Reviewer 1 comments:

Specific comments:

Title: The title is grammatically incorrect – please correct.

- Proposed change to: "Fugitive natural gas emissions in York, United Kingdom: Updating the parameters of existing algorithms to be based on instrumental limitations."

Abstract, line 16: I don't understand why fugitive emissions can have a direct impact on citizens, as stated here. Please rephrase, or otherwise better explain what this sentence is trying to convey.

- This sentence will be removed, the context it was trying to add does not make sense at the abstract stage.

Line 20: The sentence "This has led to reduced enhancement parameters as well as reduced time clustering parameters" makes no sense to the reader at the abstract stage. Either explain better what this is here, and why its relevant, or delete from the abstract. It would be confusing to any reader of the abstract alone.

- Sentence has been removed

Line 30: The growth rate of methane has not slowed recently as stated. This is a worrying and misleading oversight/inaccuracy. The Kirschke paper cited is 12 years old, and was published after the end of a rare methane plateau, which ended around 2007. Methane's rise has been accelerating in the past 15 years and every contemporary paper confirms this. Please correct this statement and cite a more recent paper which discusses the global average concentration trends.

- We have removed this reference and replaced it with *(Saunois et al., 2025)* and have updated this section accordingly.

Line 39, Nisbet et al., 2025 published a guide to achieving agricultural methane reductions – it would be useful to cite and discuss that paper in this paragraph. https://doi.org/10.1098/rspa.2024.0390

- Reference to *(Nisbet et al., 2025)* has been added to manuscript

Line 82: The units in the equation need to be defined in the text (i.e. what are the specific units of concentration and emission rate), as specific units will only be valid for the numerical constants in the equation. Importantly, the reader needs to know if concentration units are not columns (i.e ppm.metres as defined in the earlier equation on line 75) to avoid confusion.

- Units in this equation, and others have been defined

Section 2.1.2: The discussion of the lab-based baseline is useful. However, it is unclear how the survey baseline was derived – i.e. why was a value of 1.05 times the baseline selected? And how? And can you be confident that a value of 1.05 times a baseline would be appropriate in all survey conditions? Can you be sure that a singular value is appropriate for all types of survey? If not, what should surveyors look at in the data to check that a value of 1.05 is appropriate for their survey?

- Section has been rewritten to provide more clarification into how the baseline is derived. Standard deviation from the methane measurements on each mobile survey was calculated, enhancement criteria was proposed as 5 times this standard deviation divided by the median methane mixing ratio. This calculation was repeated for each mobile survey conducted in the campaign and the final enhancement criteria to be used throughout this campaign was the median of all calculated enhancement criteria. This would recommend an enhancement criteria of 1.01 times the background, however to ensure that the campaign does not focus on the smallest, most diffuse enhancements, enhancement criteria was set as 1.05 times the background. This is still a reduction from the majority of previous literature's enhancement criteria of 1.1 times the background.

Section 2.3 Line 158 – By this point, I have no idea which of the published detection algorithms described in Section 1.1 have been adopted for use in the study. Which one have you used/adapted? It will be useful to remind readers of the algorithm you are referring to, i.e. cross-reference to the earlier section where this is introduced.

- This has been corrected, this section now has a clear recap of the algorithm used within the York campaign.

None of the figures are referred to, or discussed in the narrative at all. This is not acceptable. Every figure must be referred to and discussed in the text.

- Figures are now referenced and discussed within the manuscript.

