Response to each referee's comments on "High potential for CH_4 emission mitigation from oil infrastructure in one of EU's major production regions"

Response to anonymous referee #1 comments

We thank the anonymous referee for taking the time to thoroughly review our manuscript and for providing us with insightful and constructive comments that have greatly contributed to its improvement. Responses to individual referee comments are below. This document is organized as follows: the Referee's comments are in plain text, our responses are in blue, and all the revisions in the manuscript are shown in blue *italic*. Line numbers refer to the updated WORD manuscript with tracked changes.

Anonymous Referee #1

The authors combined ground-based measurements conducted by Delre et al and Korben et al with new UAV measurements to better estimate methane emissions from oil (and gas) production in Romania and to identify leaking components on well sites. They find that methane emissions from oil infrastructure in Romania exceed inventory estimates by a factor of two or much greater and that most of the emissions come from operational venting. These findings are timely and provide essential information from rapidly reducing methane emissions from oil infrastructure in Romania.

The authors have performed an in-depth statistical analysis of measurement distributions and the role of measurement method and the detection limit/proportion of non-detects. I particularly like their analysis of detection limits and non-detects, which unfortunately are not highlighted in the main paper.

We thank the reviewer for the positive feedback on our analysis of detection limits and non-detects described in the supplementary material. We value the observation that this information was not adequately highlighted in the main paper. In response to this comment, we have revised the manuscript to incorporate additional comprehensive details of this analysis directly into section 2.2 of the main text.

A few sentences on the history and development oil production in Romania is needed. In addition, the authors should describe how the different datasets vary in terms of geology and operational practices, including comments about representativeness, and not just measurement methods. The measurements compared were made at different well sites and points in time. Although the role of temporal variability and site-by-site variations are acknowledged, the way that the results are presented, it appears that measurement method is the most important factor.

Thank you for your valuable comments. We have carefully considered your suggestions and made the following clarifications in our paper. Firstly, we have provided additional details regarding the type of sites measured, representativeness and the development of O&G production in Romania. Secondly, we want to emphasize that the primary focus of our analysis is to integrate different datasets and characterize emission distributions, rather than explaining differences between measurement methods. Following one of your detailed comments, we have made this clarification in the main text to ensure a clear understanding of our research objectives.

One of the final concluding points is that well sites with H2S emit less methane, with the argument being that they are more tightly managed due to safety reasons. However, as noted in the detailed comments, the literature shows conflicting results, which may be due to geology, status of the well, or something else. Therefore, I suggest providing additional qualifiers and reference relevant studies (such as those noted below).

Please see our replies to the individual comments below.

Most of the revisions that are needed are related to text and wording to clarify the messages. In general, it would be helpful to provide enough description so that readers can follow the topic without referring to the SI or another publication. Although a lot of the questions in the detailed comments below are answered in the SI, it would be better if a short description is provided in the main paper.

Based on the recommendations from both reviewers, we have revised our manuscript and we have incorporated additional information from the supplementary material, as indicated in the specific comments below, into the main text.

DETAILED COMMENTS ON MAIN PAPER

1. Line 44-46: Provide ranking of Romania (e.g., third or fourth) in terms of methane emissions from the energy sector in the EU.

   Response. Done.

   Lines 45-48: *“Romania, a key O&G producer within the EU, with the second highest reported annual CH_4 emissions from the energy sector in year 2020 (Greenhouse Gas Inventory Data - Comparison by Category, 2022), can play an important role towards the EU’s emission reduction targets. “*

2. Line 54: If possible, provide a more quantitative ranking (e.g., among the top 10).

Response. Done.

Lines 54-55: *"This estimate represents the third highest when compared to measurement-based estimates of similar facilities from other production regions. "*

3. Line 56: how representative are the studied areas of all areas in Romania?

Response. As mentioned in the manuscript, the majority of Romania's oil reservoirs are located in the southern part of the country, where the ROMEO measurement campaign took place in 2019. A few other smaller oil fields are located in the north-west and north-east parts of the country, as well as in the Black Sea. However, we do not have information about the production facilities in these areas and we cannot evaluate the representativeness of our studied areas to those areas. Most of Romania's natural gas reservoirs are located in the Transylvania Basin, where the second part of the ROMEO measurement campaign took place in 2021. The results of this campaign will be presented in a future publication.

4. Line 58: should it read "the entire Romanian oil production sector"? Based on the previous sentence, it appears that the 120 ktons/year is for a subset of the Romanian oil production sector.

Response. Yes, it should be read as *"This is approximately 2.5 times higher than the reported emissions from the entire Romanian oil production sector for 2020".* Our estimated emissions of 120ktons/year account for only the subset of oil production sites. However, due to the lack of detailed information on the emissions from specific types of sites in the national inventories, we compare this result with the reported emissions from all types of infrastructure in the Romanian oil production sector.

