

Assessing Urban Methane Emissions using Column Observing Portable FTIR Spectrometers and a Novel Bayesian Inversion Framework - Responses to Review Comments

Notes and questions from reviewers are reproduced here in black. Responses to reviewers are written in red, with relevant quotes from the new version of the manuscript written in red italics.

1 Reviewer 1

Jones et al. present an inverse modeling study on methane emissions from the city of Indianapolis. While a number of studies
5 have been published to address the same problem, the study is unique and novel by using data from five portable solar-tracking
Fourier transform infrared (FTIR) spectrometers and an inversion method devised to comprehensively account for uncertainties,
especially those uncertainties that are not often adequately addressed in the literature, namely, uncertainties of the background
methane concentration, and uncertainties in the spatial pattern of the prior inventory. As such, this study is positioned to
complement previous efforts in painting a fuller picture on the issue and to help point to future directions of research. The
10 paper is well written, including key information regarding the inversion method and a concise presentation of the key results in
figures and tables. I recommend publication of the paper, and provide the following questions and comments for the authors to
consider. I realize addressing the comments below might require some additional model experiments. I encourage the authors
to do what they deem as appropriate with available resources and time, as I believe a better understanding and discussion of
these issues would further strengthen the paper.

15 I think the section on the results (Section 3) should be expanded. In its current form, with text of less than 30 lines, this section
does not go much further from summarizing the numbers from the figures and tables and comparison with previous estimates.
I suggest the authors to include more discussions, highlighting the unique findings from the study with previous studies as
the context, and offer some experiences and suggestions for future work based on the results. Here are some comments and
questions for the authors to consider:

20

*Thank you very much for appreciating the novelty of our paper. We agree that both the results and conclusions sections
should have been expanded, and both are now significantly longer.*

The authors offer an explicit approach to calculating the background concentration with STILT (Eq. 5), which to me is
25 necessary and a preferred approach in future studies. I suggest the authors further explore and demonstrate the benefit of using

this approach, e.g., by doing another inversion that does not treat the background with a background influence matrix. How would the emission estimates change?

A sensitivity analysis of the affect of changing the background parameters is now included in the Supporting Information (S5: Background Parameter Sensitivity Analysis). Although we did not test eliminating the BIM totally, we did test using very low background prior co-variance values, which removes the model's ability to adjust the background values and keeps them fixed. The results of this sensitivity analysis show that allowing the model to adjust the background does have an important effect on both the posterior emissions number and its uncertainty, but that as long as the order of magnitude of these parameters is reasonable the solution is stable and not overly sensitive.

35

One intriguing finding from the study is that two prior spatial patterns have led to drastically different results. This finding should be further explored and discussed. The study uses a small state vector with just a few terms as scaling factors for the emissions from different sectors; this approach implies considerable confidence/weight on the prior spatial distribution in fitting to the observations. An alternative approach that has been used previously would be to solve for the spatial pattern with a bigger state vector that leaves more degrees of freedom on top of the prior spatial distributions, i.e., scaling factors of emissions from different parts of the domain, for example, by dividing the domain into a number of squares a few kilometers in dimension. This approach would mean the data can have more freedom in informing the spatial pattern of the posterior. What are the considerations being such a choice of state vector? Is it because of the lack of data points from FTIR, or is it because the results make more sense this way? Have you tried, and if not, is it feasible to try another inversion with a different configuration of the state vector?

45

Yes, we did attempt to spatially resolve the emissions, but it became clear early on in our analysis that we had nowhere near enough data to do this properly. This is an important concept though, and was worth putting in the conclusions section, so the following sentences were added:

Another improvement would be to solve for different scaling factors in different parts of the city. Theoretically, the inversion can solve for each grid square independently, with correlations between squares expressed in the co-variance matrices. However, in practice this requires significantly more observations, because the results become less constrained as the state vector increases in length. We believe that our framework could easily be expanded to spatially resolve emissions if implemented with data from a permanent or semi-permanent installation of sensors, such as MUCCnet (Dietrich et al, 2020)

55

The results presented in the figures and tables need some discussions and explanations. For instance, Figure 15 shows large day-to-day variations of diffuse methane emissions of about a factor of 5 (if taking the minimum of 50 mol s⁻¹ on 5/13 and maximum of 250 mol s⁻¹ on 5/22). Can you explain the reasons behind this change? Figure 15 apparently shows that using roads as the spatial pattern for the prior, an inversion using all data gives a higher emission estimate than estimates on all individual days. Why is this?

