Referee comment

anonymous author

1 General comments

The authors addressed all concerns and questions raised by the reviewers. Most questions were answered by minor modifications of the manuscript. Only some aspects led to adaptations of the simulation setup. The modifications are well integrated in the previous version of the manuscript. The manuscript improved from these changes because the assumptions and limitations of the study are much better discussed.

A few comments remain and should be clarified prior to publication.

2 Specific comments

The supplement provides code written in R programming language, which is open source and easily applicable. The example code includes instructions, but is not running because the authors do not provide the required inputs, i.e. a forward operator or the footprints. The argument that the input can also be created using open source software is weak from my point of view, because installing and running a complex open source model, which has its own requirements, is far more an obstable than running an example script with prepared inputs. I suggest to include the required input data as part of the supplement. I do not expect a full program, but a script that can reproduce some basic results.

From authors response:
'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).'

From my point of view the transport errors should also be included in the creation of synthetic measurements (cp. e.g. Michalak et al., 2004: A geostatistical approach to sur-

face flux estimation of atmospheric trace gases; Miller et al., 2014: Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions). In a real data scenario the mismatch between simulated and measured data is not limited to instrumental noise. The noise model also includes (transport) model errors.
This aspect is important because the cross validation chooses a suitable weighting parameter λ for the inversion method based on the input data. A small noise (instrumental error only) on the input data results in a smaller optimal weighting parameter and an improved reconstruction. This approach is used in the study and likely produces results that are too optimistic.
Estimation of the transport model error is a challenge by itsself but the assumption has already been made by the authors by the choice of the covariance matrix R. A higher level of noise (transport and instrumental error) forces the inversion with a larger weighting parameter and shows a reconstruction ability that is closer of what can be expected from (future) observation systems. The situation in such an approach is still optimistic because the noise has the same characteristics as anticipated by the covariance matrix R in the reconstruction method.
If the intent of this study is to assume a perfect transport model because models may improve similar to observing systems the matrix R should reflect the reduced transport error estimate. To me omitting the transport error reduces the value of the results significantly. Such an assumption should be stated clearly in the abstract and the conclusion.

Page 4, line 20: I would not use the phrase 'dilution effect' and rather explain in half a sentence why the surface influence is smaller for column measurements.

Page 4, lines 26ff: In the definition of the footprint h_i, i.e. $h_i = (\partial y_i/\partial x)^T$, the variable x is already limited to the location of emitters and has dimension n. Thus, h_i has dimension n. The authors continue: 'We select the footprint information that corresponds to the locations of the n emitters so that h_i is also a vector of n dimension.'
To me this formulation seems a bit confusing. I suggest something like: 'h_i is then a vector of dimension n with the selected footprint information that corresponds to the locations ...'.

Page 4, lines 30-32: The background b is assumed to be constant in line 30. Implicitly, it is assumed to be zero in line 32. This discrepancy should be fixed.

Decide for one of the spellings: 'L_1-regularization' or 'L_1 regularization'. Both versions are used throughout the manuscript.

Page 6, line 13: I recommend using '(e.g. Hansen, 2010)' instead of '(Hansen, 2010)' , because it is just one of many possible sources.

Page 6, line 29: 'Evgeniou et al., 2000' does not appear in the literature list. Should it be replaced with 'Hansen, 2010' or something from the statistical literature?

Page 1, line 11 and page 11, lines 12-13.: As pointed out in Referee Comment #1 oil fields with much higher well densities exist. Since most of the analysis is carried out using the 100 emitter scenario, I am not convinced that such fields should be called dense (e.g. 'a high density of wells'). Are they dense for a particular type of gas fields (e.g. fracking)? Then, this should be specified.

Figure 5: An axis break at the y-axis could be useful. Including the thresholds is a great idea to vizualize the concept of detecting high-mode emissions. Which value of S is used for the threshold? Maybe the uncertainties in the threshold by varying S could be included in the figure, too.

Figure 9: Why is the uncertainty from the threshold increasing for a larger number of surface sites for the next-generation instrument?

Overall, the manuscript reads well and is an interesting contribution to the scientific community.