5. Line 58: Is 2020 representative of methane emissions from other years? Because of the pandemic, there were changes in methane emissions in many oil and gas producing regions.

Response. We believe that the year 2020 is representative of methane emissions from other years. There was a 3.5 % and 2.9 % decrease in the reported emissions of year 2020 compared to years 2019 and 2018 respectively. These percentages are comparable with the reported reduction of emissions throughout previous years.

6. Line 59: was the source level characterization done as a part of the measurement campaign? Or was it based on other inventory estimates?

Response. Yes, the source level characterization was done as part of the measurement campaign. As mentioned in lines 301-303, *"On the component scale, the combination of* *an Optical Gas Imaging (OGI) camera for the detection of potential leak sources and a Hi-Flow Sampler (HFS) device for the quantification of the emissions was implemented".*

7. Line 75: Replace "Whereas" with "Although"

Response. Done.

8. Line 95: add "the" in front of "largest"

Response. Done.

9. Line 96: Delete "," after the first "EU"

Response. Done.

10. Line 108: provide a list of measurement platforms in parenthesis. Are they all ground-based? It's a bit confusing because UAV measurements are also included in this paper.

Response. During the ROMEO campaign, vehicles, UAVs and manned aircrafts were used to perform measurements of the oil and gas facilities. In this paper, we analyse the emissions measured by vehicles and UAVs. We referred to these measurements as "ground-based", but we understand that the use of this term is confusing. We have corrected this throughout the text. We now refer to these measurements as "ground and drone-based" measurements. As suggested, we modified the text and added the list of the measurement platforms used in the measurement campaign.

Lines 111-116: *"The goal of this project was to characterize CH_4 emissions at a component, facility and basin scale using a variety of measurement platforms e.g., vehicles, Unmanned Aerial Vehicles (UAVs), or commonly referred to as drones, and manned aircrafts. Through the use of a range of emission quantification methods, the ROMEO campaign aimed to provide a comprehensive quantification of CH_4 emissions related to onshore O&G production in Romania."*

11. Line 117: "Ground-based" is mentioned numerous times in this paragraph. However, there is no description of what is meant by "ground-based" and what "non-ground-based" would represent. What are UAVs considered as?

Response. As stated in the reply of the previous comment, in the original preprint manuscript, measurements conducted by vehicles and UAVs were considered as ground-based measurements. Data from aircrafts were considered as non-ground-based measurements but are not part of this paper's analysis. To avoid the confusion, we now use the term "ground and drone-based" measurements.

12. Line 136: in the introduction, it sounded as though only ground-based data would be used. See previous comments.

Response. See previous two answers.

13. Line 158: What counts as a cluster of high-density of production facilities? Provide a quantitative limit.

Response. More details on the description of the clusters are provided earlier in the manuscript. Specifically, we mention in lines 127-130 "*Figure 1 shows that the O&G production infrastructure is concentrated in smaller clusters that cover areas between 2 and 120 km^2, each containing 10 to 582 oil and gas related sites such as oil wells, gas wells, compressor stations and oil parks*". For further clarification, the number of facilities is now also included in the caption of the figure.

Line 158-159: *"The grey shaded areas indicate clusters with high density of production facilities (number of facilities ranging between 10 to 582), in some cases the symbols hide the areas. "*

14. Line 162: specify whether the authors had site access or if they were screening from public roads only.

Response. Screenings were performed from public roads. This is now mentioned in the main text.

Lines 163-166: *"Screenings were performed from public roads and the goal was to identify potential emissions at the site and check site accessibility, considering factors such as roads condition, time limitations, and local restrictions imposed by operators".*

15. Line 165: define "site accessibility". Is this because operators did not give permission or because the roads are not in good condition or something else?

Response. Indeed, these two are the most important factors to consider when referring to site accessibility. In order to avoid any potential bias in the measured emissions, the operators were not informed in advance about our visit to the facility. Therefore, access to the facility was not always granted. We have now added more information on site accessibility.

Lines 163-168: *"Screenings were performed from public roads and the goal was to identify potential emissions at the site and check site accessibility, considering factors such as roads condition, time limitations, and local restrictions imposed by operators. To prevent any potential bias in the measured emissions, the operators were not informed in advance about our visit to the facility, resulting in occasional restricted site access".*

16. Line 169: is there data on farms and if so, was this data analyzed?

Response. The focus of the campaign was emissions from oil and gas infrastructure, therefore emissions from farms were not measured or analysed.

17. Line 186: how close was the emission point from the tracer gas release point?

Response. As mentioned in lines 81-87 of the Supplementary Material: *"To release the tracer gas as closely as possible to the emission point, a flexible tube was pushed to the location of the well borehole by using a rod. In cases where this was not possible, such as at large area sources, the tracer was released from the side of the fence protecting the target area".* The exact distance between the emission point from the tracer gas release point was not recorded.