60

These are both good observations that deserved more discussion in the conclusions section. Thank you for pointing these out. Two paragraphs have been added to address each of these points:

65 *Variations in the inversions results between each day are quite large, because these days have significantly different wind patterns, which means that different sections of the city are sampled on each day. Therefore, the calculated scaling factor of each day represents only the emissions of the parts of the city covered by the wind, resulting in large variations of the total daily emissions of the city. By ingesting large amounts of data over a long time frame, covering many wind patterns, this issue can be resolved. Ideally, much more data than the five days used in this experimental study would be used.*

70 *When an inversion is performed using the road-based distribution and all of the days of data, the estimate of emissions is much higher than any of the individual days. This is because the road-based distribution pulls emissions from the center of the city and puts them in the outer ring-roads, which have significantly less total footprint. This means that the product of the footprints and emissions grid is relatively low, and therefore the model gives more weight to the prior. Only when all of the days are used there are enough data to influence the model significantly. Because the road-based distribution has relatively low*
75 *prior emissions in the city center, the result is a large scaling factor. This also highlights the importance of sensor placement. Sensors should be placed close to the bulk of the suspected emissions sources so that strong influence is detected in a variety of wind conditions. If the experiment was designed with the road-based distribution in mind, sensors would be placed further apart in order to adequately sample the ring roads. We would also suggest that more sensors would be needed in this situation, as a greater portion of emissions would be spaced out over a larger area.*

80

The data used by the study, i.e., FTIR column data, is somewhat unique by its own right. It would be very beneficial for future studies if the authors can, based on their experiences, summarize and discuss a few key considerations future studies should bear in mind when using this type of data. For example, what are the favorable situations to use these data, what supplemental data are needed, and what would be the minimum number of sensors needed.

85

This is a good idea, and we have added a paragraph to the conclusion section to discuss some of these points:

Future studies utilizing these types of instruments should also explore the trade-offs involved when choosing the number of sensors needed and their placement. While there are some studies exploring this concept using similar footprint methods, they have mostly been done for in situ sensor networks, and total-column sensors require a different set of considerations. For
90 *instance, the footprints of total-column sensors are larger, which means fewer sensors are needed to sample the same area. In addition, while portable FTIR systems may not be cheaper than many in situ sensors, they do not require the presence of tall buildings or investment in tower infrastructure in order to be deployed, which makes finding suitable deployment sites easier. Data collection with total-column systems is also dependent on direct sunlight, so any study exploring the utility of these types of networks in a specific region should also consider weather, solar angles, and shading.*

95

Technical errors:

Line 54: “We focus of” should be “we focus on”

Thank you. This has been corrected.

100 Line 55: “have proven” should be “have been proven”

This has been corrected.

2 Reviewer 2

2.1 Summary

The paper presents quantification of methane emissions using a network of five portable solar-tracking Fourier transform
105 infrared (FTIR) spectrometers during a field campaign at the city of Indianapolis, USA in May of 2016. Methane emissions are
estimated using a combination of Lagrangian transport model with a Bayesian inversion framework. The approach estimated
both, surface emissions and background methane concentrations flowing into the city. Diffuse emissions, presumably leaks
from the natural gas infrastructure, were found to be 73 ± 22 mol/s, 68% higher than estimated from bottom-up methods.

2.2 General Comments

110 The paper performs a valuable experiment by trying to use only a several days of total column methane measurements from
EM27/SUN solar-tracking total column Fourier transform infrared (FTIR) spectrometers to estimate methane emissions at
Indianapolis. In comparison to the complexities of aircraft flights or of setting up long-term tower measurements, EM27/SUN
instruments are relatively easy to move around and in principle could be used at multiple urban centers over the span of several
months, which could significantly help with the goal of quick greenhouse gas emission estimation required by policymakers to
115 efficiently address the issue of climate change. A novel Bayesian inversion framework is advocated in synergy with EM27/SUN
instruments to perform the mentioned methane emission estimation. In this regard, the paper is appropriate for the journal and
helps to advance the field of carbon cycle. However, despite these positive attributes, there are a number of critical omissions
in the description of the work’s methodology that need to be fixed before this paper can be published. Additionally, the article
lacks interpretation and proper discussion of the results.

120

Thank you for pointing out the positive attributes of our paper. We have expanded the description of the work’s methodology
and added discussion of the results in the results and conclusions sections.

