Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-741-AC1, 2018

Atmospheric Chemistry and Physics
Discussions

ACPD

Interactive comment

Interactive comment on "Detecting high-emitting methane sources in oil/gas fields using satellite observations" *by* Daniel H. Cusworth et al.

Daniel H. Cusworth et al.

dcusworth@fas.harvard.edu

Received and published: 16 October 2018

Response to Comments from Anonymous Referee #1

1. No data and code that can reproduce the given results are uploaded. Doing so would increase the value of the article and invite other research groups to contribute to the topic.

We thank the reviewer for this suggestion. A worked through script will be included with the publication. However, we will leave it to the reader to generate WRF-STILT footprints as the model is open source.

We point the reader to the location of the code on page 12, line 8:

Data Availability. The WRF-STILT model is available for download at https://uataq.github.io/stilt/. A worked through example of the high-mode detection observing system simulation experiment (OSSE) described in this paper is available in the Supplementary Information for this paper.

2. Most of the article uses international standard units. However, the production rate of wells is described in Mcf/d (in text, e.g. p. 3, ll. 12-13, and Fig. 2 and 3). I suggest to convert these to SI units.

We add the unit conversion in the text:

Page 3, Line 11-12. (small: 10-100 million cubic feet per day (Mcf/d) where 1 Mcf/d = 0.028 Mm3/d; medium: 100-1000 Mcf/d; large: 1000+ Mcf/d)

3. The authors use 'a x b' to denote a scalar multiplication in some formulas (e.g., p. 6, l. 28; p. 7, l. 5; p. 8, l. 14), but not consistently. I recommend using 'a " b' or 'ab' to be consistent with standard notation. To describe the dimensions of a matrix, m x n is the standard notation.

We update the equations to be consistent with standard notation:

Page 7, Line 16. r_ij= σ_Mˆ2 exp{-d/l}exp{-t/τ} for i $\neq$j (5b)

Page 7, Line 25. r_ij= cor(i,j) σ_Mi σ_Mj exp{-d/l}exp{-t/τ} (5c)

Page 9, Line 9. α= (Σ TP+ Σ FP)(Σ TP+ Σ FN)/(Σ TP+ Σ FP+ Σ FN + Σ TN)= 1/N (Σ FP)/FAR (Σ TP)/POD

4. The phrase L1-regularization originates from using the norm of the function space Lp or the sequence space lp with p = 1. To be more consistent with the mathematical literature I recommend using the notation L1 and L2 instead of L-1 and L-2. It would be beneficial to define the L-norm/lp-norm in Eq. (1) (see p. 6, l. 1).

We include the standard definition of the L-norm in the text: Page 6, Line 20. The second term represents an adjustable parameter λ and the L-norm of x, which is a measure of the magnitude of the vector x defined as the following:

âĂŰxâĂŰ_L= √(L&Σ_(k=1)ˆn |x_k |ˆL) (2)

We change L-1 and L-2 to L1 and L2 everywhere else in the text.

5. In Section 2.2 (p. 4, ll. 20-26) I suspect there is a problem with the dimensions of hi and x. If x Rn, then hi Rn to build the scalar product. With this implicit definition hi changes for each realization of the scenario. On the other hand, hi (possibly hi, i = 1, 2,. . .) is defined as the 'archived footprint covering the complete set of observing locations and times'. This definition seems to be incorrect. I think hi is the footprint corresponding to a particular measurement restricted to the locations of potential emitters. Then, the forward model H for a particular configuration is build only by a subset of footprint indices that describe the corresponding measurements. There are several ways to define these quantities properly, but the way it is defined in the article seems incorrect.

We clarify the in the text:

Page 4, Line 28. We select the footprint information that corresponds to the locations of the n emitters so that hi is also a vector of n dimension.