Line 207 – Presumably this relationship was calculated from Figure 5, though figure 5 is not mentioned at all in the text. What is the uncertainty in the emission rate resulting from the quality of the fit to the data in Figure 5 etc? What is the goodness of the linear fit in general? This is perhaps the most significant comment I would make about the paper in terms of the novel science it offers. Why would a relationship such as this (calculated for a controlled release at Bedford) be applicable to releases in York? As stated earlier, previous studies have quantified their own relationships for other cities/environments, under very different conditions, so how is the relationship for Bedford relevant elsewhere? And what is the potential uncertainty that results in applying it elsewhere? The fact that some other studies cited seem to have been published with similarly worrying oversights does not validate the approach here. A quick look at Figures 2 to 5 show just how flaky any empirical relationship is. I do believe that mobile methods such as this can be a great way to detect fugitive emissions in urban environments in general. But I am not convinced here that they can be used to derive emissions rates with any traceable error, especially using empirical relationships derived in one location/conditions, and then applied to very different environments (e.g. with buildings/obstacles perturbing flow). It makes no conceptual dynamical sense. The relationship between enhancement and emission rate would be different for different windspeeds, surface roughness etc. With that in mind, I have no confidence in the accuracy of the leak rates presented in Figure 12 and section 3. The methods described in the paper could be very useful for leak detection, but not leak quantification. I would recommend the paper is resubmitted to describe mobile use in leak detection and source attribution (i.e. source type using ethane and NOX as discriminators for example). Alternatively, a robust emissions error quantification, accounting for differences between Bedford and York conditions, could make the results in Section 3 more meaningful.

- An alternative quantification methodology is proposed, following (Tettenborn et al., 2025) where peak area is used in the quantification equation. Results from the controlled release experiment conducted are used to demonstrate that this is an appropriate methodology to follow with the instrument package utilised in this study. We will expand on this with a modelled Lagrangian dispersion experiment (using the Graz Lagrangian Model) based on the controlled release environment to better understand whether the peak area is expected to be consistent at inlet height throughout the typical distances from source encountered for roadside emissions. We will also explore the impact of considering wind speed into the emission quantification algorithm, and whether it could / should be utilised in this type of reporting.- In order to avoid overconfidence in the reported leak rates, quantified leak rates will be assigned to one of several bins, very small (0 - 2 L min^{-1}), small (2 - 6 L min^{-1}), medium (6 - 40 L min^{-1}) and large (> 40 L min^{-1})
- We agree that over distances further than 30 m, quantification is likely to be greatly affected by atmospheric conditions, buildings or obstacles. The majority of methane emissions we expect to see outside of the 30 m are likely to not be the below-road or near-road thermogenic methane emissions associated with the natural gas distribution grid. With this in mind, we intend to move away from quantifying the non-thermogenic methane emissions, instead reporting on these sources will only be on the basis of the number/frequency of detection.

Technical corrections:

Line 49: Change to “UK’s”.

- Change made

Line 51. Full stop in wrong place? Sentence does not make sense.

- Full stop moved to after reference

Line 84. Full stop needed after …”emissions”.

- Change made

Line 86 Change “has” to “have”.

- Change made

Line 135- sentence does not make sense.

- In line with this comments and previous comments on this section, the section has been rewritten in order to provide more clarity around this variation experiment.

Line 142_ Change “drives,…” to “mobile surveys” to avoid confusion. Check other instances.

- Change has been made to all instances

Figure 2 is not referred to explicitly in the text. Please add to Section 2.2.

- This and reference to all other figures has been added.

Line 160 – space needed between number and unit “5 s”, as has been done correctly on line 158 – please check and correct throughout. “44.5m” on line 161 is another example correction. There are many more.

- This and all similar mistakes have been corrected.

At this point, I will have to stop listing typos, grammar mistakes and other technical corrections. It is not a reviewer’s job to correct a paper in detail. The paper needs a thorough proof-read. Some of the mistakes do affect the potential meaning of the narrative. There are several senior co-authors on this paper – it is disappointing that they did not help the lead author a little more with this, or insist it was checked well prior to submission.

- A thorough rewrite and proof reads have taken place to ensure the corrected manuscript does not feature these mistakes.

Line 283 – equation has a black square – something went wrong in the version online?

- Correction to the final pdf copy has been made, black square is no longer visible at point of submission.