2.3 Specific Comments

125 Lines 174-175: If there are no particles released above 2500 m, is there an assumption that layers above are homogeneous or somehow not important to the total column methane measurements? What about potential transport of significant methane plumes in the mid-troposphere?

130 Yes, this is a potential issue. In our study, we found that the model was able to resolve the major background plumes that were obvious in the data, but if there is a wind shear above the 2500 meter level, and a large plume above that, the model could run into some difficulty. We have added a paragraph to the methods section to make the readers aware of this caveat:

The particle releases used to create the BIM are the same as those used to create the footprints described earlier, and so are from 14 different altitudes up to 2500 meters. Transport above this level is not modeled, which means there is a possibility that plumes present in the mid-troposphere will not be accounted for effectively.

135 Line 186: Variable s and indices i and j are not clearly defined. Please define.

This paragraph has been edited for clarity. It now reads:

140 *The column footprints are then multiplied by the a priori emissions field $F(x, y, s)$ and summed to produce a row n in the matrix \mathbf{A} at the time step m . F is a function of latitude, x , longitude y , and emissions sector, s , but is not temporally resolved. The resulting matrix \mathbf{A} will have a number of columns equal to the number of observation time points being simulated and a number of rows equal to the number of emission sectors in the inventory.*

145 Lines 205-215: the units of vector \mathbf{b} and matrix \mathbf{B} are not properly defined. Text indicates that units of vector \mathbf{b} are ppb, while units of \mathbf{B} are hr^{-1} (according to Figure 5). Given such units, equation 5 does not work (since all the sums must maintain units of ppb). It is possible that authors treat \mathbf{B} as unitless, a fraction of particles affected by the defined edges. But this needs to be clearly stated. Units and variables must always be defined in a proper mathematical notation.

150 Thank you for pointing out this inconsistency. The matrix \mathbf{B} is indeed unitless, but the values were divided by the background time step length of 15 minutes for Figure 5 to make the PDFs more intuitive. Language has been added to clarify this:

155 *Here \mathbf{B} is the background influence matrix (BIM) which relates every observation to a distribution of influence from each background time point. Matrix \mathbf{B} is unitless, as it merely transforms values from the boundary background (\mathbf{b} in ppb) to background values at each observation point (in ppb). An additional error term, ϵ_b , represents error (in ppb) due to uncertainties in the BIM.*

Also, the caption to Figure 5 has been expanded:

Although entries in the BIM are unitless, here they are shown divided by the background time step. These are therefore residence time probability density functions, with the integrals over the graphs each equal to 1.

160 Line 213: Please explain how the error term, ϵ_b is calculated?

The two error terms are not known. A sentence was added to clarify this at this point:

It is important to note that because the values of the error terms ϵ_a and ϵ_b are unknown, and can be quite large, a careful statistical inversion must be performed to retrieve valid estimates of emissions.

165

Line 221: The definition implies that you are solving for matrix x , that is for an emission scaling factor as (unitless) and background time series b (ppb), is that correct?

Yes, that is correct. A sentence was added for clarification:

170 *The emissions scaling factors and domain background are then combined into a single state vector (x).*

Line 224: Background error was not described previously (please see Line 213 comment).

The previous paragraph has been amended to explain the background error more thoroughly.

175

Line 273: Part of a sentence is missing, please fix.

This sentence has been fixed. It now reads:

180 *The prior state vector (x_a) contains initial estimates of the scaling factors that for the emissions from each sector, as well as initial estimates of the background methane concentration at every relevant point in time.*

Equation 10: if a priori scaling factors are all 1 and a prior background mixing ratio is 1.84 ppm, what is the point of the following conditions: $1 \leq i \leq n_{sec}$ and otherwise? Define i .

185 The index term i has now been added to the left-hand side of the piece-wise equation in the correct place. Equation 10 is now annotated to explain each case.

Equation 11: the conditions (such as $i, j \leq n_{sec}$ and $i = j$) are not explained.

190 Equation 11 is now annotated to explain each case.

Line 284: It is not clear what is meant by observations i and j . Are these the same as in section 2.4?

These indices are in reference to Equation 11, which is now referenced in this sentence.

195

Lines 301-302: On these lines there is a following statement, "...the diagonals of which describe the model framework's ability to reduce the uncertainties of the priors..." Please explain what do you mean by an ability to reduce the uncertainties of the priors? How an uncertainty could be reduced given that it is determined subjectively (as described in the methodology section)?

