Utrecht surveys and about 2,500 km during the Hamburg campaign. In Utrecht, some re-visits were carried out several months to a year after the initial surveys in order to check on the persistence of the LIs. In Hamburg, revisits were also performed within the 4-week intensive measurement period. Further details about the driving logistics are provided in the SI (Sect. S.1.6, Table S1 and Table S2). It is possible that pipeline leaks that were detected during the initial survey were repaired before the revisit, and the chance of this occurring increases as the time interval between visits gets longer.

2.1.4 Air sample collection for attribution

In addition to the mobile measurement of C_2H_6 and CO_2 for LIs attributions purposes, samples for lab isotope analysis of $\delta^{13}C\text{-}CH_4$ and $\delta^{2}H\text{-}CH_4$ (hereinafter $\delta^{13}C$ and δD respectively) were collected during the revisits at locations that had

displayed high CH_4 enhancements during the first surveys. Depending on the accessibility and traffic, samples were either taken inside the car (see SI, Sect. S.1.8, Figure S6a) using a tubing from the bumper inlet, or outside the car on foot using the readings from the G4302 to find the best location within the plume (see SI, Sect. S.1.8, Figure S6b). *All the samples taken in the North Elbe study area and from most of the facilities were collected when the car was parked, but the samples inside the New Elbe tunnel and close to some facilities where there was no possibility to park were taken in motion while we were within the plume.* The sampling locations across the North Elbe study area of Hamburg were determined based the untargeted surveys, and the confirmation during revisits. The C_2H_6 information was not used in the selection of sampling locations in order to avoid biased sampling. Sampling locations from the facilities were determined based on wind direction, traffic, and types of different activities. Samples for isotope analysis were collected in non-transparent aluminum-coated Tedlar Supelco, Seupel™ Inert SCV Gas Sampling Bag (2 L) and SKC, Standard FlexFoil® Air Sample Bags (3 L) using a 12 V pump and 1/4-inch Teflon tubing which pumps air with flow rate of ≈0.25 L min^{-1}. In total, 103 bag samples were collected at 24 locations in Hamburg, 14 of them in the city area North of the Elbe river and 10 at larger facilities. Usually, three individual samples were collected at each source location, plus several background air samples on each sampling day. This sampling scheme generally results in a range of mole fractions that allow source identification using a Keeling plot analysis (Keeling, 1958, 1961). Fossil CH_4 sources in the study areas of this paper (inside the ring for Utrecht and north Elbe in Hamburg) refers to emissions originating from natural gas leaks.

2.1.5 Meteorological Data

Meteorological information reflecting the large scale wind conditions during the campaigns were obtained from measurements at the Cabauw tower (51.970263° N, 4.926267° E) operated by Koninklijk Nederlands Meteorologisch Instituut (KNMI) (Van Ulden and Wieringa, 1996) for Utrecht and Billwerder tower (53.5192° N, 10.1029° E) operated by the MI at Hamburg University (Brümmer et al., 2012) for Hamburg. The wind direction and wind speed data from the masts were used for planning the surveys. Pressure and temperature measurements were used to convert volume to mass fluxes for CH_4. We also used information from the towers for the GPDM calculations of the emission rates from larger facilities, because the local wind measurements from the 2-D anemometer were not logged continuously due to failure in logging setup of the measurements. In Utrecht, the Cabauw tower is located about 20 km from the WWTP. In Hamburg Billwerder tower is about 18 km from the Soil and Compost company and about 8 km from oil production facilities. Uncertainties over the wind data will be described later.

2.2 Emission quantification

2.2.1 Data preparation and background extraction of mobile measurements

The first step of the evaluation procedure is quality control of the data from both CH_4 analyzers and the GPS records. Periods of instrument malfunction and unintended signals based on notes written during each day's measurements were removed from the raw data. Extraction of the LIs from in-situ measurements requires estimation of the background levels (see SI, Sect. S.2.1, Figure S7). We estimated CH_4 background as the median value of ± 2.5 min of measurements around each individual point as suggested in Weller et al. (2019). For estimating the CO_2 background level we used the 5^{th} percentile of ± 2.5 min of measurements around each individual point (Brantley et al., 2014; Bukowiecki et al., 2002). The background determination method for CH_4 was selected from Weller et al. (2019) to follow the emission quantification algorithm for the urban studies, and while this algorithm doesn't include background extraction for CO_2, we chose commonly adopted method of background determination for this component. These background signals were subtracted from the measurement time series

to calculate the CH_4 and CO_2 enhancements. For C_2H_6, the background was considered zero as it is normally present at a very low mole fraction; between ~0.4-2.5 ppb (Helmig et al., 2016), and is lower than the G4302 detection limit.

2.2.2 Quantification of methane emissions from leak indications

We wrote an automated MATLAB® script (available on GitHub from Maazallahi et al. (2020a)) based on the approach initially introduced in von Fischer et al. (2017), and improved in Weller et al. (2019). *This algorithm was designed to quantify CH_4 emissions from ground-level emission release locations within 5-40 m from the measurement (von Fischer et al., 2017), such as pipeline leaks and has been demonstrated that the algorithm adequately estimates the majority of those emissions from a city (Weller et al., 2018).* Using the same algorithm also ensures that results are comparable between European and US cities. The individual steps will be described below. Mapping and spatial analysis were conducted using Google Earth and ESRI ArcMap software. A flow diagram of the evaluation procedure is provided in the SI (Sect. S.2.2, Figure S8).

Following the algorithm from von Fischer et al. (2017), measurements at speeds above 70 km h^{-1} were excluded, *as the data from the controlled release experiments (von Fischer et al., 2017) were not reliable at high speed (Weller et al., 2019).* We also excluded measurements during periods of zero speed (stationary vehicle) *to avoid unintended signals coming from other cars running on compressed natural gas when the measurement car was stopped in traffic.* In order to merge the sharp 1 Hz-frequency records of the GPS with the ≈ 0.3 Hz data from the G2301 analyzer, the CH_4 mole fractions were linearly interpolated to the GPS times.

Weller et al., (2019) established an empirical equation to convert LIs observed with a Picarro G2301 in a moving vehicle in urban environments into emission rates based on a large number of controlled release experiments in various environments (Eq. (1)).

$$\mathrm{Ln}(C) = -0.988 + 0.817 * \mathrm{Ln}(Q) \quad (1)$$

In this equation, C represents CH_4 enhancements above the background in ppm and Q is the emission rate in L min^{-1}. Weller et al., (2019) used controlled releases to demonstrate that the magnitude of the observed methane enhancement is related to the emission rate and carefully characterized the limitations and associated errors of this equation. We used Eq. (1) to convert CH_4 enhancements encountered during our measurements in Utrecht and Hamburg to emission rates, and we use these estimates to categorize LIs into three classes: high (emission rate > 40 L min^{-1}), medium (emission rate 6− 40 L min^{-1}) and low (emission rate 0.5 - 6 L min^{-1}), following the categories from von Fischer et al. (2017) (Table 1).

The spatial extent of individual LIs was estimated as the distance between the location where the CH_4 mole fraction exceeded the background by more than 10 % (≈ 0.200 ppm; as used in von Fischer et al. (2017) and Weller et al. (2019)) to the location where it fell below this threshold level again. LIs which stay above the threshold for more than 160 m were excluded in the automated evaluation because we suspect that such extended enhancements are most likely not related to leaks from the NGDN (von Fischer et al., 2017).

In a continuous measurement survey on a single day, consecutive CH_4 enhancements above background observed within 5 seconds were aggregated and the location of the emission source was estimated based on the weighted averaging of coordinates (Eq. (2)). Decimal degree coordinates were converted to Cartesian coordinates (see SI, Sect. S.2.3, Figure S9) relative to local references (see SI, Sect. S.2.3, Table S7). In Utrecht, the Cathedral tower (Domtoren) and in Hamburg the St. Nicholas' Church were selected as local geographic datums. LIs observed on different days at similar locations were clustered and interpreted as one point source when circles of 30 m radius around the centre locations overlapped, similar to Weller et al., (2019). The enhancement of the cluster was assigned the maximum observed mole fraction and located as the weighted average of the geographical coordinates of the LIs within that cluster (Eq. (2) from Weller et al. (2019)), where w_i is CH_4 enhancement of each LI.

$$(\text{lon, lat}) = \frac{\sum_{i=1}^{n} w_i * (lon_i, lat_i)}{\sum_{i=1}^{n} w_i} \quad (2)$$

We compared the outputs of our software to the one developed by Colorado State University (CSU) for the surveys in US cities (von Fischer et al., 2017; Weller et al., 2019). 30 LIs were detected and no significant differences were observed (linear fit equation y = 1.00 * x - 0.00, R^2 = 0.99) (see SI, Sect. S.2.4, Figure S10). As mentioned above, in our campaign-type studies not all streets were visited twice, so this criterion was dropped from the CSU algorithm. *Instead, we used explicit source attribution by co-emitted tracers.*

The emission rate per km of road covered during our measurements was then scaled up to the city scale using the ratio of total road length within the study area boundaries derived from OSM to the length of streets covered, and converted to a per-capita emission using the population in the study areas based on LandScan data (Bright et al., 2000). Note that in this up-scaling practice, emission quantified from facilities were excluded.

To account for the emission uncertainty, similar to Weller et al. (2018) for the US city studies, we used a bootstrap technique which was initially introduced in Efron (1979, 1982), as this technique is adequate in resampling of both parametric and non-parametric problems with even non-normal distribution of observed data. Tong et al. (2012) indicated that bootstrap resampling technique is sufficiently capable in estimating uncertainty of emissions with sample size of equal or larger than 9. Efron and Tibshirani (1993) suggested that minimum of 1,000 iterations are adequate in bootstrap technique. In this study, we used non-parametric bootstrap technique to account for the uncertainty of total CH_4 emissions from all LIs in each city with 30,000 replications. As mentioned above the algorithm is based on CH_4 enhancements of measurement with 5-40 m distance from controlled release location, and can produce large uncertainty for emission quantification of individual LI (Figure 4 in Weller et al. (2019)), but with sufficient number of sample size, the uncertainty associated with total emission quantified in an urban area is more precise.

2.2.3 Quantification of methane emissions from larger facilities

Apart from the natural gas distribution network, there are larger facilities in both cities that are potential CH_4 sources within the study area. Several facilities in or around the cities were visited during the mobile surveys to provide emission estimates. *We applied a standard point source GPDM (Turner, 1969) to quantify methane emissions from these larger facilities.* A flowchart describing the steps taken during quantification from facilities in given in SI (Sect. S.2.5., Figure S11). We note that emission quantification using GPDM with data from mobile measurements is prone to large errors (factor of 3 or more) (Yacovitch et al., 2018) especially when the measurements are carried out close to the source. In this study, we also report the data obtained from larger facilities, since rough emission estimates from facilities can be obtained in the city surveys. Caulton et al. (2018) discuss uncertainties of emission quantification with GPDM. Individual facilities were visited during the routine screening measurements and during revisits for LI confirmation and air sampling.

In Utrecht, the WWTP is located in the study area and streets around this facility were passed several times during surveys. In Hamburg, we initially performed screening measurements in the harbor area (extensive industrial activities) and near an oil production site and then revisited these sites for further quantification and isotopic characterization. The data from the oil production site can be fit reasonably well with a GPDM and were therefore selected for quantification, similar to studies in a shale gas production basin in the USA (Yacovitch et al., 2015) and in the Netherlands (Yacovitch et al., 2018).

$$C(x,y,z) = \frac{Q}{2*\pi*u*\sigma_y*\sigma_z} * \{ \exp\left(\frac{-(z-z_{source})^2}{2*\sigma_z^2}\right) + \exp\left(\frac{-(z+z_{source})^2}{2*\sigma_z^2}\right)\} * \exp\left(\frac{-y^2}{2*\sigma_y^2}\right)\} \quad (3)$$

In Eq. (3), C is the CH_4 enhancement converted to the unit of g/m^3 at cartesian coordinates x, y, and z relative to the source ($[x\ y\ z]_{source} = 0$), x is the distance of the plume from the source aligned with the wind direction, y is the horizontal axis perpendicular to the wind direction, z is the vertical axis. Q is emission rate in g s^{-1}, u (m s^{-1}) is the wind speed along the x-axis, and σ_y and σ_z are the horizontal and vertical plume dispersion parameters (described below), respectively.

Determination of an effective release location is a challenge for the larger facilities. Effective emission locations for each facility were estimated based on wind direction measurements and the locations of maximum CH_4 enhancements. The facilities were generally visited multiple times under different wind conditions. The locations of the maximum CH_4 enhancements were then projected against the ambient wind, and the intersection point of these projections during different wind conditions was defined as effective emission location of the facility. *At least two measurement transects with different wind direction were used to estimate the effective location of the source. If wind directions, road accessibility or the shape of plumes were not sufficient to indicate the effective source location, the geographical coordinates of centroids of the possible sources using Google Earth imageries and field observations were used to determine the effective emission location.* For the WWTP in Utrecht we also contacted the operator and asked for the location of sludge treatment as it is the major source of CH_4 emissions (Paredes et al., 2019; Schaum et al., 2015).

Neumann and Halbritter (1980) showed that the main parameters in sensitivity analysis of GPDM are the wind speed and source emission height in close distance and the influence of emission height become less further downwind compared to the mixing layer height. In this study, the heights of emission sources were low (<10m) and estimated during surveys and/or using Google Earth imageries, and considering that such a larger measurement distance from the facilities, the main sources of uncertainty of the emission estimates for the WWTP and Compost and Soil company are most likely the mean wind speed and for the upstream facilities in Hamburg the major sources of uncertainties can be the mean wind speed and emission height. We considered 0-4 m source height for the WWTP in Utrecht, and for the upstream facilities in Hamburg we considered 0-5 m emission height for the Compost and Soil site, 0-2 m for the separator, 0-10 m for the storage tank, and 0-1 m for the oil extraction well-head. We used 1 m interval for each of these height ranges to quantify emissions in GPDM.