18. Line 214-215: did Korben et al only use GPM?

Response. Korbeń et al. (2022) performed measurements using both GPM and OTM-33A. This is now stated more clearly in the main text.

Lines 179-182: *"Here we provide a brief description of each measurement method. Delre et al. (2022) provides additional information on the deployment of TDM and GPM during the ROMEO campaign, while Korbeń et al. (2022) offers details specifically on the deployment of OTM-33A and GPM. "*

19. Line 256: are the differences in MBA between UG and EMPA applications important or can they be viewed as minor? If they are important, they should be briefly (1 sentence) described here.

Response. The way the wind is taken into account differs between the two MBA methods. EMPA's approach clusters measured methane mole fractions before kriging, where the normal wind components of continuous wind measurements were projected onto the drone positions (Morales et al., 2022), in contrast to UG, which applies the mean wind, observed throughout the flight, uniformly to all measured methane mole fractions (Vinković et al. 2022). This information has been incorporated into Section S2, which elaborates on the Mass Balance Approach, and the following changes have been implemented in the main text.

Lines 255-258: *"Both UAV-based techniques use an MBA to quantify the emission rates from sampled oil and gas facilities, but there are certain differences in the MBA between UG and EMPA application, including factors such as the treatment of wind, which are presented in the supplementary material. "*

20. Line 302: what is the detection limit of the HFS?

Response. During the ROMEO campaign, the lowest component emission rate measured with the HFS was 0.02 kg/h. Due to the low number of measurements for oil well sites using the HFS, we did not perform an in-depth analysis and therefore did not provide detailed information for this measurement method and its detection limit. For context, Allen et al. (2013) reports a detection limit of 0.55 g/h for the HFS.

21. Line 306: how many sites did the authors have access to (i.e., not from the road)?

Response. On the component scale, emissions were detected at 90 sites using the OGI camera. Out of these, emission rates from leaking components were measured at 21 sites using the HFS. Emissions from the remaining leaking sites were not quantified due to limited site access and/or the inaccessible location of the leaking components within the facility.

22. Line 310: delete "at"

Response. Done.

23. Line 312-314: provide a definition of "non-detects". Are these emissions detected by the OGI but not GPM, TDM, OTM, or UAV or vice versa? I think it's important to have one line here in the main text even if it's provided in the SI.

Response. Non-detects are measurements that show either no measurable enhancement or yield an emission rate estimate that is below the nominal detection limit of the respective method. A short definition of the "non-detects" was already provided in lines 292-294: "*The implementation of the log-normal fits requires information about the detection limit of each method and the number of sites with emissions below this value (referred to as non-detects)* ". Both component (OGI and HFS) and facility scale (GPM, TDM, OTM, MBA) quantification methods have a detection limit. As a result, they are incapable of detecting emissions that fall below this specified detection limit. As part of their standard reporting procedure, the OGI team provided information regarding the fraction of sites where they did not detect emissions. However, it is important to note that since the majority of these surveys were performed from the fence line, they have the potential to miss sources of emissions and may not capture emissions from all possible locations within the site.

24. Table 1, Figure 2: Since these data do not represent the same sites, the differences may not be due to the method but due to the difference in emissions at sites. Plus, even at the same site, emission rates can change over time. There should be some description of the sites that are included in each box plot.

Response. We agree that the differences in Figure 2 can be due to the difference in the magnitude and variability of emissions at each individual oil production site—this is why the focus of our analysis is to bring together the different datasets rather than explain differences between methods. The purpose of Table 1, Figure 2, and section 3.1 is not to directly compare the results from individual sites for each measurement method. Instead, the focus is on providing a concise overview of the emission distributions across populations of oil production sites and datasets used in our analysis and to mainly highlight that emission rates were positively skewed with a heavy tail for all measurement methods. To avoid confusion, we include a sentence clarifying the potential reasons for the differences.

Lines 344-347: *"It is important to note that the sampled oil production sites are different for each method (and sampled at different points in time), thus Figure 2 summarizes the sampled emissions distributions and the observed differences in Figure 2 may be influenced by factors such as variations in emissions magnitude and variability at each specific oil production site."*

25. Line 386: remove extra "."

Response. Done.

26. Line 399: inconsistent spacing between number of %.

Response. Done.

27. Line 401: change to "35-83% higher"

Response. Done.

28. Line 402: what are the ranges of uncertainty based on?

Response. The ranges of uncertainty are based on the 95% Confidence Interval that is calculated for each estimated emission factor.

29. Line 412: it seems as though "wells" and "sites" are being used interchangeably. On a given well site, there are many components, including the well but also other infrastructure such as tanks. Provide a definition of both and stick with one term (e.g., "well site"). In addition, well sites can vary substantially in size and complexity. For example, Tyner and Johnson (2021) have a nice classification of well sites as on-site and off-site (https://doi.org/10.1021/acs.est.1c01572).

Response. Thank you for this remark. We incorporate the term "oil production sites" throughout the paper. We now also provide additional details regarding the oil production sites included in the analysis.