200

This is a good point. Although we believe our prior uncertainty estimates are reasonable, there are arguments to be made for selecting different values. Different choices for these parameters would affect our result, and so a sentence has been added here to acknowledge that:

205 *"It should be noted that this posterior will likely be dependent on the prior uncertainties chosen, and that a rigorous inversion should have reasonable prior estimates of error."*

Lines 315-316: Although the diffuse emission value makes sense with all the days combined, individually the emissions look unrealistic as they are shown in Table 4. How can emissions jump from 40.4 mol/s to 146.3 mol/s in 2 days? And what is the meaning of negative emissions? It is not clear why the uncertainty is only 22 mol/s given a sample of only 5
210 days (each with significantly different values). Physically this is highly improbable. As an experimental product this result may be of an interest, but it should not be interpreted as an actual value of diffuse methane emissions for Indianapolis. As a cross check, it may be a good idea to take data from INFLUX towers and perform an inversion for the same days to see if the emission results will be comparable. Also, aircraft can be used to perform mass balance when available. Now, it is understandable that these comparisons may be outside of the scope of this work, but then the presented results must note that
215 they are experimental in nature and are not to be interpreted as actual natural gas emissions from Indianapolis. Please address this in the results/discussion section.

We agree, and have added paragraphs to both the results and discussion sections to highlight this point. Results:

220 *Although we are confident that the model framework was able to detect meaningful methane fluxes using all five days of data, the daily inversions yield emission numbers that vary greatly and are sometimes non-physical, such as the negative emissions reported on May 18. This shows that a single day of measurements is unlikely to produce a robust result, and even the combination of all five days should only be viewed as an experimental proof-of-concept, and continuous measurements for many months should be used to produce a more compelling final emission number.*

Conclusion:

225 *Although our measurement and data analysis techniques are experimental and newer than the other methods shown in Figure 16, we believe that our result supports the findings shown in these other studies. Since all of these methods have different*

drawbacks and make different sets of assumptions, any attempt to truly quantify emissions from such a complex source should consider as many independent methods as possible.

230 Figure 11: What part of edge do dashed optimized background values represent? Is that combined optimization of background from all the instruments?

The dashed line represents the retrieved background at the domain boundary using measurements at all stations (**b**). The caption has been changed to clarify this.

235 *Optimized domain boundary background values (gray, dashed) and optimized total column values, including urban emissions (colored lines) also shown.*

Figure 13: This figure seems cluttered, is it possible to present these data in the time series format? Also, what is the significance of the selected bins?

240

The data were binned in order to test the significance of the correlation. We agree that the plot is cluttered, but believe it is important to show the data in this form because it represents small, noisy, but detectable signal that the model is able to discern. Additional language explaining the point of binning the data was added to the Results section :

245 *In order to test the significance of the correlation between observed and modelled enhancements, points were binned into 0.5 ppb wide bins based on model enhancement values. The mean observed enhancement values for the 0-0.5 ppb and 1-1.5 ppb bins were then tested to see if they were significantly distinct, resulting in a p-value of 0.003. This result suggests that even with such a small signal level, and in the presence of large background concentration variations, the model was able to detect and quantify enhancements in methane due to urban emissions.*

250 Figure 14: It is not exactly clear how the model decides whether to adjust diffuse emissions or background mixing ratios (the choice seems subjective as it should depend on the prescribed a priori error). It would be great to have a sensitivity analysis where it could be clearly shown what happens when the errors of both, diffuse emissions and background mixing ratios, are allowed to vary to see if there is any stability in the shown solution.

255 A sensitivity analysis of the background parameters, including a figure and a paragraph of explanation, has been added to the Supporting Information (S5: Background Parameter Sensitivity Analysis). The results of this analysis show that the inclusion of these parameters, and the order of magnitude of their values, do play an important role in shaping the output of the inversion, but the model is not overly sensitive to the exact numbers chosen, and the result is stable.

260 Lines 326-331: Varying the prior of the spatial distribution of emissions significantly changes the result of the total diffuse emission estimation. That in itself shows that each posterior result should be approached with caution.

We agree, and have added language to the discussion to emphasise this point:

265 *It is important to note that the inversion framework described here does not have a mechanism for incorporating spatial
uncertainty and cannot redistribute emissions around the domain. Testing multiple plausible distributions gives a sense of the
sensitivity of the model, and we suggest that any comprehensive analysis include model inversions using several distributions.*

Figure 16: Please define on one of the maps what is precisely meant by “Marion County”.

270 *The outline of Marion County has now been added to Figure 1.*