Cross wind horizontal dispersions σ_y were estimated from the measured plumes by fitting a Gaussian curve to the *individual plumes* from each set during each day's survey. A set of plumes is defined as a back to back transects during a period of time downwind each facility on different days. Later average emissions from all sets of plumes were used to report CH_4 emission for each of the facilities. A suitable Pasquill–Gifford stability class was then determined by selecting a pair of parameters (Table 1-1 in EPA, 1995) that matches best and give the closest number to the with the fitted value of σ_y. Vertical dispersions σ_z were then estimated using the identified Pasquill–Gifford stability class in the first step, using the distances to the source locations (Table 1-2 in EPA, 1995). Uncertainties due to these estimates will be discussed below. Mass emission rates were calculated using the metric volume of CH_4 at 1 bar of atmospheric pressure (0.715 kg m^{-3} at 0 °C and 0.666 kg m^{-3} at 20 °C, P. 1.124 in IPCC, 1996), and linear interpolation was used for temperatures in between.

Due to technical issues, local wind data were not logged continuously and thus we used wind data from two towers *which are 8 to 20 km away from the facilities we focused for emission quantifications.* These distances introduce extra uncertainties in analyzing the emissions using GPDM mainly on the wind speed. By comparing some of the local high-quality wind data to data from the towers, we estimated that the local wind speed is within the *range of ± 30 % of the collected tower data.* This range was adopted to estimate the wind speed for emission quantifications for the set of plumes measured downwind of the facilities. The wind directions were aligned at local scale of each facility based on the locations of sources and locations of maxima of average CH_4 enhancements from a set of transects in each day's survey and we considered ± 5° uncertainty in wind direction for the GPDM quantification.

2.3 Emission attribution

2.3.1 Mobile C_2H_6 and CO_2 measurements

During the Utrecht campaign, the overall mole fraction of CH_4 and C_2H_6 in the NGDN was ≈ 80 % and ≈ 3.9 % (STEDIN, *personal communication*) and in Hamburg the mole fraction of CH_4 and C_2H_6 in the NGDN was about ≈ 95

% and ≈ 3.4 % (GasNetz Hamburg, *personal communication*) respectively. This ratio can vary depending on the mixture of gas compositions from different suppliers, but should meet the standards on the gas compositions in the Netherlands (65 – 96 mol-% for CH_4 and 0.2 – 11 mol-% for C_2H_6 (ACM, 2018)) and in Germany (83.64 – 96.96 mol-% for CH_4 and 1.06 – 6.93 mol-% for C_2H_6 (DVGW, 2013)). *Compressed natural gas vehicles can be mobile CH_4 emission sources (E. K. Nam et al., 2004; Curran et al., 2014; Naus et al., 2018; Popa et al., 2014) and in this study we also observed CH_4 signals from vehicles.* For example, the point to point C_2H_6:CH_4 ratio (C_2:C_1) calculated from road measurements of a car exhaust shown in Figure S12 (see SI, Sect. S.2.6) is 14.2 ± 7.1 %. During the campaigns in Utrecht and Hamburg the C_2:C_1 of NGDNs was less than 10 % and in our study, we removed all the locations where the C_2:C_1 ratio was greater than 10 %. CH_4 emissions from combustion processes are always accompanied by large emissions of CO_2 and can therefore be identified based on the low CH_4:CO_2 emission ratio. In this study, LIs with CH_4:CO_2 ratio between 0.02 and 20 with R^2 greater than 0.8 were attributed to combustion.

2.3.2 Lab isotopic analysis of $\delta^{13}C$ and δD

After sample collections, the bag samples were returned to the IMAU for analysis of both $\delta^{13}C$ and δD (Brass and Röckmann, 2010) and some samples were analyzed at the Greenhouse Gas Laboratory (GGL) in the department of Earth Sciences, Royal Holloway University of London (RHUL) for $\delta^{13}C$ (Fisher et al., 2006) (see SI, Sect. S.2.7, Figure S13).

At the IMAU, we used isotope ratio mass spectrometry (IRMS) instrument of ThermoFinnigan MAT DeltaPlus XL (Thermo Fisher Scientific Inc., Germany). We used a reference cylinder calibrated against Vienna Pee Dee Belmnite (V-PDB) for $\delta^{13}C$ and Vienna Standard Mean Ocean Water (V-SMOW) for δD at the at the Max Planck Institute for Biogeochemistry (MPI-BGC), Jena, Germany (Sperlich et al., 2016). The cylinder contained CH_4 mole fraction of 1975.5 ± 6.3 ppb, $\delta^{13}C$ = -48.14 ± 0.07 ‰ vs V-PDB and δD = -90.81 ± 2.7 ‰ vs V-SMOW. The samples were pumped through a magnesium perchlorate ($Mg(ClO_4)_2$) dryer before the CH_4 extraction steps. Each sample was measured at least 2 times (up to four times) for each isotope. Every other sample, the reference gas was also measured 3 times for $\delta^{13}C$ and δD. Each measurement, from the CH_4 extraction to the mass spectrometer, took ≈ 30 minutes.

At the GGL, Flex foil SKC bag samples were each analyzed for methane mole fractions and $\delta^{13}C$. Methane mole fractions were determined using a Picarro G1301 CRDS, which measured every 5 seconds for 2 minutes resulting in a precision ± 0.3 ppb (Lowry et al., 2020; France et al., 2016; Zazzeri et al., 2015). Each sample was then measured for stable isotopes ($\delta^{13}C$-CH_4) using an Elementar Trace gas and continuous-flow gas chromatography isotope ratio mass spectrometry (CF-GC-IRMS) system (Fisher et al., 2006), which has an average repeatability of ± 0.05 ‰. CH_4 extraction was preceded by drying process using $Mg(ClO_4)_2$. Each sample was measured 3 times for $\delta^{13}C$-CH_4, where the duration of each analysis was ≈ 20 minutes. Both instruments are calibrated weekly to the WMO X2004A methane scale using air filled cylinders that were measured by the National Oceanic and Atmospheric Administration (NOAA), and cylinders that were calibrated against the NOAA scale by the MPI-BGC (France et al., 2016; Lowry et al., 2020).

The analytical systems for isotope analysis have been described, used and/or compared in several previous publications (Fisher et al., 2011; Röckmann et al., 2016; Umezawa et al., 2018; Zazzeri et al., 2015). Measurement uncertainties in $\delta^{13}C$ and δD are 0.05-0.1 ‰ and 2-5 ‰ respectively.

After the LIs were analyzed and quantified, the measurements of C_2H_6, CO_2, and isotopic composition from the air samples were used for source attribution. We characterize the observed LIs as of fossil origin when they had a concomitant C_2H_6 signal between 1 % and 10 % of the CH_4 enhancements and when the isotopic composition was in the range -50 to -40 ‰ for $\delta^{13}C$ and -150 to -200 ‰ for δD. A LI was characterized as microbial when there was no C_2H_6 signal (<1 % of the CH_4 enhancements larger than 500 ppb), $\delta^{13}C$ was between -55 ‰ and -70 ‰ and δD was between -260 and -360 ‰ (Figure 7 in Röckmann et al., 2016). LIs with enhancements of CH_4 lower than 500 ppb and no C_2H_6 signals were categorized as

unclassified. LIs with no C_2H_6 signals, no significant CH_4:CO_2 ratio, and no information on $\delta^{13}C$ and δD were also categorized as unclassified. The source signatures for each sampling location were determined by a Keeling plot analysis of the three samples collected in the plumes and a background sample taken on the same day.

3 Results

3.1 Quantification of CH_4 emissions across Utrecht and Hamburg

Table 2 summarizes the main results from the surveys in Hamburg and Utrecht. The amount of km of roads covered in Hamburg is roughly a factor of 2 larger than in Utrecht, and also the number of detected LIs is roughly a factor of 2 larger, for all three categories. This shows that the overall density of LIs (km covered per LI) in both cities is not very different. Specifically, a LI is observed every 5.6 km in Utrecht and every 8.4 km in Hamburg. While not all streets were visited twice in both cities (see SI, Sect. S.1.5, Table S4) 80 % of LIs in Utrecht and 69 % of LIs in Hamburg were revisited which account for 91 % and 86 % of emissions respectively in the study areas. During revisits, 60 % of CH_4 emissions in Utrecht and 46 % of emissions in Hamburg were confirmed. In both cities, all LIs in the high emission category were re-observed. In some cases, re-visits were carried out several months after first detection, and the LIs were still confirmed (e.g. see SI, Sect. S.1.7, Figure S5).

The distribution of CH_4 LIs across the cities of Utrecht and Hamburg is shown in Figure 2. As shown in Table 2, a total of 145 significant LIs were detected in Hamburg and 81 in Utrecht; these LIs cover all three LI categories. Two LIs in Hamburg and one LI in Utrecht fall in the high (red) emission category; the highest LI detected in Utrecht and Hamburg corresponded to emission rates of ≈ 100 L min^{-1} and ≈ 70 L min^{-1}, respectively. *Noted that estimates for individual leaks with the Weller et al. (2019) algorithm can have large error, thus these results are indicative of large leaks, but the precise emission strength is very uncertain.* Six LIs in Utrecht and 16 LIs in Hamburg fall in the middle (orange) emission category, and 127 LIs in Hamburg and 74 LIs in Utrecht fall in the low (yellow) emission category. The distribution of emissions over the three categories is also similar between the two cities, with roughly one third of the emissions originating from each category (Figure 2), but the number of LIs in each category is different. The contribution of LIs in the high emission category is about a third of the total observed emissions (35 % in Utrecht is (1 LI) and in 30 % in Hamburg (2 LIs)).

CH_4 emitting locations were categorized based on the roads where the LIs were observed (Figure 1, Figure 2, Figure 3, and Table S8 in SI, Sect. S.3.1). Average emission rates per LI as derived from equation (1) are similar for the two cities with 3.6 L min^{-1} LI^{-1} in Utrecht and 3.4 L min^{-1} LI^{-1} in Hamburg, but they are distributed differently across the road (Figure 1). In Utrecht, emitting locations on level 2 roads contributed the most (50 % of emissions) to the total emissions while in Hamburg the majority of the emissions occurred on residential roads (56 % of total emissions). This shows that the major leak indications may happen on different road classes in different cities and there is no general relation to the size of streets between these two cities.

In Figure 4, we compare cumulative CH_4 emissions for Utrecht and Hamburg to numerous US cities (Weller et al., 2019). After ranking the LIs from largest to smallest, it becomes evident that the largest 5 % of the LIs account for about 60 % of emissions in Utrecht, and 50 % of the emissions in Hamburg.

As mentioned above, the observed total emission rates observed on roads in urban environment in the two cities are relatively similar when normalized by the total amount of km covered, 0.64 L min^{-1} km^{-1} for Utrecht and 0.4 L min^{-1} km^{-1} for Hamburg (Table 2). Using these two emission factors, the observed emission rates (≈110 t yr^{-1} in Utrecht and ≈180 t yr^{-1} in Hamburg) were up-scaled to the entire road network in the two cities, ≈ 650 km in Utrecht and ≈ 3,000 km in Hamburg. This includes the implicit assumption that the pipeline network is similar to the street network. Total up-scaled emission rates based on mobile measurements on roads in urban environment before considering attribution analysis over LI locations are 150 t yr^{-}

[1] and 440 t yr^{-1} across the study areas of Utrecht and Hamburg respectively. Distributing the calculated emission rates over the population in the city areas yields emission rates of 0.54 ± 0.15 kg yr^{-1} $capita^{-1}$ for Utrecht and 0.31 ± 0.04 kg yr^{-1} $capita^{-1}$ for Hamburg (see SI, Sect. S.3.2, Figure S14).

3.2 Attribution of CH_4 emissions across Utrecht and Hamburg

Figure 5 shows the results of the isotope analysis for the 21 locations in Hamburg where acceptable Keeling plots were obtained (see SI, Sect. S.3.3, Table S9 and Table S10). The results cluster mostly in three groups, which are characterized by the expected isotope signatures for fossil, microbial, and pyrogenic samples as described in Röckmann et al., (2016).

Average isotope signatures for the LIs in the city of Hamburg were $\delta^{13}C$ = -52.3 ± 5.1 ‰ and δD = -298.4 ± 30.3 ‰ for the samples characterized as microbial and $\delta^{13}C$ = -41.9 ± 1.0 ‰ and δD = -196.1 ± 10.6 ‰ for the samples characterized as fossil (Figure 5). One sample from the Hamburg city area displays a very high source signature of $\delta^{13}C$ = -23 ‰ and δD = -153 ‰. The origin of CH_4 with such an unusual isotopic signature could not be identified and it is considered an outlier. In Hamburg, 10 % of the LI locations (38 % of emissions) on the north side of Elbe were sampled for isotope analysis. The lab isotopic attributions show that the LIs with the higher emission rates are mostly caused by emission of fossil CH_4. 79 % of the inferred emissions at 38 % of the LIs were identified as of fossil origin, 20 % of emissions at 54 % of the LIs as of microbial origin (for an identified source see SI, Sect. S.3.3, Figure S15), 1 % of emissions at 8 % of LIs as of pyrogenic origin.

In Hamburg, during three passes through the new Elbe tunnel (see SI, Sect. S.3.4, Figure S16) a CH_4:CO_2 of 0.2 ± 0.1 ppb:ppm was derived for combustion-related emission. During the surveys of open roads, clear CH_4:CO_2 correlations were observed for several LIs and an example of a measurement of car exhaust is shown in Figure S12a (see SI, Sect. S.2.6) with CH_4:CO_2 = 1.6 ppb:ppm. Previous studies have shown relatively low CH_4:CO_2 ratios of $4.6*10^{-2}$ ppb:ppm (Popa et al., 2014), 0.41 ppb ppm^{-1} (E. K. Nam et al., 2004), and 0.3 ppb:ppm (Naus et al., 2018) when cars work under normal conditions. During cold engine (Naus et al., 2018) or incomplete combustion conditions, the fuel to air ratio is too high, which results in enhanced emission of black carbon particles and reduced carbon compounds, so higher CH_4:CO_2 ratios. Hu et al. (2018) reported 2 ± 2.1 ppb:ppm in a tunnel, but 12 ± 5.3 ppb:ppm [1] on roads. In addition to car exhaust, there are other combustion sources which can affect CH_4 and CO_2 mole fractions at the street level including natural gas water heater (CH_4:CO_2 ratio of ≈ 2 ppb:ppm; Lebel et al., 2020), restaurant kitchens, etc. Based on the CH_4:CO_2 ratio (ppb:ppm) criterion defined above (see Sect. 2.3.1), 17 % of LIs (10 % of emissions) can be attributed to combustion (see SI, Sect. S.3.4, Figure S17) with a mean CH_4:CO_2 ratio of 3.2 ± 3.9 ppb:ppm (max = 18.7 and min = 0.8 ppb:ppm). The C_2:C_1 ratio for these LIs attributed to combustion in Hamburg was 7.8 ± 3.5 %. In Utrecht 7 % of LIs (2 % of emissions) are attributed to combustion with a mean CH_4:CO_2 ratio of 9.8 ± 5.8 ppb:ppm (max = 16.7 and min = 3.0 ppb:ppm).