Lines 150-151: *"The oil production sites included in the study were usually relatively simple, consisting of pump jacks and additional production equipment."*

30. Line 422: what are "open-ended lines"? Since they represent such a large fraction, it is important to describe what they are.

Response. The definition of an open-ended line is now included in the text.

Lines 423-425: "*An open-ended line refers to a pipe or tubing that is not sealed at one end, and therefore remains open to the atmosphere, allowing all gas to be vented to the atmosphere".*

31. Line 438: provide definition of an "oil park".

Response. The definition of an oil park is now included in the text.

Lines 439-440: "An oil park is a facility designed to gather, store, and distribute oil produced from multiple individual wells in the surrounding area. "

32. Line 443: replace "with" with "by"

Response. Done.

33. Line 454: add "Methane" in front of "Emissions".

Response. Done.

34. Line 485: replace with "we now confirm this discrepancy is large for Romania"

Response. Done.

35. Line 492: this sentence suggests that gas flaring was practiced in Romania before. Why was it stopped? Or is this a typo?

Response. Gas flaring was not widely practiced in Romania in the past, but there has been a reduction in emissions from gas flaring the past years. Based on the emissions reported to the UNFCCC in category 1.B.2.c (sub-category *ii: flaring*), CH_4 emissions from gas flaring were 0.3 kt for the base year 1989, whereas for the latest inventory year 2020, CH_4 emissions were decreased to 0.09 kt. However, we do not know the exact reasons for this reduction.

36. Line 503-509: it is important to clearly define what is meant by site to make these comparisons.

Response. In response to a previous comment, we have included the term "oil production sites" consistently throughout our paper to refer to the specific type of sites analyzed in our study. Therefore, we compare our results to other studies that have focused on similar oil and gas production sites.

Lines 503-506: *"EFs estimated for the studies used for our comparison range between 1.2 and 8.2 kg h^{-1} site^{-1} for O&G production sites (e.g., oil well and/or gas well sites), with the majority of the EFs being below 3 kg h^{-1} site^{-1} (see Table S13)."*

37. Line 511: I would just say third or fourth highest. Clearly the Barnett and Fayetteville (where there's more unconventional development), the distributions are more heavy tailed. Could we group the regions by production type or some other factor? It may be that Romania has the highest level of skewness among conventional production regions.

Response. Indeed, looking at Figure 7, Romania presents the fourth highest level of skewness. We now provide this ranking in our manuscript (line 511). As suggested, when categorizing regions based on production type, Romania could potentially exhibit the highest level of skewness among conventional production regions. Nonetheless, our objective in this section of the paper is to emphasize that the Southern Romanian production region exhibits a heavy-tailed distribution and a level of skewness comparable to findings from other studies in regions across US and Canada. Therefore, rather than making a direct comparison, our intention is to show that Romania shares similar characteristics with those regions.

38. Line 523-526: provide more descriptions of the open-ended lines and what the non-identified components below the ground may be.

Response. As suggested in a previous comment, a definition of the open-ended lines in now provided in lines 423-425. To maintain a coherent flow in our paper, we include a short description of the possible non-identified components below the ground in the same paragraph as the definition of the open-ended lines.

Lines 427-429: *"It should be noted that the inaccessible and, as a result, non-identified components below the ground may consist of valves, pumps, connectors, or potentially open-ended lines".*

39. Line 527-535: It seems that this is an area for further research. One study based in Canada found that H2S emitting wells emit more methane: https://www.sciencedirect.com/science/article/abs/pii/S0048969722005836

Response. We thank the reviewer for the remark. It is important to note that the investigated wells emitting both H_2S and high levels of CH_4 in the mentioned paper primarily consist of two abandoned wells and one potentially undocumented well (El Hachem and Kang 2022). In contrast, the oil production sites with associated H_2S emissions examined in our study are active and operational where H_2S levels are most likely monitored, and safety protocols are in place in comparison to abandoned wells where maintenance or monitoring is absent. This aligns with the findings of the suggested paper in the following comment which indicate that reduction strategies focusing on olfactory compounds in Peace River have proven beneficial in reducing and maintaining lower CH_4 emissions, despite not being specifically designed for CH_4 reduction purposes (Lavoie et al. 2022). However, we do agree that this is an area for further research and that, to our knowledge, published literature lacks measurements of H_2S emissions from O&G production sites and no clear relationship between CH_4 and H_2S emission rates has been established. We now include the following information in the main text.

Lines 536-541: "These findings are consistent with the research conducted by Lavoie et al. (2022), which showed that reduction strategies focusing on olfactory compounds in Peace River have proven beneficial in reducing and maintaining lower CH4 emissions, despite not being specifically designed for CH4 reduction purposes (Lavoie et al. 2022). However, it is important to note that further research is needed to establish a clear relationship between CH4 and H2S emission rates".