6. As described in Sect. 1 the Barnett Shale has 20000 well pads in the 300 km by 300 km domain, i.e. a well density of 0.22 wells/km2. Other oil fields, like the Kern River Oil Field near Bakersfield, CA, have much larger well densities (> 200 wells/km2). Are the chosen well densities representative for certain types of oil fields? Also, if well pads (and other possibly emitting infrastructure) are not homogeneously distributed, the local density may be much larger. The densities analyzed in this study are thus more to the lower bound of what is required. The concept of spatial tolerance is an interesting extension. The analysis is carried out using the much lower well density (0.04). How do the results compare to the 0.2-case? I expect that the results in Fig. 7 are too optimistic for many oil fields with densely distributed infrastructure.

Printer-friendly version

Discussion paper

We thank the reviewer for these useful points on context and update the manuscript accordingly:

Page 10, Line 13. Actual fields can be even denser but we are limited in our investigation by the 1.3 $\times$ 1.3 km2 resolution of the WRF simulation.

7. Section 2.2 describes how the pseudo-observations are created. It seems that no transport error is considered in the noise (p. 4, ll. 29-31). However, transport errors are mentioned when describing the inversion methods (p. 6, ll. 18-20). Are transport errors included in the study? I think they should!

Non-perfect transport is anticipated and accounted for in the inversion through R (page 7, line 4). Transport errors are not included in the pseudo-observation creation (page 5, line 1), as this error is related to the instrument characteristics (Table 1).

8. The study considers column measurements by satellites but also a network of in-situ observations. The advantage of column measurements is that the in-flowing background concentration (= b, see p. 4., l. 26) is measured and the assumption that it is constant (or known) is justified. When considering only the in-situ network the background concentration is unknown, which is an additional challenge in the inversion. This aspect could be mentioned to support the assumption of a constant boundary in this study.

We address this issue in the text:

Page 5, Line 26. We assume that these sites report hourly data with 1 ppb precision and that the background concentration in surface air is constant, consistent with the assumption made for satellite observations. A variable background would complicate the problem but could be retrieved as part of the inversion (Wecht et al., 2014).

9. Also, I wonder at which altitude the in-situ analyzers are placed? I expect that local low mode emitters may have a significant influence on observations taken close to the surface.

We clarify in the text:

Page 4, Line 12. Surface observations are taken in the lowest model layer (centered at 28 m above ground) and the corresponding footprints are obtained by releasing and tracking back in time 100 particles at the observation location and time.

10. In Sect. 2.5 high-mode emitters are defined via the standard deviation, whereas in Sect. 2.1 high-mode emitters are those that exceed 40 kg/h. Which definition applies for the results? And what are the reasons to use that definition over the other?

We refer the reviewer to the discussion on page 8, line 10 of the manuscript:

"In a real-world application we would not know the actual pdfs of emissions (Figure 1), so we need to diagnose the occurrence of high-mode emitters on the basis of anomalies in the distribution of x ÌĆ."

We expand this discussion on page 8, line 13:

Using anomaly detection on x ÌĆ instead of a fixed threshold (e.g., 40 kg h-1) allows for generalization to other emission fields where the mean normal and high modes may be different than the Barnett Shale.

11. The concept of L2-regularization is well described in many textbooks covering inverse problems. I think that the given reference, i.e. Evgeniou et al. (2000), is not very helpful for applied atmospheric sciences. My recommendation for applied researchers would be P. C. Hansen, Discrete Inverse Problems: Insight and Algorithms, 2010, but many other options exist.

We thank the reviewer for the suggestion and change the reference to Hansen (2010).

12. Using L1-regularization to exploit the sparsity of the problem is a great idea, which turns out to give better results than the standard approach. This concept has rarely been used in atmospheric inverse modeling studies and is probably new to many in the research community. The references provided are helpful. Still, some questions

remain: - How does a solution produced by L1-regularization differ from one produced by L2-regularization? The answer is described in Sect. 3.1, but I think a figure comparing both solutions for some representative realization would be useful. The same figure could also be used to illustrate how high-mode emitters are detected from the emission estimate. - I assume and I hope that 5% of high-mode emitters is generally a large estimate of failing systems. Further, I suspect that the identification of high-mode emitters improves for a smaller percentage of failing systems. Are there consequences on the solution produced by L1-regularization if the solution is less sparse, i.e. more high-mode emitters? Is a low degree of sparsity important for the algorithm to perform well?