Based on the C_2H_6 signals, 64 % of the emissions (33 % of LIs) were characterized as fossil, while 25 % of emissions (20 % of LIs) were identified as microbial. Due to low CH_4 and C_2H_6 enhancements, 47 % of the locations (11 % of emission) were considered unclassified. The C2:C1 ratio for the LIs attributed to emissions from NGDNs in Hamburg study area (North Elbe) is 4.1 ± 2.0 %. The oil production site in south-east Hamburg had a higher C_2:C_1 ratio of 7.1 ± 1.5 %.

In Utrecht, C_2H_6 was measured only during four surveys in February, April, and June 2019 (revisits of 2-day surveys across the city center and 2 days to LIs with high emission rates) as the CH_4 - C_2H_6 analyzer was not available during the first campaign. The C_2:C_1 ratios from this limited survey indicates that 93 % of emissions (69 % of the LIs across the city centre, including combustions) are likely from fossil sources (Table 2) and 73 % of emissions (43 % of the LIs, including combustion) out of all LIs. In Utrecht, the C_2:C_1 ratio for the LIs attributed to NGDNs is 3.9 ± 0.8 %.

3.3 Quantification of CH_4 plume from larger facilities

Table 3 shows the emission rate estimates from the larger facilities in Utrecht and Hamburg. CH_4 plumes from the WWTP (Figure 6 and in SI, Sect. S.1.6., Table S5) were intercepted numerous times during the city transects, and the error estimate in Table 3 represents one standard deviation of 5 sets of measurements where each measurement comprises 2-4 transects during three measurement days (12-Feb.-2018, 24-Apr.2018, and 07-Jan.-2019). Figure 7 shows an example of a fit of a Gaussian plume to the measurements from the Utrecht WWTP. The derived distance to the source was 215 ± 90 m, the hourly average wind speed was 3.5 ± 1.1 m s^{-1} and the wind direction was 178 ± 5 degrees (see SI, Sect. S.1.6, Table S5).

The total emission rate of the WWTP in Utrecht was estimated at 160 ± 90 t yr^{-1}. The reported errors include stability classes, wind speed and directions, and effective point source coordinates. Not all transects provided datasets that allowed an adequate Gaussian fit, these were not included in total estimates from the facilities, e.g. measurements during the visits of the harbor area in Hamburg were excluded. In Hamburg, plumes from several facilities were also intercepted several times (see SI, Sect. S.1.6, Table S6). For a Compost and Soil Company in Hamburg we estimate an emission rate of 70 ± 50 t yr^{-1}. The mobile quantifications at the upstream sites in Hamburg from a separator, a tank, and an oil well yield annual CH_4 emission of 4.5 ± 3.7 t yr^{-1}, 5.2 ± 3.0 t yr^{-1}, and 4.8 ± 4.0 t yr^{-1} respectively.

4 Discussion

4.1 Detection and quantification

As mentioned above (see Sect. 2.2.2), we used methods similar to the ones introduced by von Fischer et al. (2017) and updated in Weller et al. (2019) that were used to characterize CH_4 emission from local gas distribution systems in the US. An important difference is that we did not visit each street twice in the untargeted survey, and the revisits were specifically targeted at locations where we had found a LI during the first visit. A consequence of the different sampling strategy is that we do not base our city-level extrapolated emissions estimates on "confirmed" LIs, as done in Weller et al. (2019) but on all the LIs observed. In our study, 60 % of CH_4 LIs in Utrecht and 46 % of LIs in Hamburg were confirmed. This number may be biased high, since we preferentially revisited locations that had shown higher LIs, and the percentage of confirmed LIs may have been lower if we had visited locations with smaller LIs. Von Fischer et al. (2017) reported that LIs in the high emission rate category have a 74 % chance of detection, which decreased to 63 % for the middle category and 35 % frequency for the small category. In our study, all LIs within the high emission rate category (n = 1 and n = 2 LIs in Utrecht and Hamburg respectively) were confirmed in both cities. Overall, the confirmation rates found in Hamburg and Utrecht were similar to the ones reported in the US cities by von Fischer et al. (2017), suggesting that the results from both driving strategies can be compared when we take into account an overall confirmation percentage of roughly 50 %.

In 13 US cities the "LI density" ranged from 1 LI per 1.6 km driven to 1 LI per ≈ 320 km driven (EDF, 2019). This illustrates that cities within one country can be very different in their NGDN infrastructure. In Utrecht, one LI was observed every 5.6 km of street covered and in Hamburg every 8.4 km covered. Note that we normalize the number of LIs per km of road covered, not km of road driven, since the revisits were targeted to confirm LIs, which would bias the statistics if we normalize by km of road driven. After accounting for the confirmation percentage of 50 %, the LI densities in Utrecht and Hamburg become 1 LI per 11.2 km covered in Utrecht, and 1 LI per 16.8 km covered in Hamburg. When we take into account the attributions (fraction fossil/total LIs is 43 % in Utrecht and 31 % in Hamburg), confirmed LIs from the NGDN are found every 26 km in Utrecht and every 54 km in Hamburg. The highest 1 % of the LIs in Utrecht and Hamburg account for approximately 30 % of emissions, emphasizing the presence of a skewed distribution of emissions. The emissions distribution is even more skewed for these two European cities than for countrywide US cities, where approximately 25 % of emissions

comes from the highest 5 % of the LIs. Skewed emission distributions appear to be typical for emissions from the oil and gas supply chain across different scales. For example, a synthesis study reviewing the distribution of upstream emissions from the US natural gas system shows that in the US 5 % of the leaks are responsible for 50 % of the emissions (Brandt et al., 2016).

4.2 Attribution

Four different approaches were combined in Hamburg for emission source attribution, which allows an evaluation of their molecular consistency. Figure 5 shows that measurements of the C_2:C_1, δD, and $\delta^{13}C$ provide a very consistent distinction between fossil and microbial sources of CH_4. Except for one outlier with a very enriched $\delta^{13}C$ and δD contents and no C_2H_6 signal, all samples that are classified as "microbial" and depleted in $\delta^{13}C$ and δD signatures contain no measurable C_2H_6. Samples that are characterized as "fossil", based on $\delta^{13}C$ and δD signatures, bear a C_2H_6 concomitant signal. This strengthens the confidence in source attribution using these tracers. The fossil $\delta^{13}C$ signature of bag samples from natural gas leaks in Hamburg ($\delta^{13}C$ = -41.9 ± 1.0 ‰) is higher than recent reports from the city of Heidelberg, Germany ($\delta^{13}C$ = −43.3 ± 0.8 ‰ (Hoheisel et al., 2019)). This shows that within one country, $\delta^{13}C$ from NGDNs can vary from one region to another. *These numbers do not agree within combined errors, but are also not very different. $\delta^{13}C$ values of CH_4 from the NGDN can vary regionally and temporally, e.g. due to differences in the mixture of natural gas from various suppliers for different regions in Germany (DVGW, 2013). In a comprehensive study at global scale, it is also shown that how $\delta^{13}C$ values of fossil fuel CH_4 have significant variabilities in different regions within an individual basin (Figure 4 in Sherwood et al. (2017)).*

In Hamburg both C_2:C_1 and CH_4:CO_2 analysis along with $\delta^{13}C$ and δD signatures suggest that ≈ 50 % to ≈ 80 % of estimated emissions (≈ 30 % and ≈ 40 % of LIs respectively) originate from NGDNs, whereas CH_4:CO_2 analysis and the smaller sample of C_2:C_1 measurements in Utrecht suggests that the overwhelming fraction (70 - 90 % of emissions; 40 – 70 % of LIs) originated from NGDNs. We note that although it is widely assumed that microbial CH_4 is not associated with ethane, some studies have reported microbial production of ethane, so it may not be a unique identifier (Davis and Squires, 1954; Fukuda et al., 1984; Gollakota and Jayalakshmi, 1983; Formolo, 2010). The online C_2:C_1 analysis to attribute LIs is fast and can be used at larger scale, but with the instrument we used we were not able to clearly attribute sources with CH_4 enhancements of less than 500 ppb. Isotopic analysis by IRMS can attribute sources for smaller LIs (down to 100-200 ppb) but is clearly more labor intensive, and it would be a considerable effort to take samples from all LIs observed across an urban area. *Overall, C_2H_6 and CO_2 signals are very useful in eliminating non-fossil LIs in mobile urban measurements and with improvements in instrumentations, analyzing signals of these two species along with evaluation of CH_4 signals can make process of detecting pipeline leaks from NGDN more efficient.*

In Hamburg, most of the LIs were detected in the city center (Figure 1). This means that the LI density is higher than the average value in the center, but much lower than the average value in the surrounding districts and residential areas. Many of the LIs in the city center were attributed to combustion and microbial sources, thus they do not originate from leaks in the NGDN. *Many of the microbial LIs encountered in Hamburg are around the Binnenalster lake (see SI, Sect. S.3.3, Figure S15), which suggests that anaerobic methanogenesis (Stephenson and Stickland, 1933; Thauer, 1998) can cause these microbial emission in this lake, as seen in other studies focused on emissions from other lakes (e.g., DelSontro et al., 2018; Townsend-Small et al., 2016). Microbial CH_4 emissions from sewage system (Guisasola et al., 2008) can also be an important source of in this area, as seen in US urban cities (Fries et al., 2018). Fries et al. (2018) performed direct measurement of CH_4 and nitrous oxide (N_2O) from a total of 104 sites, and analyzed $\delta^{13}C$ and δD signatures of samples from 27 of these locations, and attributed 47 % of these locations to microbial emissions in Cincinnati, Ohio, USA.*

4.3 Comparison to national inventory reports

In the national inventory reports, total upscaled emissions from NGDNs are based on sets of emission factors for different pipeline materials (e.g., grey cast iron, steel, or plastic) at different pressures (e.g., <= 200 mbar or >200 mbar). The reported emission factors are based on IPCC tier 3 approach (Buendia et al., 2019). However, emission estimates do not exist for individual cities including Utrecht and Hamburg. Also, it is not possible to calculate a robust city-level estimate using the nationally reported emission factors because there is no publicly available associated activity data, i.e., pipeline materials and lengths for each material, at the level of individual cities. As a result, a robust direct comparison between nationally reported emissions and our measurements, akin to a recent study in the United States (Weller et al., 2020), is currently not possible. The following juxtaposition of our estimates and national inventory downscaling to city-level is therefore provided primarily as illustration of the data gaps rather than a scientific comparison. In Utrecht, we attributed 70 – 90 % of the mobile measurement inferred emissions of ≈ 150 t yr^{-1} to the NGDN, thus 105 – 135 t yr^{-1}.

The Netherlands National Institute for Public Health and the Environment (RIVM) inventory report derived an average NGDN emission factor of ≈ 110 kg km^{-1} yr^{-1} using 65 leak measurements from different pipeline materials and pressures in 2013. *This weighted average ranged from a maximum of 230 kg km^{-1} yr^{-1} for grey cast iron pipelines to a minimum of 40 kg km^{-1} yr^{-1} for pipelines of other materials with overpressures <= 200 mbar (for details, see P. 130 in Peek et al. (2019)).* This results in an average CH_4 emissions of ≈ 70 t yr^{-1} (min = 30 t yr^{-1} and max = 150 t yr^{-1}) for the study area of Utrecht, assuming ≈ 650 km of pipelines inside the ring, and further assuming that Utrecht's NGDN is representative of the national reported average (see qualifiers above). The average emissions for the Utrecht study, based on emissions factors reported for the Netherlands, is smaller by a factor of 1.5 - 2 compared to the emissions derived here. The variability factor of 5, from the reported emission (resulting from the variability in pipeline materials) highlights the need for city-level specific activity data for a robust comparison. In Hamburg, 50 – 80 % of the upscaled emissions of 440 t yr^{-1} (220 – 350 t yr^{-1}), can be attributed to the emission from NGDN. The national inventory from the Federal Environment Agency (UBA) in Germany, reports an average CH_4 emission factor for NGDN from low pressure pipelines as ≈ 290 kg km^{-1} yr^{-1} (max = 445 kg km^{-1} yr^{-1} (grey cast iron) and min = 51 kg km^{-1} yr^{-1} (plastic)) based on measurements from the 1990s (Table 169 in Federal Environment Agency (2019)). Assuming ≈ 3000 km of pipelines in the targeted region, and further assuming that Hamburg's NGDN is representative of the national reported average (see qualifiers above), results in an estimated NGDN CH_4 emissions average of ≈ 870 t yr^{-1} (min = 155 t yr^{-1} and max = 1350 t yr^{-1}). While this study's estimate (220 – 350 t yr^{-1}) falls in the lower end of this range, the reported emissions variability factor of 9 (resulting from the variability in pipeline materials) highlights again the need for city-level specific activity data for a robust comparison. To put the national inventory comparison into perspective, it should be noted that GasNetz Hamburg detected and fixed leaks at 20 % of the fossil LIs in this study, which accounted for 50 % of emissions. In Utrecht and Hamburg, the natural gas consumption in our target area were retrieved through communications with LDCs. In the Utrecht and Hamburg study areas, natural gas consumption is 0.16 bcm yr^{-1} (STEDIN, personal communication) and 0.75 bcm yr^{-1} (GasNetz Hamburg, personal communication) respectively. The estimated emissions from NGDNs in our study is between 0.10 – 0.12 % in Utrecht and between 0.04 – 0.07 % in Hamburg of total the annual natural gas consumptions in the same area. In the US, where the majority of natural gas consumption is from residential and commercial sectors, Weller et al. (2020) reported emissions of 0.69 Tg $year^{-1}$ (0.25 - 1.23 with 95 % confidence interval), with a sum of ≈ 170 Tg $year^{-1}$ (U.S. EIA, 2019), showing 0.4 % (0.15 % - 0.7 %) loss from NGDNs. The US NGDNs loss is about four times larger than our reported loss in Utrecht, and is about ten times larger than the loss for Hamburg. Considering the population of Utrecht (≈ 0.28 million) and Hamburg (≈ 1.45 million), the natural gas consumption densities in these study areas are ≈ 570 m^3 $capita^{-1}$ yr^{-1} and ≈ 520 m^3 $capita^{-1}$ yr^{-1}, where in the US (population ≈ 330 million (US Census Bureau, 2020)) the density is about ≈ 730 m^3 $capita^{-1}$ yr^{-1} (see SI, Sect. S.3.2, Figure S14). This shows that annual natural gas

consumption per capita in the US is about 30 % and 40 % higher than in Utrecht and Hamburg respectively. The emission per km of pipeline in Utrecht is between 0.45 – 0.5 L min^{-1} km^{-1} and in Hamburg is between 0.2 – 0.32 L min^{-1} km^{-1}. In the US, based on 2,086,000 km km of local NGDN pipeline (Weller et al., 2020), this emission factor will be between 0.32 – 1.57 L min^{-1} km^{-1}. This shows higher emissions per km pipeline in the countrywide studies of US compared to just two European cities of Utrecht and Hamburg (see qualifiers above). This can be partly explained by pipeline material, maintenance protocols, and higher use of natural gas consumption in the US. However, the substantial variability in emission rates across US cities, as wells as the annual variability of gas consumption over the year, again restricts a direct comparison of two cities with a national average measured over multiple years.