Addendum to review above:

I would like to offer the team a way to prove me wrong if they are confident that I may have misunderstood the empirical emissions estimation method. I offer this as this paper could also be a really interesting investigation of the pros and cons and accuracy of empirical emissions methods from mobile sampling - a discussion that has been all too lacking in many other studies that have attempted to apply it, but which has gone mostly unchallenged in the literature thus far. I am concerned there is a group think going on with this method, as the first of its kind slipped the net. I could be wrong though. This paper could be a great contribution to that debate if it refocuses to be a narrative about how (and how not) such methods could be used.

My assertion that an emissions relationship derived under a specific set of environmental conditions is not transferrable to any other (significantly different) set of conditions is based on how any emission plume would be conceived to advect and disperse due to wind, turbulence, convection and diffusion. Any change in any one of these parameters would affect the downwind concentrations from a point source sampled downwind and mapped by a mobile vehicle (whether near to a source, or further afield). As the empirical emissions method solely relies on measured concentration enhancements and wind at an unknown distance from source downwind, the emissions would therefore be calculated to be different if any of these environmental parameters varied, compared to the conditions under which the empirical relationship was derived. Is this correct so far? If not, I apologise completely as I've misunderstood the method and I should be corrected.

Assuming the above assumptions I have made are correct(?), the team could attempt to calculate the emissions error that would result if any of the operational (i.e. a later survey using the relationship) environmental conditions varied in some defined parameter space, by using a physical advection-dispersion relationship (such as a Gaussian Plume) as a control. For example, a simple simulation/test could be set up using a 2D Gaussian plume equation (or convection/diffusion equation if very local scales are preferred). For plumes with the same simulated emission rate in all cases, but with different environmental conditions (wind and surface roughness etc), how would the emissions from the same derived empirical relationship change as a function of those conditions if you pseudo-sampled the simulated plume at ground level using the instrumentation described in the paper? If you were able to test this in such a simulation, you could get a handle on the potential errors that you might get in the York environment versus the validation survey by making some informed assumptions/observations about the wind and roughness conditions in York versus the validation environment. I appreciate that the whole idea of an empirical relationship is to sidestep the impossible complexity that comes with such varying conditions, but for it to be useful, a conceptual validation must be shown at the very least. If you were able to do this in this paper, it would not only place meaningful uncertainties on the emissions derived in Section 3, but it would also provide an important commentary and resource on the utility of this method for others to follow. If you did it, it would turn this into a very impactful paper I believe and a very useful course correction to the method.

I wanted to offer this idea as I feel awful about recommending the paper is rejected. I recognise the effort of the fieldwork. If the team would prefer to take this idea offline and revise the paper with my help along the above lines, I'd be happy to drop anonymity if you would like to withdraw the paper and work on it together. Alternatively, if I've totally misunderstood the method and you can defend it conceptually and convincingly in review here, I'd be very open to being shown why I am wrong. If that were the case, the other aspects of the paper raised in my review could be corrected via major corrections. Well done again with the fieldwork and I look forward to seeing what you think about these ideas.

- We feel that our previous comments have dealt with these points and hope this satisfies the reviewer

<u>Reviewer 2 comments:</u>

<u>Sections</u>

Section 1.1 Good initial summary of previous algorithms, but this section can be improved, but the equations from von Fischer et al. (2017) and Weller et al. (2019) are presented without clear explanation of terms (e.g., ppm.metres index). A brief description would help non-specialist readers. Please also read and refer to other relevant studies in which they also derived similar equations, Tettenborn et al.(2025) and Joo et al. (2024). Authors can make an intercomparison emission rate quantifications derived from these equations in a table and discuss the results.

- Greater description will be added to these terms as well as the addition of Tettenborn et al. (2025) and Joo et al., (2024).

Sections 2.1.1 and 2.1.2 Are these efforts made to address instruments drifts? If so, explanations and descriptions can be provided in a better format.