We include a new figure (now Figure 5) and add the following text:

Page 8, Line 4. Figure 5 shows the distribution x ÌĆ from a single realization of emissions, GeoCARB 4×/day pseudo-observations, and both L1 and L2 regularization. In this simulation, L1 regularization enables the retrieval of high-mode emitters while L2 regularization is more restrictive in allowing excursions from the low-mode mean.

Page 8, Line 15. Figure 5 shows thresholds for classifying high-mode emitters using anomaly detection and a fixed value of 40 kg h-1. The L1 threshold is larger than the L2 threshold, but smaller than the 40 kg h-1. Had the fixed threshold been used, some high-mode emitters (relative to x ÌĆ) would have not been classified as such.

13. It could be argued that 20% failure of an alarm system is still a lot, but some criteria for success needs to be applied. However, this criteria is neither mentioned in the abstract nor in the conclusions, when systems are defined successful or not. I think it should be briefly mentioned in both sections.

Page 1, Line 24. "...are successful (>80% detection rate, <20% false alarm rate) at locating high-emitting sources..."

Page 11, Line 20. GeoCarb shows little difference in success rate (Equitable Threat

Printer-friendly version

Discussion paper

Score (ETS) > 0.65) for 2 or 4 overpasses per day.

14. I recommend uploading scripts that reproduce the given results. The study could serve as a test environment for new modeling approaches and as an interesting project for students.

See Response to Comment #1.

15. P. 4, l. 2: A comma is missing in '(small, medium, large)'.

Fixed

16. P. 4, l. 20: ... 1.3 x 1.3 km2 pixels

Fixed

17. P. 5, l. 28: remove brackets around x ÌĆ

Fixed

Response to Comments from Anonymous Referee #2

1. I have significant concerns about the effect of meteorological conditions (especially wind speed) on the detection of methane plumes by satellite observations. The high-wind condition promotes the dispersal and dilution of methane in the atmosphere, which makes the methane enhancement relative to background smaller than that under a low-wind condition, and the resulting low concentrations are much difficult to be detected by satellite. Therefore I wonder if the results of this paper are sensitive to the meteorological conditions used. This study was performed based on only a 1-week simulation using the WRF-STILT model, however, it used a very strong statement on the findings in the abstract and conclusion parts. It cannot convince me that the 7 days meteorological fields are representative enough. I suggest that the authors give a detailed discussion on the potential influence of meteorological conditions on their results.

We thank the reviewer for bringing this point to light. We address this issue on page 5, line 6:

The mean daytime 10 m horizontal wind speed inside the observing domain during the simulated week is 5.4 m/s. Stronger winds could further dilute plumes within an observing domain, making the ability for satellite detection of emitters more difficult; on the other hand, the model transport error is less for stronger winds (Varon et al., 2018).

Page 11, Line 15. "Our results in these meteorological conditions can be summarized usefully in terms of answers to questions that a field manager might have:"

2. It's not clear to me how this study used the WRF-STILT model because the model configurations are not described in details in Sect. 2.2. How many theoretical particles are released from each receptor, and what are the receptor heights? How is the column footprint calculated? Is it integrated from the footprints of different vertical layers? It is important that the method part is self-contained and does not require the reader to go through another source.

We add more information in the text:

Page 4, Line 9. Footprints for each column were obtained by releasing and tracking back in time 100 particles from vertical levels centered at 28 m, 97 m, 190 m, 300 m above ground, and 8 additional levels up to 14 km altitude spaced evenly on a pressure grid.

3. In the inverse method, the observational error covariance matrix accounts for instrument and model transport errors (line 18, page 6). What about the representation error? How is it considered for the satellite observations that have a different pixel size from the resolution of transport model?