Normalized LIs emissions per capita in Utrecht (0.54 ± 0.15 kg yr^{-1} $capita^{-1}$) are almost double the emission factor in Hamburg (0.31 ± 0.04 kg yr^{-1} $capita^{-1}$). This metric may be useful to compare cities, assuming that the emission quantification method is equally effective for different cities. CH_4 emissions can vary among different cities, depending on the age, management and material of NGDNs, and/or the management of local sewer systems. In our study, we only surveyed two cities, and the above number may not be adequate for extrapolation to the country scale (McKain et al., 2015).

4.4 Interaction with utilities

After the city surveys, locations with the highest emissions (high and medium categories) were shared with STEDIN Utrecht and all LI locations were reported to GasNetz Hamburg. The utilities repair teams were sent to check whether LIs could be detected as leaks from NGDN and fixed. The LDCs follow leak detection procedures based on country regulations (e.g., for GasNetz Hamburg in SI, Sect. S.4.1, Table S11). GasNetz Hamburg also co-located the coordinates of the detected reported LIs with the NGDN and prioritized repairs based on safety regulations mentioned in Table S12 (see SI, Sect. S.4.1). This interaction with the LDCs resulted in fixing major NGDN leaks in both cities. In Utrecht the only spot in the high emission category was reported to STEDIN, but the pipelines on this street had been replaced, which most likely fixed the leak, as it was not found later by the gas company nor in our later survey with the CH_4 - C_2H_6 analyzer. In Utrecht, half of the LIs in the medium category were found and repaired.

A routine leak survey (detection and repair) had been performed by GasNetz Hamburg between 1-5 months before the campaign, for the different regions (see SI, Sect. S.4.1., Table S11). The timing of any routine detection and repair likely influences the absolute number of LIs measured during independent mobile measurements, and the survey by GasNetz Hamburg thus likely has influenced the absolute number of LIs measured in our campaign. We then reported the LI latitude/longitude coordinates to GasNetz Hamburg about 4 months after our campaign. Additionally, we provided map images of the LIs immediately after the campaign. The comparison of the number of reported LIs (and emission rates) during our campaign with those identified by GasNetz Hamburg post-campaign assumes that the leaks continued to emit gas until they were detected and fixed by GasNetz Hamburg (if they were detected).

Depending on how close the gas leaks are located to a building, the LDCs prioritize the leaks into four classes from the highest to lowest priority: A1, A2, B, and C (see SI, Sect. S.4.1, Table S12). In Hamburg, both LIs in the high category were identified as A1 gas leaks and fixed by GasNetz Hamburg immediately. Most of the Hamburg LIs that were detected and identified as fossil are in close proximity to the natural gas distribution pipelines (see SI, Sect. S.4.2, Table S13). Investigation of the pipeline material shows that most of NGDN emissions are due to leaks from steel pipelines (see SI, Sect. S.4.2, Table S14), which are more prone to leakage because of pipeline corrosion (Zhao et al., 2018). Nevertheless, only 7 of the 30 LIs (23 %) that were positively attributed to fossil CH_4 were detected and fixed by the LDC. If we assume that the fraction fossil / total LIs determined in Hamburg (≈ 35 %) is representative for the entire population of LIs encountered (thus also for the ones that were not attributable), about 50 of the 145 LIs are likely due to fossil CH_4. The LDC found and fixed leaks at 10 of these locations (≈ 20 %). A recent revisit (January 2020) to these locations confirmed that no LIs were detected at 9 out of

these 10 locations. For the 10th location a smaller LI was detected in close proximity, and GasNetz Hamburg confirmed that this was a leak from a steel pipeline. The whole pipeline system on this street dates back to the 1930s and is targeted for replacement in the near future.

In summary, about 20 % of the LIs including the two largest LIs that were attributed to a fossil source were identified as NGDN gas leaks (see SI, Sect. S.4.2, Figure S18), and were repaired by GasNetz Hamburg, but these accounted for about 50 % of fossil CH_4 emissions of Hamburg, similar to what was observed in the US studies (Weller et al., 2018). Possibly, smaller leakages that can be detected with the high sensitivity instruments used in the mobile surveys cannot be detected with the less sensitive equipment of LDCs. Another possible explanation for the fact that the LDC did not detect more leaks may be that reported LI locations do not always coincide with the actual leak locations, although Weller et al. (2018) reported that the median distance of actual leak locations to the reported ones was 19 m. Combined measurements with GasNetz Hamburg are planned to investigate why the majority of the smaller LIs reported in mobile surveys is not detected in the regular surveys of the LDC.

The average C_2:C_1 ratio for LIs with a significant C_2H_6 signals across Hamburg was 5.6 ± 3.9 %. For the spots where the LDC found and fixed leaks this ratio was 3.9 ± 2.6 %. Thus, some of the locations where CH_4 enhancements were found were influenced by sources with an even higher C_2:C_1 ratio than the gas in the NGDN. One confirmed example is the very high ratio found in exhaust from a vehicle as shown in Figure S12 (see SI, Sect. S.2.6). The abnormal operation of this vehicle is confirmed by the very high CH_4:CO_2 ratio of 5.5 ppb:ppm (SI, section S2). This is more than 20 times higher than CH_4:CO_2 ratios of 0.2 ± 0.1 ppb:ppm observed during passages through the Elbe tunnel, a ratio that agrees with previous studies (SI, section S2).

Repairing gas leaks in a city has several benefits for safety (preventing explosions), sustainability (minimizing GHG emissions) and economics. Gas that is not lost via leaks can be sold for profit, but gas leak detection and repair is expensive and is usually associated with interruptions of the infrastructure (breaking up pavements and roads). Also, as reported above, and in agreement with the studies in US cities, for small LIs the underlying leaks are often not found by the LDCs, possibly because their equipment is less sensitive and aimed for finding leak rates that are potentially dangerous.

Our measurements in Hamburg demonstrate that in particular smaller LIs may originate from biogenic sources, e.g. the sewage system, and not necessarily from leaks in the NGDN. In this respect, attribution of LIs prior to reporting to the LDCs may be beneficial to facilitate effective repair. *Figure S19 (see SI, Sect. S.5) illustrates how the individual measurement components can be efficiently combined in a city leak survey program.*

4.5 Large facilities

The WWTP in Utrecht emits 160 ± 90 t yr^{-1}, which is similar to the total detected emissions (150 t yr^{-1}) inside the study area of Utrecht. The emissions reported for this facility from 2010 until 2017 are 130 ± 50 t yr^{-1} (Rijksoverheid, 2019), in good agreement with our measurements. CH_4 emission from a single well in Hamburg was estimated at 4.4 ± 3.5 t yr^{-1}, which is in the range of median emissions of 2.3 t yr^{-1} reported for gas production wells in Groningen, NL (Yacovitch et al., 2018), and average emissions of all US oil and gas production wells 7.9 ± 1.8 t yr^{-1} (Alvarez et al., 2018). In Hamburg, the emissions from a Compost and Soil Company amount to about 10 % of the total emissions in the city target region, whereas a wellhead, a storage tank and a waste-oil separator contribute only about 1 % each. This shows that individual facilities can contribute significantly to the total emissions of a city. The contribution of each source is dependent on infrastructure, urban planning and other conditions in the city (e.g. age and material of pipeline, maintenance programs, waste management, sewer system conditions, etc.), which may change the source mix from one city to another. For example, in Utrecht the WWTP is located within our domain of study. The wastewater treatment in Hamburg most likely causes CH_4 emissions elsewhere.

Therefore, facility-scale CH_4 emissions should be reported on a more aggregated provincial or national level. For emissions from the NGDN, the urban scale is highly relevant, as the emission can only be mitigated at this scale.

5 Conclusions

Mobile measurements provide a fast and accurate technique for observing and identifying even relatively small CH_4 enhancements (i.e., tens of ppb) across cities and are useful for detecting potential gas leaks. During our intensive measurement campaigns, 81 LIs were observed in Utrecht (corresponding to emissions of ≈110 t CH_4 yr^{-1}) and 145 LIs (≈180 t CH_4 yr^{-1}) in Hamburg. These estimates, based on the streets covered, were then up-scaled to the total study area, *using the road network map as a proxy for the length of the pipeline network* which then yielded total emissions of 150 t yr^{-1} and 440 t yr^{-1} across the study area of Utrecht and Hamburg respectively. The isotopic signature of CH_4 in air samples and continuous mobile measurement of CO_2 and C_2H_6 mole fraction show that not all the LIs observed across the two cities have fossil origin. In Utrecht, C_2:C_1 and CH_4:CO_2 analyses show that 70 -90 % of emissions were fossil. In Hamburg, C_2:C_1, CH_4:CO_2, and $\delta^{13}C$-δD analyses suggests that 50 - 80 % of emissions originate from natural gas pipelines. For the locations where samples for isotope analysis were collected, 80 % of emissions were identified as fossil. A large fraction of emissions in both cities originated from few high emitting locations. The LDC in Hamburg (GasNetz Hamburg) detected and fixed leaks at 20 % of the locations that likely due to fossil sources, but these accounted for 50 % of emissions. Large LIs were generally confirmed as gas leaks from steel pipelines. The C2:C1 ratio at the locations where gas leaks were fixed by GasNetz Hamburg was 3.9 ± 2.6 %. The mobile measurement technique is less labor and time intensive than conventional methods and can provide extensive coverage across a city in a short period. Based on our experience for the Netherlands and Germany a protocol could be developed that aids LDCs in guiding their leak detection and repair teams. The use of emission categories and source attribution can help target repair activities to the locations of large fossil emissions. Emission quantification from large facilities shows that these emissions may be equivalent to total CH_4 emissions from NGDN leaks in urban environments. In order to analyze discrepancies between spatial explicit measurement-based estimates as presented here with reported annual average national emissions by sectors a coordinated effort with national agencies is necessary to address the lack of publicly available activity data (e.g., pipe material) disaggregated from the national-level (e.g., at the city-level).

Author contributions

H. M. performed the mobile measurements, wrote the MATLAB® code, analyzed the data, and together with T. R. drafted the manuscript. J. M. F. and M. M. contributed with air sampling and isotope analysis. D. Z. -A. and S. S. contributed to the scientific interpretation and comparison between European and US cities. Z. D. W. and J. C. v. F. facilitated comparison to US cities and contributed to the statistical analysis. H. D. v. d. G. and T. R. provided instruments, equipment, and supervised the measurements and data analysis. T. R. developed the research idea and coordinated the city campaigns. All authors contributed to the interpretation of the results and the improvement of the manuscript.

Competing interests: The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the Climate and Clean Air Coalition (CCAC) Oil and Gas Methane Science Studies (MMS) hosted by the United Nations Environment Programme. Funding was provided by the Environmental Defense Fund, Oil and Gas Climate Initiative, European Commission, and CCAC. This project received further support from the H2020 Marie Skłodowska-Curie project Methane goes Mobile – Measurements and Modelling ($MEMO^2$; https://h2020-memo2.eu/), grant number 722479. Dr. Daniel Zavala-Araiza and Dr. Stefan Schwietzke were funded by the Robertson Foundation. We thank

Dr. Rebecca Fisher who supervised RHUL contribution to the isotopic analysis of Hamburg campaign. Special thanks to Prof. Stefan Bühler from the Meteorological Institute of Hamburg University and Dr. Stefan Kinne from the Max Planck Institute for Meteorology for hosting our team during the Hamburg city measurement surveys. We would like to extend our appreciation to the anonymous referees for the insightful comments which led to improvements of the manuscript. We appreciate continuous efforts from executive and management boards of GasNetz Hamburg, Dr. Luise Westphal, Michael Dammann, Dr. Ralf Luy, and Christian Feickert who facilitated productive communications, provided information on the gas infrastructure in Hamburg and organized leaks repairs with their teams in study area of Hamburg. We also thank asset manager of STEDIN Utrecht, Ricardo Verhoeve who provided information and planned leaks repairs by STEDIN in Utrecht. We thank Charlotte Große from DBI Gas and Environmental Technologies GmbH Leipzig (DBI GUT Leipzig) who helped with clarifying information on reported emission factors provided in national inventory reports. We thank the former MSc students of Utrecht University, Laurens Stoop and Tim van den Akker who helped with the measurements in Utrecht study area.

References

ACM: Authority for Consumers and Markets in the Netherlands, Low NOx Burgners (LNBs) gas code, [online] Available from: https://wetten.overheid.nl/BWBR0037935/2018-05-26, 2018.

Allen, D. T., Torres, V. M., Thomas, J., Sullivan, D. W., Harrison, M., Hendler, A., Herndon, S. C., Kolb, C. E., Fraser, M. P., Hill, A. D., Lamb, B. K., Miskimins, J., Sawyer, R. F. and Seinfeld, J. H.: Measurements of methane emissions at natural gas production sites in the United States, Proc. Natl. Acad. Sci., 110(44), 17768–17773, doi:10.1073/pnas.1304880110, 2013.

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. 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. P.: Assessment of methane emissions from the U.S. oil and gas supply chain., Science, 361(6398), 186–188, doi:10.1126/science.aar7204, 2018.

Brandt, A. R., Heath, G. A. and Cooley, D.: Methane Leaks from Natural Gas Systems Follow Extreme Distributions, Environ. Sci. Technol., 50(22), 12512–12520, doi:10.1021/acs.est.6b04303, 2016.