- Section 2.1.1 is designed to understand the response rate of the TILDAS, in order to find the true sampling rate. All other instrumentation used had well defined response rates. However, as the TILDAS had been adapted for use in mobile sampling, its response rate was unknown. The TILDAS is capable of measurements at 10 Hz but this is not a guarantee of true 10 Hz measurements. The response rate found in section 2.1.1 can then be used to inform the time clustering parameter of the algorithm (i.e., the time within which separate enhancements cannot be reliably separated and have to be considered as the same enhancement)
- Section 2.1.2 describes the variation present within the measurements of methane from the MGGA, this section is used to inform the algorithm's enhancement parameter.
- Instrument drift is not addressed as all instruments record to a central computer, instrument lag (between MGGA and TILDAS) is addressed in section 2.4.2.
-We have improved the clarity of these sections.

Section 2.2 It appears that the measurements only took place on the main roads? Why didn't you choose the residential alleys and avenues? In urban surveys those narrower streets have high significance.

- This survey was intended as an initial overview of the efficacy of the updated algorithm and therefore stuck to major roads in order to focus on finding the largest fugitive emissions.
- The aim of this sampling was to get good spatial coverage of the city and maintain as constant a driving speed as possible. By removing the slower, time consuming transects through smaller streets we were able to achieve that.
- Additionally, our previous UK studies in *(Vogel et al., 2024)* focussed on covering the main thoroughfare for each 1 km x 1km grid square but when a street by street approach was undertaken, the majority of emissions detected were from domestic boiler emissions as opposed to pipeline leaks. A smaller survey of intensive driving in York also showed this, the results of which can be added to the SI. However, as losses from gas distribution was our aim with this research, this city-wide route was selected.

Section 2.4.1 This is an important part of your work and should be discussed and described in more details. See comments about the intercomparison with other equations introduced previously.

- More detail will be added to this section. A larger amount of detail will be added in terms of the discussions of different quantification equations. Additionally, another section will be added to the results showing how each quantification equation affects the resulting calculated leak rates.

Section 2.4.2 This section including the graphs can be placed in the supplementary information and main results can be used in the algorithm in the mani manuscript.

- Plots will be moved to the SI while description of the results will be kept in the main manuscript.

Section 2.5 The attribution part using mobile ethane and carbon dioxide is of high relevant. Are the values in X-axis are methane readings (including the background) or the values show methane enhancements? This is important because it is unclear how you classified the values around the 2 ppm into the source specific categories? For further use of ethane and carbon dioxide, I refer you to Fig. S12 (See SI from Maazallahi et al. (2020)) and Sect. S3 and Sect. S5from Maazallahi et al. (2023).

- Values shown in the figure are raw CH_4 measurements, sources are not assigned of single point measurements (i.e. a point measuring 2 ppm CH_4) rather, they are measured across a 10 second window of the enhancement, a linear regression is used to find the gradient of C_2:C_1 over this window, which is then used to assign sources. This section will be rewritten in order to provide more clarity. We also intend to include the further use of CO_2 and C_2H_6 that you recommend from the above references.

Section 3.2 Do you intend to only discuss the pyrogenic sources here? I can see from Fig. 12 that other sources are also shown in the figure.

- The section will be renamed, as Pyrogenic sources are not the only ones discussed in this section.

Section 3.3 Please see comments with regards to the use of other equations. An intercomparison can be made here and discussed in details. This can be a very significant part of your work

- This comparison will be added. It will be especially relevant in the adapted manuscript as we will move to quantification using peak area.

Lines:

L1-3: Title is clear but could be more concise. Consider: “Adapting mobile methane survey algorithms to instrument specifications: a case study in York,UK”.

- Proposed change to: “Fugitive natural gas emissions in York, United Kingdom: Updating existing algorithm parameters based on instrumental limitations.”

L19-20: “This study aims to adapt existing algorithms parameters...”.grammatically awkward. Suggest: “This study adapts existing algorithm parameters by investigating the limitations of the mobile survey platform instrumentation.”

- This change will be made

L21: The claim that old methods may underpredict LIs by 53.5% is striking but should be contextualized: is this due to sensitivity changes or source filtering?

- Context will be added, the change is due to source filtering.

L23: Emission reduction from 185.10 to 60.23 L min^{-1} is very significant.

- This will likely change during corrections as we move towards a peak area quantification and non-specific quantification where LIs are assigned to leak rate bins

L28-30: This sentence reads awkward.