40. However, another study found something similar to what was found here: https://www.sciencedirect.com/science/article/abs/pii/S0048969721059143

Response. See previous comment.

41. Line 545-549: Would it be safe to say that zero gas production wells are "abandoned wells"? How do the emission rates compare with those for abandoned wells in published literature? https://pubs.acs.org/doi/abs/10.1021/acs.est.0c04265

Response. Thank you for the question and the suggested paper. The oil production sites analyzed in our work are not abandoned wells as investigated in (Williams, Regehr, and Kang 2021). They are still operating and producing oil, but without collecting associated gas. The associated natural gas is considered a byproduct and might be deemed uneconomical to collect and sell. The vented gas is not metered and therefore zero gas production is reported from these sites.

42. Line 566: "practices"

Response. Done.

43. Line 575: as noted earlier, with the mention of "ground-based quantification methods", I expect some discussion or comments about the role of non-ground-based quantification methods. Also, UAVs are included in this analysis, which I would consider as aerial-based quantification methods. So, there's a need to change the terminology and discuss the limitations and benefits of "ground-based quantification methods" compared to others.

Response. As mentioned in previous comments, to avoid confusion, we now use the term "ground and drone-based" measurements to refer to the data analysed in this manuscript. We are currently finishing the data analysis of the aircraft measurements and in a future publication, the synthesis of the results from the ground, drone and aircraft measurements will be discussed. Therefore, we would prefer to include the discussion concerning the role of ground and aerial-based quantification methods in the upcoming paper.

44. Line 577: replace with "oil well sites". I suggest using this term throughout the paper.

Response. Thank you for your suggestion. To enhance clarity and eliminate any potential confusion, we have decided to consistently incorporate the term "oil production sites" throughout the paper.

45. Line 582-585: This point may not be fully supported by literature. The role of H2S should be studied more systematically.

Response. See previous comments.

46. Line 584: H2S emissions is not a component. replace "this component" with "H2S emissions".

Response. Done.

DETAILED COMMENTS ON SUPPLEMENTAL INFORMATION

47. Line 82: What about components other than the well borehole that may be leaking methane? It is discussed later on but an introductory paragraph on the range of potentially leaking components is needed.

Response. The main text already incorporates a paragraph discussing individual leaking components from oil production sites in section 3.3. As mentioned in a previous comment, the oil production sites included in the study were usually relatively simple, consisting of pump jacks and additional production equipment. It can be, therefore, inferred that most emission points for such oil production sites are situated near the pump jack and well borehole. This is further supported by the infrared videos captured, where emissions were observed in close proximity to the pump jack. Since the TDM team conducted facility scale measurements, knowledge of the exact type of leaking component is not necessary. This is the case for the rest of the measurement methods described in section S2 which only focus on the facility scale.

48. Line 145 and section: What are the pros and cons of the UG and EMPA methods? It seems that the biggest difference is how wind is taken into account.

Response. Due to their adaptability and flexibility, the two UAV-based methods have the advantage of being able to sample at locations that are hard to reach for traditional measurement methods, allowing quick adaptation of the measurement strategy according to changing wind conditions (Andersen et al. 2021; Morales et al. 2022). This possibility of rapid adaptation to changing wind conditions is highly valuable, especially when a large number of sources have to be quantified in a short amount of time, as was the case during the ROMEO measurement campaign. High sensitivity to wind conditions and a relatively short flight time can be seen as a slight drawback of the UAV-based quantifications since the individual flight represents a brief snapshot of the plume (Morales et al., 2022).

Indeed, the difference between these two UAV-based MB approaches is in how the wind is being taken into account; EMPA projects the normal wind component to each measured methane mole fraction, while UG applies the mean wind to all sampled methane mole fractions, as it has been previously explained in comment 19 for line 240.

The above information has been included in Section S2 of the supplementary material.

Line 147-155: "Two different UAV-based systems using a Mass Balance Approach (MBA) were used to quantify the emission rates from the surveyed oil and gas facilities. Due to their adaptability and flexibility, the two UAV-based methods have the advantage of being able to sample at locations that are hard to reach for traditional measurement methods, allowing quick adaptation of the measurement strategy according to changing wind conditions (Andersen et al. 2021; Morales et al. 2022). This possibility of rapid adaptation to changing wind conditions is highly valuable, especially when a large number of sources have to be quantified in a short amount of time, as was the case during the ROMEO measurement campaign. High sensitivity to wind conditions and a relatively short flight time can be seen as a slight drawback of the UAV-based quantifications since the individual flight represents a brief snapshot of the plume (Morales et al., 2022). Here we describe the differences in the MBA between the active AirCore system from the University of Groningen (UG) and the Quantum Cascade Laser Absorption Spectrometer (QCLAS) from the Swiss Federal Institute for Materials Science and Technology (EMPA)."

49. Line 161: How is the wind speed and direction determined? Is it measured directly in the field and if so, is it measured using the UAV or something else?