Brantley, H. L., Hagler, G. S. W., Kimbrough, E. S., Williams, R. W., Mukerjee, S. and Neas, L. M.: Mobile air monitoring data-processing strategies and effects on spatial air pollution trends, Atmos. Meas. Tech., 7(7), 2169–2183, doi:10.5194/amt-7-2169-2014, 2014.

Brass, M. and Röckmann, T.: Continuous-flow isotope ratio mass spectrometry method for carbon and hydrogen isotope measurements on atmospheric methane, Atmos. Meas. Tech., 3(6), 1707–1721, doi:10.5194/amt-3-1707-2010, 2010.

Bright, E. A., Coleman, P. R. and Dobson, J. E.: LandScan : A Global Population database for estimating populations at risk, [online] Available from: https://www.semanticscholar.org/paper/LandScan-%3A-A-Global-Population-database-for-at-risk-Bright-Coleman/17e6076b6761788684434d1e14e85e8877fc0146 (Accessed 23 September 2019), 2000.

Brümmer, B., Lange, I. and Konow, H.: Atmospheric boundary layer measurements at the 280 m high Hamburg weather mast 1995-2011: mean annual and diurnal cycles, Meteorol. Zeitschrift, 21(4), 319–335, doi:10.1127/0941-2948/2012/0338, 2012.

Buendia, E. C., Guendehou, S., Limmeechokchai, B., Pipatti, R., Rojas, Y., Sturgiss, R., Tanabe, K., Wirth, T., Romano, D., Witi, J., Garg, A., Weitz, M. M., Cai, B., Ottinger, D. A., Dong, H., MacDonald, J. D., Ogle, S. M., Rocha, M. T., Sanchez, M. J. S., Bartram, D. M. and Towprayoon, S.: 2019 refinement to the 2006 IPCC guidelines for national greenhouse gas inventories. [online] Available from: https://www.ipcc.ch/report/2019-refinement-to-the-2006-ipcc-guidelines-for-national-greenhouse-gas-inventories/, 2019.

Bukowiecki, N., Dommen, J., Prévôt, A. S. H., Richter, R., Weingartner, E. and Baltensperger, U.: A mobile pollutant measurement laboratory - Measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution, Atmos. Environ., 36(36–37), 5569–5579, doi:10.1016/S1352-2310(02)00694-5, 2002.

Caulton, D. R., Li, Q., Bou-Zeid, E., Fitts, J. P., Golston, L. M., Pan, D., Lu, J., Lane, H. M., Buchholz, B., Guo, X., McSpiritt, J., Wendt, L. and Zondlo, M. A.: Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods, Atmos. Chem. Phys., 18(20), 15145–15168, doi:10.5194/acp-18-15145-2018, 2018.

Chamberlain, S. D., Ingraffea, A. R. and Sparks, J. P.: Sourcing methane and carbon dioxide emissions from a small city: Influence of natural gas leakage and combustion, Environ. Pollut., 218, 102–110, doi:10.1016/J.ENVPOL.2016.08.036, 2016.

Chen, J., Dietrich, F., Maazallahi, H., Forstmaier, A., Winkler, D., Hofmann, M. E. G., Denier van der Gon, H. and Röckmann, T.: Methane emissions from the Munich Oktoberfest, Atmos. Chem. Phys., 20(6), 3683–3696, doi:10.5194/acp-20-3683-2020, 2020.

Curran, S. J., Wagner, R. M., Graves, R. L., Keller, M. and Green, J. B.: Well-to-wheel analysis of direct and indirect use of natural gas in passenger vehicles, Energy, 75, 194–203, doi:10.1016/j.energy.2014.07.035, 2014.

Davis, J. B. and Squires, R. M.: Detection of Microbially Produced Gaseous Hydrocarbons Other than Methane., Science, 119(3090), 381–2, doi:10.1126/science.119.3090.381, 1954.

DelSontro, T., Beaulieu, J. J. and Downing, J. A.: Greenhouse gas emissions from lakes and impoundments: Upscaling in the face of global change, Limnol. Oceanogr. Lett., 3(3), 64–75, doi:10.1002/lol2.10073, 2018.

DVGW: Technische Regel – ArbeitsblattDVGW G 260 (A), Bonn. [online] Available from: https://shop.wvgw.de/var/assets/leseprobe/508866_lp G 260.pdf, 2013.

E. K. Nam, T. E. Jensen, A. and Wallington, T. J.: Methane Emissions from Vehicles, Environ. Sci. Technol., doi:10.1021/ES034837G, 2004.

EDF: Local leaks impact global climate, [online] Available from: https://www.edf.org/climate/methanemaps (Accessed 5 November 2019), 2019.

Efron, B.: Bootstrap Methods: Another Look at the Jackknife, Ann. Stat., 7(1), 1–26, doi:10.1214/aos/1176344552, 1979.

Efron, B.: The Jackknife, the Bootstrap and Other Resampling Plans, Society for Industrial and Applied Mathematics., 1982.

Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, Champman & Hall, London., 1993.

EPA: User's guide for the industrial source guide complex (ISC3) dispersion models, volume II - Description of model algorithms., 1995.

Etheridge, D. M., Steele, L. P., Francey, R. J. and Langenfeld, R. L.: Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability, J. Geophys. Res., 103, 979–993, doi:0148-0227/98/98JD-00923, 1998.

Etminan, M., Myhre, G., Highwood, E. J. and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophys. Res. Lett., 43(24), 12,614-12,623, doi:10.1002/2016GL071930@10.1002/(ISSN)1944-8007.2016GRLEDHIGH, 2016.

Federal Environment Agency: National Inventory Report for the German Greenhouse Gas Inventory 1990 – 2017. [online] Available from: https://unfccc.int/documents/194930, 2019.

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(7), 4091–4099, doi:10.1021/acs.est.6b06095, 2017.

Fisher, R., Lowry, D., Wilkin, O., Sriskantharajah, S. and Nisbet, E. G.: High-precision, automated stable isotope analysis of atmospheric methane and carbon dioxide using continuous-flow isotope-ratio mass spectrometry, Rapid Commun. Mass Spectrom., 20(2), 200–208, 2006.

Fisher, R. E., Sriskantharajah, S., Lowry, D., Lanoisellé, M., Fowler, C. M. R., James, R. H., Hermansen, O., Lund Myhre, C., Stohl, A., Greinert, J., Nisbet-Jones, P. B. R., Mienert, J. and Nisbet, E. G.: Arctic methane sources: Isotopic evidence for atmospheric inputs, Geophys. Res. Lett., 38(21), n/a-n/a, doi:10.1029/2011GL049319, 2011.

Formolo, M.: The Microbial Production of Methane and Other Volatile Hydrocarbons, in Handbook of Hydrocarbon and Lipid Microbiology, pp. 113–126, Springer Berlin Heidelberg., 2010.

France, J. L., Cain, M., Fisher, R. E., Lowry, D., Allen, G., O'Shea, S. J., Illingworth, S., Pyle, J., Warwick, N., Jones, B. T., Gallagher, M. W., Bower, K., Le Breton, M., Percival, C., Muller, J., Welpott, A., Bauguitte, S., George, C., Hayman, G. D., Manning, A. J., Myhre, C. L., Lanoisellé, M. and Nisbet, E. G.: Measurements of $\delta^{13}C$ in CH_4 and using particle dispersion modeling to characterize sources of Arctic methane within an air mass, J. Geophys. Res. Atmos., 121(23), 14,257-14,270, doi:10.1002/2016JD026006, 2016.

Fries, A. E., Schifman, L. A., Shuster, W. D. and Townsend-Small, A.: Street-level emissions of methane and nitrous oxide from the wastewater collection system in Cincinnati, Ohio, Environ. Pollut., 236, 247–256, doi:10.1016/j.envpol.2018.01.076, 2018.

Fukuda, H., Fujii, T. and Ogawa, T.: Microbial Production of C_2 -Hydrocarbons, Ethane, Ethylene and Acetylene, Agric. Biol. Chem., 48(5), 1363–1365, doi:10.1080/00021369.1984.10866323, 1984.

Gallagher, M. E., Down, A., Ackley, R. C., Zhao, K., Phillips, N. and Jackson, R. B.: Natural Gas Pipeline Replacement Programs Reduce Methane Leaks and Improve Consumer Safety, Environ. Sci. Technol. Lett., 2(10), 286–291, doi:10.1021/acs.estlett.5b00213, 2015.

Gioli, B., Toscano, P., Lugato, E., Matese, A., Miglietta, F., Zaldei, A. and Vaccari, F. P.: Methane and carbon dioxide fluxes and source partitioning in urban areas: The case study of Florence, Italy, Environ. Pollut., 164, 125–131, doi:10.1016/j.envpol.2012.01.019, 2012.

Gollakota, K. G. and Jayalakshmi, B.: Biogas (natural gas?) production by anaerobic digestion of oil cake by a mixed culture isolated from cow dung, Biochem. Biophys. Res. Commun., 110(1), 32–35, doi:10.1016/0006-291X(83)91255-X, 1983.

Guisasola, A., de Haas, D., Keller, J. and Yuan, Z.: Methane formation in sewer systems, Water Res., 42(6–7), 1421–1430, doi:10.1016/j.watres.2007.10.014, 2008.

Heilig, G. K.: The greenhouse gas methane (CH4): Sources and sinks, the impact of population growth, possible interventions, Popul. Environ., 16(2), 109–137, doi:10.1007/BF02208779, 1994.

Helfter, C., Tremper, A. H., Halios, C. H., Kotthaus, S., Bjorkegren, A., Sue, C., Grimmond, B., Barlow, J. F. and Nemitz, E.: Spatial and temporal variability of urban fluxes of methane, carbon monoxide and carbon dioxide above London, UK, Atmos. Chem. Phys, 16, 10543–10557, doi:10.5194/acp-16-10543-2016, 2016.

Helmig, D., Rossabi, S., Hueber, J., Tans, P., Montzka, S. A., Masarie, K., Thoning, K., Plass-Duelmer, C., Claude, A., Carpenter, L. J., Lewis, A. C., Punjabi, S., Reimann, S., Vollmer, M. K., Steinbrecher, R., Hannigan, J. W., Emmons, L. K., Mahieu, E., Franco, B., Smale, D. and Pozzer, A.: Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production, Nat. Geosci., 9(7), 490–495, doi:10.1038/ngeo2721, 2016.

Hendrick, M. F., Ackley, R., Sanaie-Movahed, B., Tang, X. and Phillips, N. G.: Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments, Environ. Pollut., 213, 710–716, doi:10.1016/j.envpol.2016.01.094, 2016.

Hmiel, B., Petrenko, V. V, Dyonisius, M. N., Buizert, C., Smith, A. M., Place, P. F., Harth, C., Beaudette, R., Hua, Q., Yang, B., Vimont, I., Michel, S. E., Severinghaus, J. P., Etheridge, D., Bromley, T., Schmitt, J., Faïn, X., Weiss, R. F. and Dlugokencky, E.: Preindustrial 14 CH 4 indicates greater anthropogenic fossil CH 4 emissions, Nature, 578, doi:10.1038/s41586-020-1991-8, 2020.

Hoheisel, A., Yeman, C., Dinger, F., Eckhardt, H. and Schmidt, M.: An improved method for mobile characterisation of δ^{13} CH_4 source signatures and its application in Germany, Atmos. Meas. Tech., 12(2), 1123–1139, doi:10.5194/amt-12-1123-2019, 2019.

Hopkins, F. M., Kort, E. A., Bush, S. E., Ehleringer, J. R., Lai, C.-T., Blake, D. R. and Randerson, J. T.: Spatial patterns and source attribution of urban methane in the Los Angeles Basin, J. Geophys. Res. Atmos., 121(5), 2490–2507, doi:10.1002/2015JD024429, 2016.

Hu, N., Liu, S., Gao, Y., Xu, J., Zhang, X., Zhang, Z. and Lee, X.: Large methane emissions from natural gas vehicles in Chinese cities, Atmos. Environ., 187, 374–380, doi:10.1016/j.atmosenv.2018.06.007, 2018.

IPCC: Guidelines for national greenhouse inventories. [online] Available from: https://www.ipcc-nggip.iges.or.jp/public/gl/guidelin/ch1ref8.pdf, 1996.

Jackson, R. B., Down, A., Phillips, N. G., Ackley, R. C., Cook, C. W., Plata, D. L. and Zhao, K.: Natural gas pipeline leaks across Washington, DC, Environ. Sci. Technol., 48(3), 2051–2058, doi:10.1021/es404474x, 2014.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R. and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophys. Res. Lett., 40(16), 4393–4397, doi:10.1002/grl.50811, 2013.

Keeling, C. D.: The concentration and isotopic abundances of atmospheric carbon dioxide in rural areas, Geochim. Cosmochim. Acta, 13(4), 322–334, doi:10.1016/0016-7037(58)90033-4, 1958.

Keeling, C. D.: The concentration and isotopic abundances of carbon dioxide in rural and marine air, Geochim. Cosmochim. Acta, 24(3–4), 277–298, doi:10.1016/0016-7037(61)90023-0, 1961.

Lamb, B. K., Cambaliza, M. O. L., Davis, K. J., Edburg, S. L., Ferrara, T. W., Floerchinger, C., Heimburger, A. M. F., Herndon, S., Lauvaux, T., Lavoie, T., Lyon, D. R., Miles, N., Prasad, K. R., Richardson, S., Roscioli, J. R., Salmon, O. E., Shepson, P. B., Stirm, B. H. and Whetstone, J.: Direct and Indirect Measurements and Modeling of Methane Emissions in Indianapolis, Indiana, Environ. Sci. Technol., 50(16), 8910–8917, doi:10.1021/acs.est.6b01198, 2016.

LBEG: Geoinformation of Lower Saxony and Schleswig-Holstein, [online] Available from: https://nibis.lbeg.de/cardomap3/, 2018.

Lebel, E. D., Lu, H. S., Speizer, S. A., Finnegan, C. J. and Jackson, R. B.: Quantifying Methane Emissions from Natural Gas Water Heaters, Environ. Sci. Technol., 54(9), 5737–5745, doi:10.1021/acs.est.9b07189, 2020.

Lowry, D., Fisher, R. E., France, J. L., Coleman, M., Lanoisellé, M., Zazzeri, G., Nisbet, E. G., Shaw, J. T., Allen, G., Pitt, J. and Ward, R. S.: Environmental baseline monitoring for shale gas development in the UK: Identification and geochemical characterisation of local source emissions of methane to atmosphere, Sci. Total Environ., 708, 134600, doi:10.1016/j.scitotenv.2019.134600, 2020.