- Changed to: “Signatories of the pledge aim to reduce their 2020 methane emissions by 30 % by 2030. This change was brought about due to increasing concern surrounding methane’s role as a climate forcer.”

L45-61: The transition to UK natural gas network is relevant, but the paragraph could better link to the study's objectives. Better to reduce non-necessary information and provide information about emissions.

- Excess information on the NTS and GDNs will be removed. Instead a focus on the overview of the UKs distribution network, followed by a description on the study's focus on detection of emissions from the GDNs.

L59: The sentence on "many unknowns" is vague. Consider rephrasing to emphasize the need for improved detection methods.

- Change to: "The large distances of natural gas pipework leads to large uncertainties in where a fugitive emission may occur, this highlights the need for mobile measurement surveys, in order to locate these emissions."

L76–85: The introduction of source attribution (CO_2, C2H6 and isotopes) is well-placed and sets up the study's contribution.

L86–90: The final paragraph clearly states the paper's aims. Well done.

L91–105: Instrument details are thorough, but the ethane calibration formula is presented without defining "mean response." The calibration procedure can be moved to supplementary information.

- Additional discussion on the TILDAS will be added.

L112–119: The description of TILDAS valve modifications is technical and unclear. However, the rationale for aiming for "true 10Hz" vs. 5Hz is not fully justified. It is also unclear why other parameters are important in terms of emission rate quantifications. These descriptions belong to supplementary information and main findings can be placed in the mani manuscript.

- Why did you the measurements in normalized format?
- Better show the whole measurement graphs.

- More description is required here in order to provide clarity. However, these modifications are not important for quantification of enhancements, but rather for the source attribution of enhancements themselves.

L160-162: Here you point out an important relation between two parameters;driving speed vs sampling frequency. With lower speed and higher measurement frequency, there is higher possibility to capture methane enhancement signals.

- We agree, although the measurement frequency must be due to the response rate and flow rate of the instrument, not the data acquisition speed. This is the key discussion point in 2.1.1

L169-170: do you suggest that a quantification equation should be introduced for each city and instrument used for individual survey? This means that in every urban surveys (city and instrument dependency) a set of control release experiments should be performed.

- We initially wished to present that quantification equations were unique to measurement platforms and that quantification equations based on peak height would never result in a quantification equation that would be applicable to a different platform. However, after comments from the first reviewer and through discussion amongst the authors, we want to move towards a peak area quantification equation, that we believe would be applicable to other platforms regardless of instrumentation.

<u>Equations</u>

Equations are not numbered throughout the manuscript.

- Equation numbers and references to them will be added to the corrected manuscript.

L283 – L208 – L202 – L106 etc.: equations are not numbered.

L283: what are those back boxes in the equation?

- Artifacts appear to have been added during the conversion to a .pdf, this has been corrected.

Figures

Figures are not referenced within the text. Many figures can be moved to the supplementary information and main figures can be kept in the main manuscript.

- References to figures have been added throughout the manuscript, discussion amongst the authors will take place as to which figures should be moved to the SI,

Figure 10 – Use of wind direction and LI assignment are very unclear. It appears that the wind speed values are not corrected and include driving speed as well. The figure is very vague and values are questionable.

- Wind speed is corrected, however, it is expected that wind speed is appearing abnormally high due to disturbance from other vehicles on road, the figure will be amended to remove wind speed and any discussion involving wind speed will also be removed.

Figure 11: texts are in small fronts. What are the use of those LI numbers? How do you interpret those numbers with regards to e.g. Fig. 10. It is better to discuss specific LIs in details.

- Both figures will be adapted to discuss the LIs in more specific terms

Tables:

Table 1 – Results provided in this table require further, detailed and clear explanation and discussion in the main text. Results from other algorithms can be also included in the table.

- This section will be restructured, providing clearer reference as to which results refer to the algorithm referred to in *(von Fischer et al., 2017; Weller et al., 2019)*, which are from the new method and which are derived from partially updating the old algorithm to the newer algorithm.