We clarify in the text:

Page 7, Line 5. Representation errors are negligible due to the model grid resolution finer or the same resolution as the instrument pixels (Turner et al., 2018).

Response to Comments from Anonymous Referee #3

1) My first major concern is regarding the treatment of clouds. The authors claim that they "assume clear-sky conditions to simplify the discussion" but this is a serious assumption, which, in my opinion, is unacceptable. The caveat at the end of the conclusions that states "as long as skies are clear" is really problematic. Clouds have a major impact on observational coverage. Would accounting for clouds enhance the disparity between the next-generation satellite and TROPOMI and GeoCARB? Actually, my main concern here is whether accounting for clouds would reduce the POD of the next generation satellite (as envisaged here) to less than 0.8, which would mean that even such an instrument would be unable to detect dense high-emitting sources. It is critical that the authors account for the impact of clouds in their analysis.

We thank for the reviewer for bringing up this important point. We expand upon our decision to consider only clear-sky on page 5, line 13.

Successful methane retrievals from satellites require clear sky. The probability of clear sky in a partly cloudy domain depends greatly on pixel size (Remer et al., 2012). Results for a partly cloudy condition would depend on the particular cloud configuration and would be difficult to generalize. Here we assume clear-sky conditions to avoid this complication, but the detection probability for high-mode emitters should then be viewed as an upper limit. In particular, it should be recognized that no detection from satellite is possible for a cloudy domain.

We explain how it relates to denser fields on page 10, line 13.

Actual fields can be even denser but we are limited in our investigation by the 1.3×1.3 km2 resolution of the WRF simulation.

2) My other major concern is with the treatment of model transport error. The authors assumed an error of 4 ppb for both the surface and satellite observations. However, that is not a justifiable assumption. The model transport errors at the surface, in the

vicinity of point sources, will be very different from that in the CH4 column. Assuming the same transport errors for these two types of measurements does not allow for a fair and meaningful comparison of the satellite and surface measurements.

We update our results and figures to account for this discrepancy:

Page 7, Line 10. "...σM is the model transport error standard deviation previously estimated to be 4 ppb for methane columns (Turner et al., 2018). Given the order of magnitude difference in sensitivity between satellite columns and surface measurements (Figure 3), we assume σM to be 40 ppb for surface measurements."

Page 7, Line 25. r_ij= cor(i,j) σ_Mi σ_Mj exp{-d/l}exp{-t/τ} (5c)

We update Figures 6 and 9 (previously 5 and 8). The new analysis did not change conclusions in the text.

3) My third concern is with the lack of discussion of the impact of systematic errors in the satellite data. Since the launch GOSAT it has become clear that systematic errors in the greenhouse gas retrievals pose a major challenge for the use of the data. I appreciate that it would be challenging for the authors to reasonably address the issue of biased retrievals in their OSSEs, but they should at least add some discussion in the manuscript about how systematic errors could confound the detection of the CH4 sources.

We add more information about systematic errors in the text:

Page 5, Line 3. SWIR instruments may also suffer from systematic errors but we do not account for those here in the absence of information. The largest source of systematic error on our scale would likely be the inhomogeneity in surface reflectivity (Pfister et al., 2005).

Technical Comments 1) Page 2, line 8: Satellites do not measure the atmospheric columns of methane. They measure backscattered solar radiation from which the atmospheric columns of methane are retrieved. Please change the wording here.

We fixed the wording in the text:

Page 2, Line 8. Satellites measure backscattered solar radiation in the shortwave infrared (SWIR) from which atmospheric columns of methane can be retrieved with near uniform sensitivity down to the surface under clear-sky conditions (Jacob et al., 2016).

2) Table 1: The row is not properly aligned for the TROPOMI entry.

Fixed

Please also note the supplement to this comment:
https://www.atmos-chem-phys-discuss.net/acp-2018-741/acp-2018-741-AC1-supplement.zip

---

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-741, 2018.