Lyon, D. R., Zavala-Araiza, D., Alvarez, R. A., Harriss, R., Palacios, V., Lan, X., Talbot, R., Lavoie, T., Shepson, P., Yacovitch, T. I., Herndon, S. C., Marchese, A. J., Zimmerle, D., Robinson, A. L. and Hamburg, S. P.: Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region, Environ. Sci. Technol., 49(13), 8147–8157, doi:10.1021/es506359c, 2015.

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, Environ. Sci. Technol., 50(9), 4877–4886, doi:10.1021/acs.est.6b00705, 2016.

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.: MATLAB® code for evaluation of Urban Surveys, Zenodo, doi: 10.5281/zenodo.3928972, 2020a.

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.: Utrecht and Hamburg city measurements data, ICOS, https://doi.org/10.18160/RAJS-KZZQ, 2020b.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A. and Elkins, J.: Law Dome CO2 , CH4 and N2O ice core records extended to 2000 years BP, Geophys. Res. Lett., 33(14), L14810, doi:10.1029/2006GL026152, 2006.

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., Herndon, S. C., Nehrkorn, T., Zahniser, M. S., Jackson, R. B., Phillips, N. and Wofsy, S. C.: Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts, Proc. Natl. Acad. Sci., 112(7), 1941–1946, doi:10.1073/PNAS.1416261112, 2015.

Mitchell, A. L., Tkacik, D. S., Roscioli, J. R., Herndon, S. C., Yacovitch, T. I., Martinez, D. M., Vaughn, T. L., Williams, L. L., Sullivan, M. R., Floerchinger, C., Omara, M., Subramanian, R., Zimmerle, D., Marchese, A. J. and Robinson, A. L.: Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results, Environ. Sci. Technol., 49(5), 3219–3227, doi:10.1021/es5052809, 2015.

Myhre, G., Shindell, D., Bréon, F. M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J. F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T. and Zhan, H.: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York, NY, USA. [online] Available from: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter08_FINAL.pdf, 2013.

Naus, S., Röckmann, T. and Popa, M. E.: The isotopic composition of CO in vehicle exhaust, Atmos. Environ., 177, 132–142, doi:10.1016/J.ATMOSENV.2018.01.015, 2018.

Neumann, G. and Halbritter, G.: Sensitivity analysis of the Gaussian plume model, in Studies in Environmental Science, vol. 8, pp. 57–62, Elsevier., 1980.

Noël, S., Weigel, K., Bramstedt, K., Rozanov, A., Weber, M., Bovensmann, H. and Burrows, J. P.: Water vapour and methane coupling in the stratosphere observed using SCIAMACHY solar occultation measurements, Atmos. Chem. Phys., 18(7), 4463–4476, doi:10.5194/acp-18-4463-2018, 2018.

O'Shea, S. J., Allen, G., Fleming, Z. L., Bauguitte, S. J.-B., Percival, C. J., Gallagher, M. W., Lee, J., Helfter, C. and Nemitz, E.: Area fluxes of carbon dioxide, methane, and carbon monoxide derived from airborne measurements around Greater London: A case study during summer 2012, J. Geophys. Res. Atmos., 119(8), 4940–4952, doi:10.1002/2013JD021269, 2014.

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, Environ. Sci. Technol., 50(4), 2099–2107, doi:10.1021/acs.est.5b05503, 2016.
Paredes, M. G., Güereca, L. P., Molina, L. T. and Noyola, A.: Methane emissions from anaerobic sludge digesters in Mexico: On-site determination vs. IPCC Tier 1 method, Sci. Total Environ., 656, 468–474, doi:10.1016/j.scitotenv.2018.11.373, 2019.
Peek, C. J., Montfoort, J. A., Dröge, R., Guis, B., Baas, K., Huet, B. van, Hunnik, O. R. van and Berghe, A. C. W. M. van den: Methodology report on the calculation of emissions to air from the sectors Energy, Industry and Waste, as used by the Dutch Pollutant Release and Transfer Register., 2019.
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, Environ. Pollut., 173, 1–4, doi:10.1016/j.envpol.2012.11.003, 2013.
Popa, M. E., Vollmer, M. K., Jordan, A., Brand, W. A., Pathirana, S. L., Rothe, M. and Röckmann, T.: Vehicle emissions of greenhouse gases and related tracers from a tunnel study: CO : CO2 , N2O : CO2 , CH4 : CO2 , O2 : CO2 ratios, and the stable isotopes 13C and 18O in CO2 and CO, Atmos. Chem. Phys., 14(4), 2105–2123, doi:10.5194/acp-14-2105-2014, 2014.
Prather, M. J., Holmes, C. D. and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophys. Res. Lett., 39(9), n/a-n/a, doi:10.1029/2012GL051440, 2012.
Rijksoverheid: Emissieregstratie. [online] Available from: http://www.emissieregistratie.nl/erpubliek/erpub/facility.aspx (Accessed 9 December 2019), 2019.
Röckmann, T., Eyer, S., van der Veen, C., Popa, M. E., Tuzson, B., Monteil, G., Houweling, S., Harris, E., Brunner, D., Fischer, H., Zazzeri, G., Lowry, D., Nisbet, E. G., Brand, W. A., Necki, J. M., Emmenegger, L. and Mohn, J.: In situ observations of the isotopic composition of methane at the Cabauw tall tower site, Atmos. Chem. Phys., 16(16), 10469–10487, doi:10.5194/acp-16-10469-2016, 2016.
Schaum, C., Lensch, D., Bolle, P. Y. and Cornel, P.: Sewage sludge treatment: Evaluation of the energy potential and methane emissions with cod balancing, J. Water Reuse Desalin., 5(4), 437–445, doi:10.2166/wrd.2015.129, 2015.
Schmidt, G. A. and Shindell, D. T.: Atmospheric composition, radiative forcing, and climate change as a consequence of a massive methane release from gas hydrates, Paleoceanography, 18(1), n/a-n/a, doi:10.1029/2002PA000757, 2003.
Schwietzke, S., Sherwood, O. A., Bruhwiler, L. M. P., Miller, J. B., Etiope, G., Dlugokencky, E. J., Michel, S. E., Arling, V. A., Vaughn, B. H., White, J. W. C. and Tans, P. P.: Upward revision of global fossil fuel methane emissions based on isotope database, Nature, 538(7623), 88–91, doi:10.1038/nature19797, 2016.
Sherwood, O. A., Schwietzke, S., Arling, V. A. and Etiope, G.: Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Burning Sources, version 2017, Earth Syst. Sci. Data, 9(2), 639–656, doi:10.5194/essd-9-639-2017, 2017.
Sperlich, P., Uitslag, N. A. M., Richter, J. M., Rothe, M., Geilmann, H., van der Veen, C., Röckmann, T., Blunier, T. and Brand, W. A.: Development and evaluation of a suite of isotope reference gases for methane in air, Atmos. Meas. Tech., 9(8), 3717–3737, doi:10.5194/amt-9-3717-2016, 2016.
Stephenson, M. and Stickland, L. H.: Hydrogenase: The bacterial formation of methane by the reduction of one-carbon compounds by molecular hydrogen, Biochem. J., 27(5), 1517–1527, doi:10.1042/bj0271517, 1933.
Thauer, R. K.: Biochemistry of methanogenesis: a tribute to Marjory Stephenson:1998 Marjory Stephenson Prize Lecture, Microbiology, 144(9), 2377–2406, doi:10.1099/00221287-144-9-2377, 1998.
Tong, L. I., Chang, C. W., Jin, S. E. and Saminathan, R.: Quantifying uncertainty of emission estimates in National Greenhouse Gas Inventories using bootstrap confidence intervals, Atmos. Environ., 56, 80–87, doi:10.1016/j.atmosenv.2012.03.063, 2012.
Townsend-Small, A., Disbennett, D., Fernandez, J. M., Ransohoff, R. W., Mackay, R. and Bourbonniere, R. A.: Quantifying emissions of methane derived from anaerobic organic matter respiration and natural gas extraction in Lake Erie, Limnol. Oceanogr., 61(S1), S356–S366, doi:10.1002/lno.10273, 2016.
Turner, A. J., Frankenberg, C. and Kort, E. A.: Interpreting contemporary trends in atmospheric methane, Proc. Natl. Acad. Sci., 116(8), 2805–2813, doi:10.1073/PNAS.1814297116, 2019.
Turner, D. B.: Workbook of Atmospheric Dispersion Estimates, U.S. Environmental Protection Agency. [online] Available from: https://nepis.epa.gov/Exe/ZyPDF.cgi/9101GKEZ.PDF?Dockey=9101GKEZ.PDF, 1969.
U.S. EIA: Natural gas consumptions in the United States, [online] Available from: https://www.eia.gov/energyexplained/natural-gas/use-of-natural-gas.php (Accessed 16 June 2020), 2019.
Van Ulden, A. P. and Wieringa, J.: Atmospheric boundary layer research at Cabauw, Boundary-Layer Meteorol., 78(1–2), 39–69, doi:10.1007/BF00122486, 1996.
Umezawa, T., Brenninkmeijer, C. A. M., Röckmann, T., van der Veen, C., Tyler, S. C., Fujita, R., Morimoto, S., Aoki, S., Sowers, T., Schmitt, J., Bock, M., Beck, J., Fischer, H., Michel, S. E., Vaughn, B. H., Miller, J. B., White, J. W. C., Brailsford, G., Schaefer, H., Sperlich, P., Brand, W. A., Rothe, M., Blunier, T., Lowry, D., Fisher, R. E., Nisbet, E. G., Rice, A. L., Bergamaschi, P., Veidt, C. and Levin, I.: Interlaboratory comparison of δ13C and δD measurements of atmospheric CH4 for combined use of data sets from different laboratories, Atmos. Meas. Tech., 11(2), 1207–1231, doi:10.5194/amt-11-1207-2018, 2018.
UNI MISKOLC and ETE: A register of all gas regulations and norms concerning the necessary gas quality for allowing the transport in the natural gas grid. [online] Available from: https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/redubar_a_register_of_all_gas_regulations.pdf, 2008.

US Census Bureau: U.S. and World Population Clock, [online] Available from: https://www.census.gov/popclock/ (Accessed 20 June 2020), 2020.

Weller, Z., Hamburg, S. P. and von Fischer, J. C.: A national estimate of methane leakage from pipeline mains in natural gas local distribution systems, Environ. Sci. Technol., doi:10.1021/acs.est.0c00437, 2020.

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, Environ. Sci. Technol., acs.est.8b03135, doi: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 data, edited by M. Mauder, PLoS One, 14(2), e0212287, doi:10.1371/journal.pone.0212287, 2019.

West, J. J., Fiore, A. M., Horowitz, L. W. and Mauzerall, D. L.: Global health benefits of mitigating ozone pollution with methane emission controls., Proc. Natl. Acad. Sci. U. S. A., 103(11), 3988–93, doi:10.1073/pnas.0600201103, 2006.

Xu, L. and Jiang, C.: Initial desorption characterization of methane and carbon dioxide in coal and its influence on coal and gas outburst risk, Fuel, 203, 700–706, doi:10.1016/J.FUEL.2017.05.001, 2017.

Yacovitch, T. I., Herndon, S. C., Roscioli, J. R., Floerchinger, C., McGovern, R. M., Agnese, M., Pétron, G., Kofler, J., Sweeney, C., Karion, A., Conley, S. A., Kort, E. A., Nähle, L., Fischer, M., Hildebrandt, L., Koeth, J., McManus, J. B., Nelson, D. D., Zahniser, M. S. and Kolb, C. E.: Demonstration of an Ethane Spectrometer for Methane Source Identification, Environ. Sci. Technol., 48(14), 8028–8034, doi:10.1021/es501475q, 2014.

Yacovitch, T. I., Herndon, S. C., Pétron, G. P., Kofler, J., Lyon, D., Zahniser, M. S. and Kolb, C. E.: Mobile Laboratory Observations of Methane Emissions in the Barnett Shale Region, , doi:10.1021/es506352j, 2015.

Yacovitch, T. I., Neininger, B., Herndon, S. C., Van der Gon, H. D., Jonkers, S., Hulskotte, J., Roscioli, J. R. and Zavala-Araiza, D.: Methane emissions in the Netherlands: The Groningen field, Elem Sci Anth, 6(1), 57, doi:10.1525/elementa.308, 2018.

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., Proc. Natl. Acad. Sci. U. S. A., 112(51), 15597–602, doi:10.1073/pnas.1522126112, 2015.

Zazzeri, G., Lowry, D., Fisher, R. E., France, J. L., Lanoisellé, M. and Nisbet, E. G.: Plume mapping and isotopic characterisation of anthropogenic methane sources, Atmos. Environ., 110, 151–162, doi:10.1016/j.atmosenv.2015.03.029, 2015.

Zhao, W., Zhang, T., Wang, Y., Qiao, J. and Wang, Z.: Corrosion Failure Mechanism of Associated Gas Transmission Pipeline., Mater. (Basel, Switzerland), 11(10), doi:10.3390/ma11101935, 2018.

Zimmerle, D. J., Williams, L. L., Vaughn, T. L., Quinn, C., Subramanian, R., Duggan, G. P., Willson, B., Opsomer, J. D., Marchese, A. J., Martinez, D. M. and Robinson, A. L.: Methane Emissions from the Natural Gas Transmission and Storage System in the United States, Environ. Sci. Technol., 49(15), 9374–9383, doi:10.1021/acs.est.5b01669, 2015.

Zimnoch, M., Necki, J., Chmura, L., Jasek, A., Jelen, D., Galkowski, M., Kuc, T., Gorczyca, Z., Bartyzel, J. and Rozanski, K.: Quantification of carbon dioxide and methane emissions in urban areas: source apportionment based on atmospheric observations, Mitig. Adapt. Strateg. Glob. Chang., 24(6), 1051–1071, doi:10.1007/s11027-018-9821-0, 2019.