Response. Both UG and EMPA measured local meteorological conditions using the 3D sonic anemometer placed in the near vicinity of the source at heights of ~ 3 m (UG) and ~ 5 m (EMPA), with a sampling frequency of 10 Hz (UG) and 20 Hz (EMPA). This information has been incorporated in section S2, lines 180-184 of the supplementary material.

50. Line 173: How is the horizontal vector, u(y,z), determined?

Response. The horizontal wind vector u(y,z), referred to as projected wind in Morales et al. (2022), involved taking the 1 s average normal wind component and projecting it onto the measurement plane by matching the timestamp of the anemometer to the GPS location of the UAV during the time of measurement. Since the anemometer was placed in close proximity to the investigated source, the wind measurements were assumed to be representative of the conditions encountered by the UAV. This information has been incorporated in section S2, lines 184-189 of the supplementary material.

51. Line 237: Reading later on, the role of Detection Limit in the likelihood function is described. However, there should be some information here or earlier on.

Response. We briefly acknowledge the importance of the Detection Limit and the number of measurements below the detection limit in this section, referencing the more comprehensive discussion provided later in sections S5 and S8.

Lines 239-241: *"The role and significance of selecting the detection limit and the impact of the number of measurements below that limit are discussed in detail in sections S5 and S8. "*

52. Line 245: Describe the direct search algorithm used here and provide a reference. Is this described in Zavala-Araiza et al (2015)? If so, it should be cited in this sentence.

Response. Yes, the entire methodology followed for the calculation of the emission factors is thoroughly explained in Zavala-Araiza et al (2015). Here we provide a short description of the statistical approach used. As suggested, we have appropriately referenced the citation in the sentence.

53. Line 313-314: Is this due to the fact that detection limit emission rate is applied to non-detects? If non-detects were set to zero, the opposite would be true.

Response. This is correct. As previously mentioned, the number of sites with emissions below the detection limit of each measurement method are referred to as non-detects. Therefore, values between 0 kg CH_4/h and the determined detection limit emission rate were randomly assigned with equal probability based on the fraction of non-detects estimated for each measurement method. However, as seen in Equation 8 of section S4 in the Supplementary Material, the parameters affecting the emission distribution are not the precise value of the assigned emission rates, but rather the fraction of non-detects and the determined detection limit (see section S8. Sensitivity analysis of the statistical estimator). If the fraction of non-detects is set to zero, the emission distribution will become narrower compared to the case where a non-zero value is assigned to non-detects. This effect is explained in detail in section S8.

54. Table S3: Provide definition of S_0 in the caption again.

Response. Done.

55. Line 357: What is the difference between east and west regions? In general, a better description of the basin and the general setting could be helpful in this paper.

Response. The basin was arbitrarily split into east and west regions, i.e. north-east and -west of the city of Bucharest, since all four quantification methods (TDM, OTM-33A, GPM, MBA) were performed in each part of the basin. The landscape in the eastern region is predominantly flat, with a steppe plain area in the middle and mountains in the north. This region is heavily industrialized and contains the main oil reserves in Romania. Apart from that, the chemical, rubber, food, and textile industries are also present. The west region is relatively flat and dominated by agriculture; alongside this, oil is also being extracted. Around the city of Craiova (Fig. 1), the largest city in southwestern Romania, automotive, heavy electrical, and transport equipment industries are located (ROMEO web 2023).

56. Line 470-472: The importance of getting the low-end distribution right is something that is not widely discussed and is worth highlighting in the main paper.

Response. Upon your suggestion, we include a few sentences in the main text discussing the sensitivity of the statistical estimator to the low end of the distribution.

Lines 319-324: *"The effect of the fraction of non-detects and the detection limit on the log-normal fits and the final EFs is further explored by testing several different values (section S5). We find that reducing the detection limit or increasing the fraction of non-detects leads to higher estimated EFs due to the widening of the distribution towards the lower end. This emphasizes the importance and need of conducting a thorough investigation when selecting the values for these two parameters."*

57. Line 569: Why were only production and age evaluated? Was it because other data were not available or is there evidence to indicate that these are the most important characteristics to look at?

Response. Many previous studies have investigated the relationship between CH_4 emissions and the age, oil and natural gas production rates, as well as the sample representativeness by comparing these site-specific parameters from the sampled facilities to the overall population (Robertson et al. 2017; 2020; Zavala-Araiza et al. 2018; Omara et al. 2022). Therefore, we decided to adopt a similar approach. In addition to considering oil and gas production and facility age, we also evaluated the spatial coverage of the study area. While detailed results of this assessment are not provided in the manuscript, Figure 1 illustrates the broad spatial coverage achieved. In general, the data obtained from the operator included information about the site type, spud date (site age), spatial coordinates, status of the site and water, oil, natural gas, and gross production of the facilities in the studied areas for the duration of the measurement campaign.

