Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-741-RC1, 2018
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Chemistry and Physics Discussions
Open Access
EGU

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

Anonymous Referee #1

Received and published: 11 September 2018

1 General comments

The study addresses an interesting and current research topic, namely the monitoring of emissions from production sites in oil- and gas fields with different observing systems (satellite and in-situ). Even though these observation systems are not operated yet (with the exception of the TROPOMI instrument), it is important to investigate the capabilities of future measuring systems prior to their installation.

The study is clearly written and well structured; thus easy to follow. Also, the authors embed the study well in existing literature. Many references even for minor topics are provided.

Printer-friendly version

Discussion paper

CC BY

The study uses modern concepts of inverse modeling, namely high resolution transport models and L_1-regularization. Although its success in other applications, L_1-regularization has rarely been used in environmental inverse modeling studies so far.

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.

I recommend the publication after considering the following minor modifications.

2 Specific comments

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.

The authors use '$a \times b$' 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 '$a \cdot b$' or 'ab' to be consistent with standard notation. To describe the dimensions of a matrix, $m \times n$ is the standard notation.

The phrase L_1-regularization originates from using the norm of the function space L^p or the sequence space l^p with $p = 1$. To be more consistent with the mathematical literature I recommend using the notation L_1 and L_2 instead of $L - 1$ and $L - 2$. It

CC BY

would be beneficial to define the L-norm/l^p-norm in Eq. (1) (see p. 6, l. 1).

In Section 2.2 (p. 4, ll. 20-26) I suspect there is a problem with the dimensions of h_i and x. If $x \in \mathbb{R}^n$, then $h_i \in \mathbb{R}^n$ to build the scalar product. With this implicit definition h_i changes for each realization of the scenario. On the other hand, h_i (possibly $h_i,\ i = 1, 2, ...$) is defined as the *'archived footprint covering the complete set of observing locations and times'*. This definition seems to be incorrect. I think h_i 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.

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 $0.22\ wells/km^2$. Other oil fields, like the Kern River Oil Field near Bakersfield, CA, have much larger well densities ($> 200\ wells/km^2$). 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.

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!

Printer-friendly version

Discussion paper

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.

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.

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?

The concept of L_2-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.

Using L_1-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

Printer-friendly version

Discussion paper

CC BY

remain:
- How does a solution produced by L_1-regularization differ from one produced by L_2-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 L_1-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?

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.

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.

3 Technical corrections

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

P. 4, l. 20: ... $1.3 \times 1.3\, km^2$ pixels

Printer-friendly version

Discussion paper

CC BY

P. 5, l. 28: remove brackets around $\hat{x}$

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

CC BY