Table 1: Natural gas distribution network CH_4 emission categories

Class	CH_4 Enhancement (ppm)	Equivalent Emission Rate (L min^{-1})	Equivalent Emission Rate (≈ kg hr^{-1})	LI Location Colour (Figure 1, Figure 2, and Figure S14)
High	>7.6	>40	>1.7	Red
Medium	1.6-7.59	6 - 40	0.3 – 1.7	Orange
Low	0.2-1.59	0.5 - 6	0.0 – 0.3	Yellow

Table 2- Measurements and results summaries across the study area, inside the ring in Utrecht and north Elbe in Hamburg

Study Area				Utrecht (inside the Ring)	Hamburg (North Elbe)
≈ km street driven		Total km driven		1,000 km	1,800 km
		Driven once		220 km	900 km
		Driven more than once		780 km	900 km
≈ km street covered		Total km covered		450 km	1,200 km
		covered once		230 km	900 km
		covered more than once		220 km	300 km
LIs and emissions		Total number		81 LIs	145 LIs
		LI density		5.6 km covered LI^{-1}	8.4 km covered LI^{-1}
		Total emission rate		290 L min^{-1}	490 L min^{-1}
		Average emission rate per LI		3.6 L min^{-1} LI^{-1}	3.4 L min^{-1} LI^{-1}
		Total emission rate per year		107 t yr^{-1}	180 t yr^{-1}
LIs visited	Once	Number		16 LIs	45 LIs
		Emissions		26 L min^{-1}	68 L min^{-1}
		Average emission rate per LI		1.6 L min^{-1} LI^{-1}	1.5 L min^{-1} LI^{-1}
	More than once	Number		65 LIs	100 LIs
		Emissions		264 L min^{-1}	423 L min^{-1}
		Average emission rate per LI		4.1 L min^{-1} LI^{-1}	4.2 L min^{-1} LI^{-1}
Total LIs categorized based on von Fischer et al. (2017) categories	High (>40 L min^{-1})	Number		1 LI	2 LIs
		Emissions		102 L min^{-1}	145 L min^{-1}
		Average emission rate per LI		101.5 (L min^{-1} LI^{-1})	72.4 L min^{-1} LI^{-1}
		% of emissions		35 % of total emissions	30 % of total emissions
	Medium (6-40 L min^{-1})	Number		6 LIs	16 LIs
		Emissions		84 L min^{-1}	176 L min^{-1}
		Average emission rate per LI		14.0 L min^{-1} LI^{-1}	11 L min^{-1} LI^{-1}
		% of emissions		30 % of total emissions	36 % of total emissions
	Low (0.5-6 L min^{-1})	Number		74 LIs	127 LIs
		Emissions		105 L min^{-1}	169 L min^{-1}
		Average emission rate per LI		1.4 L min^{-1} LI^{-1}	1.3 L min^{-1} LI^{-1}
		% of emissions		36 % of total emissions	35 % of total emissions
Total LIs categorized based on OSM road classes	Level 1	Number		6 LIs	29 LIs
		Emissions		5 L min^{-1}	68 L min^{-1}
		Average emission rate per LI		0.76 L min^{-1} LI^{-1}	2.3 L min^{-1} LI^{-1}
	Level 2	Number		16 LIs	34 LIs
		Emissions		145 L min^{-1}	99 L min^{-1}
		Average emission rate per LI		9.0 L min^{-1} LI^{-1}	2.9 L min^{-1} LI^{-1}
	Level 3	Number		3 LIs	23 LIs
		Emissions		10 L min^{-1}	43 L min^{-1}
		Average emission rate per LI		3.4 L min^{-1} LI^{-1}	1.9 L min^{-1} LI^{-1}
	Residential	Number		45 LIs	52 LIs
		Emissions		93 L min^{-1}	274 L min^{-1}
		Average emission rate per LI		2.1 L min^{-1} LI^{-1}	5.3 L min^{-1} LI^{-1}
	Unclassified	Number		11 LIs	7 LIs
		Emissions		38 L min^{-1}	6 L min^{-1}
		Average emission rate per LI		3.4 L min^{-1} LI^{-1}	0.8 L min^{-1} LI^{-1}
Attribution	C_2:C_1 ratio analysis	Fossil (Inc. combustion)	% of emissions	93 % of total emissions	64 % of total emissions
			% of LIs	69 % of LIs	33 % of LIs
		Microbial	% of emissions	6 % of total emissions	25 % of total emissions
			% of LIs	10 % of LIs	20 % of LIs
		Unclassified	% of emissions	1 % of total emissions	11 % of total emissions
			% of LIs	21 % of LIs	47 % of LIs
	$\delta^{13}C$ and δD analysis	Fossil	% of emissions	----------	79 % of total emissions
			% of LIs	----------	38 % of LIs
		Microbial	% of emissions	----------	20 % of total emissions
			% of LIs	----------	54 % of LIs
		Other	% of emissions	----------	1 % of total emissions
			% of LIs	----------	8 % of LIs (Pyrogenic)

	CH_4:CO_2 ratio analysis	Combustion	% of emissions	2 %	10 %
			% of LIs	7 %	17 %
		Other	% of emissions	98 %	90 %
			% of LIs	93 %	83 %
	C_2:C_1 ratio, CH_4:CO_2 ratio, and $\delta^{13}C$ - δD analyses	Fossil	% of emissions	73 %	48 %
			% of LIs	43 %	31 %
		Combustion	% of emissions	2 %	10 %
			% of LIs	7 %	17 %
		Microbial	% of emissions	8 %	35 %
			% of LIs	4 %	33 %
		Unclassified	% of emissions	16 %	7 %
			% of LIs	46 %	19%
Average emission rate per km driven				0.29 L min^{-1} km^{-1}	0.27 L min^{-1} km^{-1}
km driven / total LIs				12.5 km LI^{-1}	12.36 km LI^{-1}
Emission factors to scale-up emissions per km covered				0.64 L min^{-1} km^{-1}	0.40 L min^{-1} km^{-1}
km covered per LIs	km covered / total LIs			5.6 km LI^{-1}	8.4 km LI^{-1}
	km covered / red LIs			454.8 km LI^{-1}	611.4 km LI^{-1}
	km covered / orange LIs			75.8 km LI^{-1}	76.4 km LI^{-1}
	km covered / yellow LIs			6.1 km LI^{-1}	9.6 km LI^{-1}
km road from OSM (≈ km pipeline)				≈ 650 km	≈ 3000 km
Up-scaled methane emissions to total roads				420 L min^{-1} (≈150 t yr^{-1})	1,200 L min^{-1} (≈440 t yr^{-1})
Bootstrap emission rate estimate and error				420 ± 120 L min^{-1}	1,200 ± 170 L min^{-1}
Population in study area				≈ 0.28 million	≈ 1.45 million
Average LIs emissions per capita (kg yr^{-1} $capita^{-1}$)				0.54 ± 0.15	0.31 ± 0.04
Yearly natural gas consumption				≈ 0.16 bcm yr^{-1}	≈ 0.75 bcm yr^{-1}
Fossil emission factors	C_2:C_1 ratio attribution analysis		Average emission rate per km gas pipeline	0.60 ± 0.2 L min^{-1} km^{-1}	0.26 ± 0.04 L min^{-1} km^{-1}
			Average emission rates per capita	0.50 ± 0.14 kg yr^{-1} $capita^{-1}$	0.20 ± 0.03 kg yr^{-1} $capita^{-1}$
	$\delta^{13}C$ and δD attribution analysis		Average emission rates per km gas pipeline	----------	0.32 ± 0.05 L min^{-1} km^{-1}
			Average emission rates per capita	----------	0.25 ± 0.04 kg yr^{-1} $capita^{-1}$
	C_2:C_1 ratio, CH_4:CO_2 ratio, and $\delta^{13}C$ - δD analyses		Average emission rates per km gas pipeline	0.47 ± 0.14 L min^{-1} km^{-1}	0.19 ± 0.03 L min^{-1} km^{-1}
			Average emission rates per capita	0.39 ± 0.11 kg yr^{-1} $capita^{-1}$	0.15 ± 0.02 kg yr^{-1} $capita^{-1}$
			Average emission rates / yearly consumption	0.10 – 0.12 %	0.04 – 0.07 %

Table 3- CH_4 Emissions from larger facilities in Utrecht and Hamburg estimated with the Gaussian Plume model

Facility	Emission rate (t yr^{-1})
Utrecht	
Waste Water Treatment Plant (52.109791° N, 5.107605° E)	160 ± 90
Hamburg	
F: Compost and Soil Company (53.680233° N, 10.053751° E)	70 ± 50
Upstream D1: 53.468774° N,10.184481° E (separator) D2: 53.468443° N,10.187408° E (storage tanks) D3: 53.466694° N,10.180647° E (oil well)	 D1: 4.5 ± 3.7 D2: 5.2 ± 3.0 D3: 4.8 ± 4.0

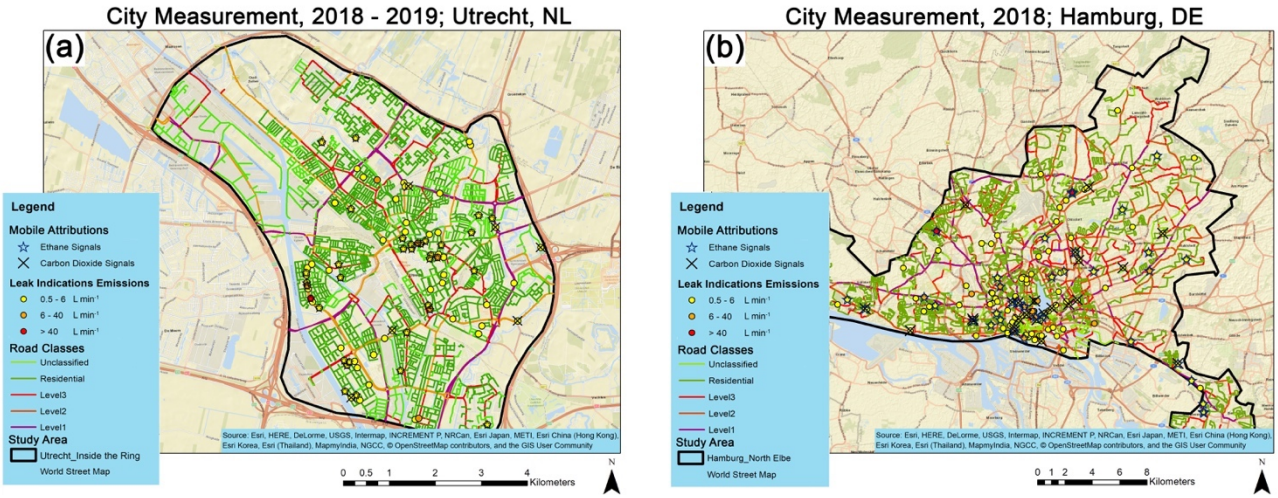

Figure 1: Locations of significant LIs for the categories on different street classes in (a) Utrecht and (b) Hamburg. Road colors indicate the street classes according to the OSM. Black polygons show urban study areas.

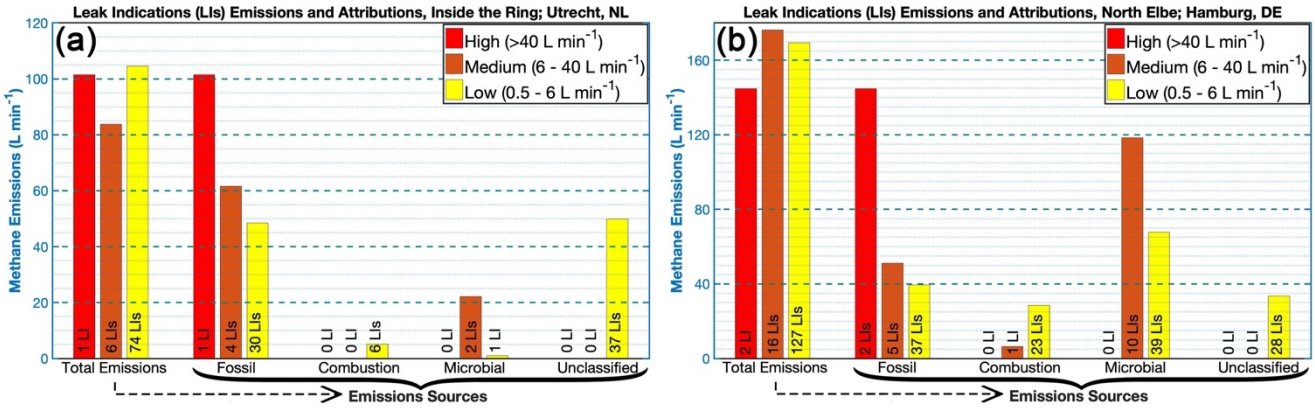

Figure 2: Total CH_4 emission rates from different sources in (a) Utrecht and (b) Hamburg; ***the arrow shows how the emissions are attributed to different sources***

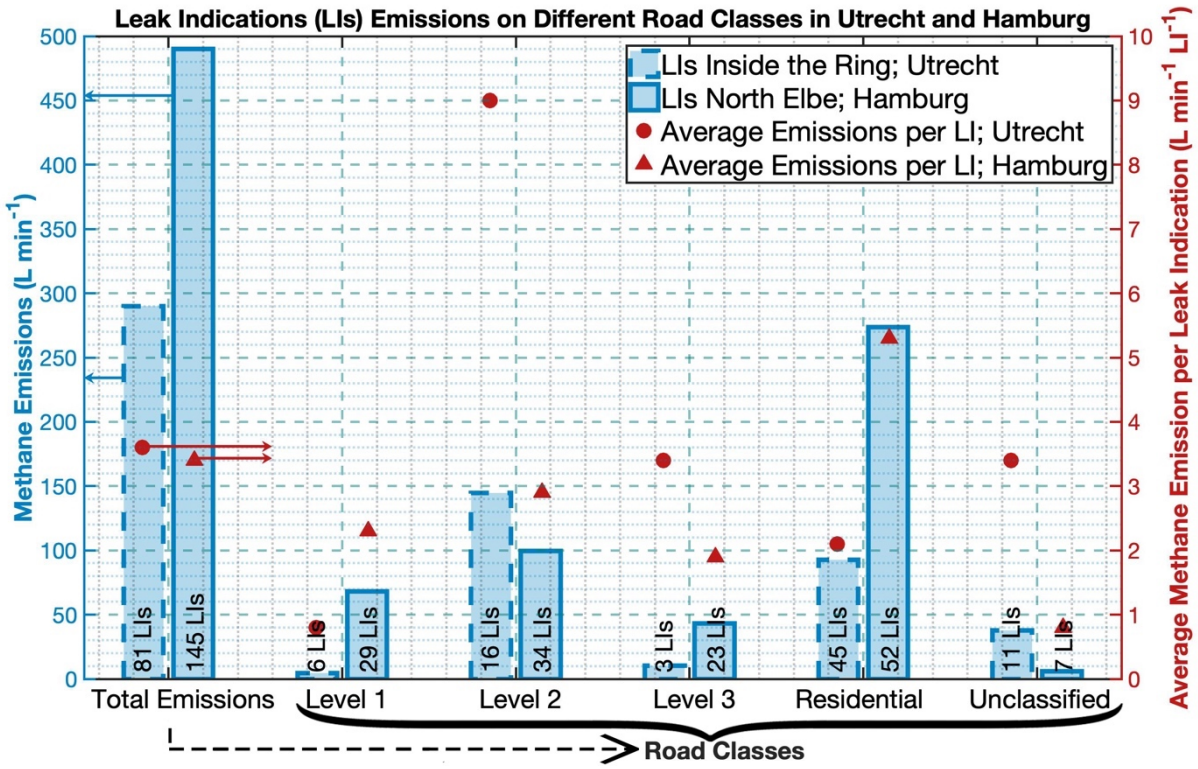

Figure 3: Total CH_4 emissions in Utrecht and Hamburg; ***the arrow shows how the total emissions are distributed on different road classes***