58. Line 597: Replace “for” with “of”

Response. Done.

59. Line 605-606: I would argue that what is being shown here is that emitting wells are more likely to have higher gas production and associated gas, which is consistent with the finding that oil wells are more likely to be non-emitting. Therefore, I suggest revising the sentence: “no significant differences were found between emitting and non-emitting sites”.

Response. It is true that emitting oil wells have slightly higher gas production ($9.5x10^3$ scm) and a lower fraction of wells reporting zero gas production (70%) compared to non-emitting oil well sites (values equal to $7.5x10^3$ scm and 80% respectively). We assumed that this difference was not significant, but following your suggestion we revised the sentence, and modified the paragraph. Half of the 181 IR screened oil wells had detected emissions, but we have not assessed other type of facilities with a statistically meaningful sample size, so we cannot conclude that oil wells are more likely to be non-emitting.

Lines 604-617:"Similarly, a summary of the characteristics from the IR screened oil production sites and from the total population of oil production sites in Romania is shown in Table S15. We find that emitting oil production sites have a slightly higher gas production and a lower fraction of sites reporting zero gas production compared to non-emitting oil production sites. For the gas production, approximately 70 % of emitting and 82 % of non-emitting oil production sites visited report zero gas production or had no gas production in 2019. These percentages are higher than the average percentage of the total population of oil production sites in the country. Emitting oil production sites had an average age of 36 years, average gas production of 9,500 scm per year and average oil production of 48 tons per year. We found a slightly lower range of values for non-emitting sites (see Table S15). Overall, the oil production sites visited were representative of the total population of sites in the country in terms of age. However, measurements leaned more towards the high oil, but very low gas producing end of the spectrum. "

60. Line 609: Replace “oi” with “oil”

Response. Done.

References

Allen, David T., Vincent M. Torres, James Thomas, David W. Sullivan, Matthew Harrison, Al Hendler, Scott C. Herndon, et al. 2013. “Measurements of Methane Emissions at Natural Gas Production Sites in the United States.” *Proceedings of the National Academy of Sciences* 110 (44): 17768–73. https://doi.org/10.1073/pnas.1304880110.

Andersen, Truls, Katarina Vinkovic, Marcel de Vries, Bert Kers, Jaroslaw Necki, Justyna Swolkien, Anke Roiger, Wouter Peters, and Huilin Chen. 2021. “Quantifying Methane Emissions from Coal Mining Ventilation Shafts Using an Unmanned Aerial Vehicle (UAV)-Based Active AirCore System.” *Atmospheric Environment: X* 12 (December): 100135. https://doi.org/10.1016/j.aeaoa.2021.100135.

El Hachem, Khalil, and Mary Kang. 2022. “Methane and Hydrogen Sulfide Emissions from Abandoned, Active, and Marginally Producing Oil and Gas Wells in Ontario, Canada.” *Science of The Total Environment* 823 (June): 153491. https://doi.org/10.1016/j.scitotenv.2022.153491.

Lavoie, Martin, Jennifer Baillie, Evelise Bourlon, Elizabeth O’Connell, Katlyn MacKay, Ian Boelens, and David Risk. 2022. “Sweet and Sour: A Quantitative Analysis of Methane Emissions in Contrasting Alberta, Canada, Heavy Oil Developments.” *Science of The Total Environment* 807 (February): 150836. https://doi.org/10.1016/j.scitotenv.2021.150836.

Morales, Randulph, Jonas Ravelid, Katarina Vinkovic, Piotr Korbeń, Béla Tuzson, Lukas Emmenegger, Huilin Chen, Martina Schmidt, Sebastian Humbel, and Dominik Brunner. 2022. “Controlled-Release Experiment to Investigate Uncertainties in UAV-Based Emission Quantification for Methane Point Sources.” *Atmospheric Measurement Techniques* 15 (7): 2177–98. https://doi.org/10.5194/amt-15-2177-2022.

Omara, Mark, Daniel Zavala-Araiza, David R. Lyon, Benjamin Hmiel, Katherine A. Roberts, and Steven P. Hamburg. 2022. “Methane Emissions from US Low Production Oil and Natural Gas Well Sites.” *Nature Communications* 13 (1): 2085. https://doi.org/10.1038/s41467-022-29709-3.

Riddick, Stuart N., Denise L. Mauzerall, Michael A. Celia, Mary Kang, Kara Bressler, Christopher Chu, and Caleb D. Gum. 2019. “Measuring Methane Emissions from Abandoned and Active Oil and Gas Wells in West Virginia.” *Science of The Total Environment* 651 (February): 1849–56. https://doi.org/10.1016/j.scitotenv.2018.10.082.

Robertson, Anna M., Rachel Edie, Robert A. Field, David Lyon, Renee McVay, Mark Omara, Daniel Zavala-Araiza, and Shane M. Murphy. 2020. “New Mexico Permian Basin Measured Well Pad Methane Emissions Are a Factor of 5–9 Times Higher Than U.S. EPA Estimates.” *Environmental Science & Technology* 54 (21): 13926–34. https://doi.org/10.1021/acs.est.0c02927.