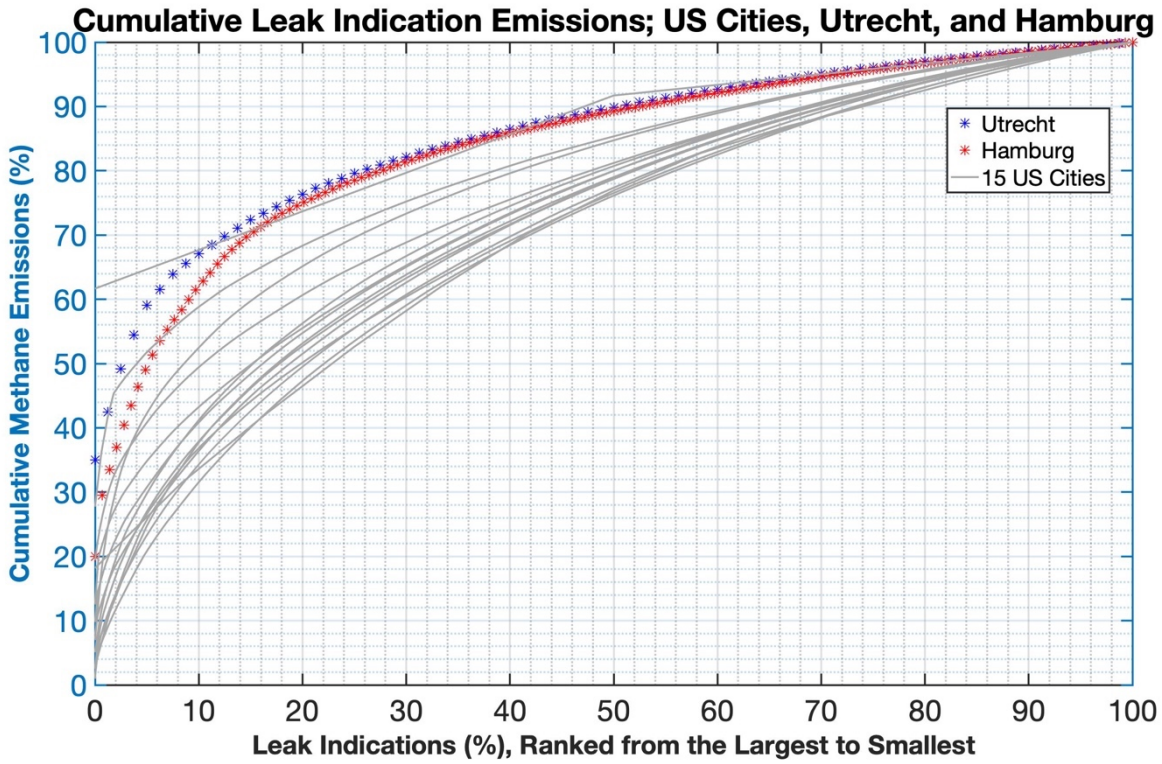

Figure 4: Cumulative plot of CH_4 emissions across US cities, Utrecht, and Hamburg; datasets for the US cities are from Weller et al. (2019)

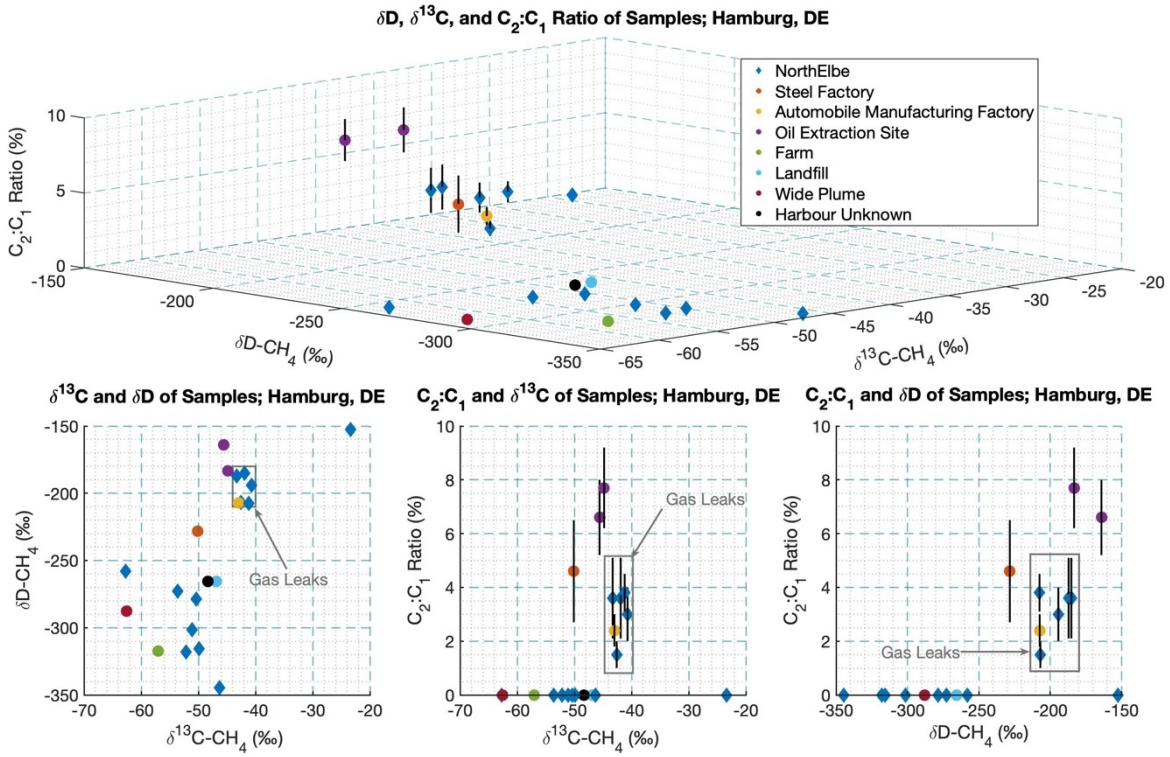

Figure 5: Results from the attribution measurements in Hamburg: C_2:C_1 ratios, and isotopic signatures ($\delta^{13}C$ and δD) of collected air samples; measurement uncertainties in $\delta^{13}C$ is 0.05 - 0.1 ‰ and in δD is 2 - 5 ‰

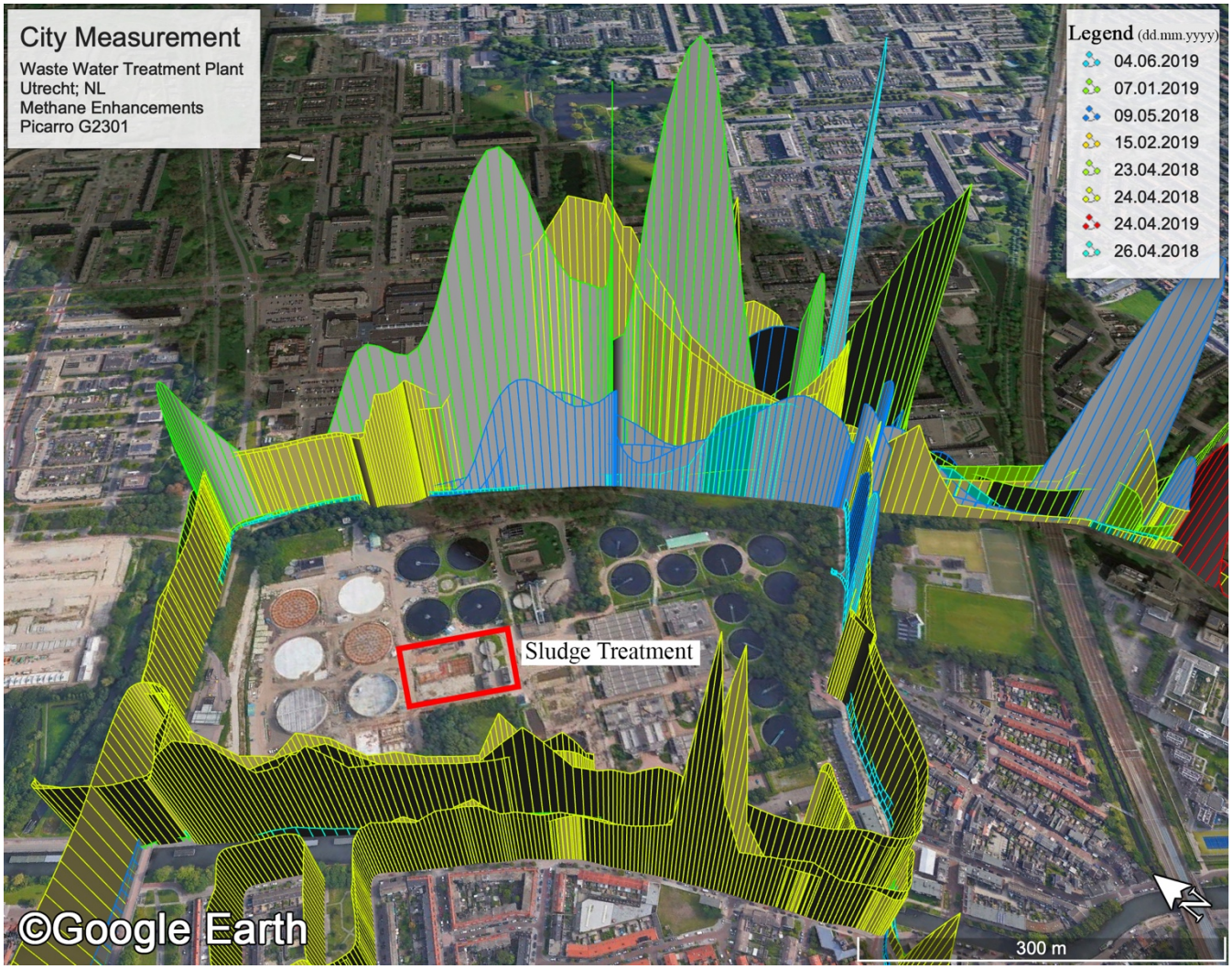

Figure 6: CH_4 enhancements measured downwind waste water treatment plant on Brailledreef street and later used for quantifications from this facility in Utrecht; the centre of the area where the sludge treatment is located was considered as the effective CH_4 emission source, the plumes are plotted on the same scale and max CH_4 enhancement is ≈ 0.3 ppm

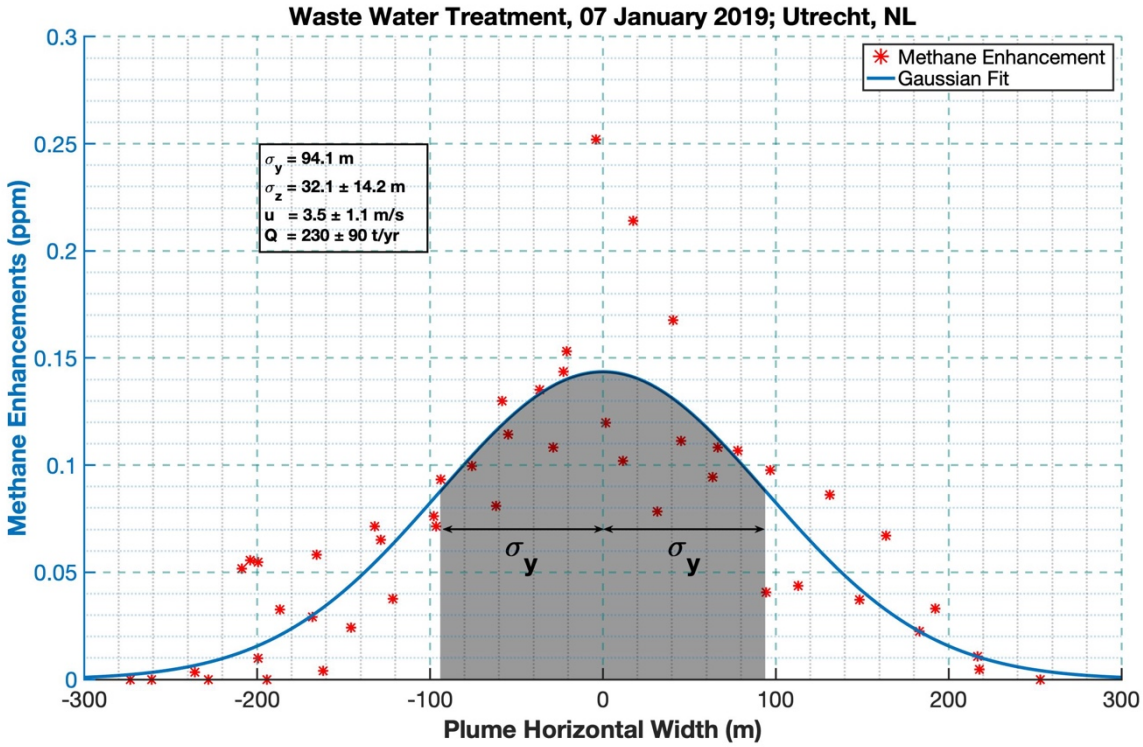

Figure 7: Gaussian curve fitted to some transects downwind the waste water treatment plant in Utrecht