Robertson, Anna M., Rachel Edie, Dustin Snare, Jeffrey Soltis, Robert A. Field, Matthew D. Burkhart, Clay S. Bell, Daniel Zimmerle, and Shane M. Murphy. 2017. “Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions.” *Environmental Science & Technology* 51 (15): 8832–40. https://doi.org/10.1021/acs.est.7b00571.

ROMEO web. 2023. “Overview - ROMEO Measurement Campaign.” 2023. http://romeo-memo2.wikidot.com/overview.

Vinković, Katarina, Truls Andersen, Marcel de Vries, Bert Kers, Steven van Heuven, Wouter Peters, Arjan Hensen, Pim van den Bulk, and Huilin Chen. 2022. “Evaluating the Use of an Unmanned Aerial Vehicle (UAV)-Based Active AirCore System to Quantify Methane Emissions from Dairy Cows.” *Science of The Total Environment* 831 (July): 154898. https://doi.org/10.1016/j.scitotenv.2022.154898.

Williams, James P., Amara Regehr, and Mary Kang. 2021. “Methane Emissions from Abandoned Oil and Gas Wells in Canada and the United States.” *Environmental Science & Technology* 55 (1): 563–70. https://doi.org/10.1021/acs.est.0c04265.

Zavala-Araiza, Daniel, Scott C. Herndon, Joseph R. Roscioli, Tara I. Yacovitch, Matthew R. Johnson, David R. Tyner, Mark Omara, and Berk Knighton. 2018. “Methane Emissions from Oil and Gas Production Sites in Alberta, Canada.” Edited by Detlev Helmig and Stefan Schwietzke. *Elementa: Science of the Anthropocene* 6 (March): 27. https://doi.org/10.1525/elementa.284.

Response to anonymous referee #2 comments

We thank the anonymous referee for the valuable feedback and comments, which have helped to improve our manuscript. Responses to individual referee comments are below. This document is organized as follows: the Referee's comments are in plain text, our responses are in blue, and all the revisions in the manuscript are shown in blue *italic*. Line numbers refer to the updated WORD manuscript with tracked changes.

Anonymous Referee #2

general comments

The work reported in this paper is from a co-ordinated field campaign conducted between many groups, measuring CH4 from oil production sites across Romania using four different approaches. The scientific question is firmly within ACP's remit, and contains well established methods and statistical analyses to interrogate the results. The report is laid out clearly, with assumptions and methods presented well. Much of the detail of the methodologies is in the SI. I note that another reviewer has recommended to bring some of this into the main text. I would support this, although I think this is somewhat down to personal preference, as I did not see any guidance on this from the journal itself.

Substantial conclusions are reached, showing that methane emissions from these sources are over twice as high as currently reported, thus highlighting the inadequacy of the methods uses to report emissions in this sector and the potential for mitigation by stopping the leaks.

We express our gratitude to the reviewer for acknowledging the important findings presented in our work and for providing positive feedback. Based on the recommendations from both reviewers, we have revised our manuscript and we have incorporated additional details and information from the supplementary material into the main text.

specific comments

1. L422 and elsewhere: open-ended lines is a term which is used, and is somewhat self explanatory, however it would help to define it for those not within this field. e.g. it is not clear whether these lines are designed like this, or if they should have a cap on the end but for some reason the cap has been left off.

Response. The definition of an open-ended line is now included in the text.

Lines 423-425: "*An open-ended line refers to a pipe or tubing that is not sealed at one end, and therefore remains open to the atmosphere, allowing all gas to be vented to the atmosphere".*

2. L554 and 577: what are the Tier 1 EFs for developing and developed countries? Presumably the currently used EFs for reporting are quite different to what you have derived. This value would be useful to compare with the EF you have derived, eg in the conclusion as a comparison (and earlier when you discuss the IPCC methodology).

   Response. Based on the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, the TIER 1 EF for fugitive emissions from the onshore conventional oil production are $1.5x10^{-6}$ to $3.6x10^{-3}$ and $1.5x10^{-6}$ to $6.0x10^{-2}$ Gg per 10^3 m^3 conventional oil production for developed and developing countries, respectively. However, our estimates only cover emissions from oil production sites and these IPCC factors include emissions from the entire upstream segment thus a direct comparison is not feasible. Therefore, to avoid confusion, we have chosen not to include these specific numerical figures in our discussion.

technical corrections

3. L67 and later: % symbol should not have a space before it, i.e. it should be 25% and not 25 %

   Response. This is now corrected throughout the text.

4. L75 L80: Is Global Methane Tracker 2022, 2022 correctly referenced? Just checking, as it looks odd with 2022 appearing twice

   Response. Thank you for pointing this out. We have updated the reference.