Geosci. Model Dev. Discuss., 7, C1442–C1447, 2014
www.geosci-model-dev-discuss.net/7/C1442/2014/

Geoscientific Model Development Discussions

Interactive comment on "FLEXINVERT: an atmospheric Bayesian inversion framework for determining surface fluxes of trace species using an optimized grid" *by* R. L. Thompson and A. Stohl

R. L. Thompson and A. Stohl

rona.thompson@nilu.no

Received and published: 21 August 2014

This paper outlines the development of an inverse modelling software package "FLEXINVERT". The methods employed have mostly been published elsewhere and are well known. Therefore, whilst there are relatively few new insights in this article, the paper is a very thorough and clear account of FLEXINVERT system. I think it will be suitable for publication in GMD, once the following comments have been addressed.

We thank Referee 1 for his/her constructive comments and reply to them below.

General comments

One of the main findings in the paper is how sensitive certain parts of the inversion are to "background" mixing ratios. This is well known for Lagrangian model inversions, but I think it merits further discussion here. In this paper, baselines are either estimated using an Eulerian model, or from the mole fractions directly. There are number of problems that occur to me: a) In the case that optimized TM5 mole fractions were used, there is an element of circularity, because the TM5 mole fractions will have already seen the observations. b) In the case where the lower quartile of observations were used, did the authors take into account the fact that the lowest measured mole fractions can have a wide range of origins? In particular, for methane, when air enters Europe from the Southerly sector, it can be significantly depleted in methane, compared to other "baselines" from the Atlantic? If I understand their method correctly, the baselines they obtain would be rather smoothed, and would not identify short-timescale "low methane" events. Even more significant "depletions" have been observed elsewhere in the world, for example in East Asia, where the air can rapidly fluctuate between Northern and Southern hemispheric during the summer. c) In the case where these baseline mole fractions were optimized, there is also an element of circularity, as the observations themselves have been used to determine the "prior". In each of these cases, there is the potential for the choice of baseline to erroneously influence the derived emissions either through biases (that would likely not be well accounted for in the uncertainty quantification method outlined, which assumes only stochastic errors). I don't think the paper needs to solve these problems. However, I think the discussion could be expanded very slightly to further highlight some limitations.

The problem of accurately determining background mixing ratios for Lagrangian models is a challenging one. We address the two cases the referee mentions as follows: 1) using the optimized TM5 mole fractions. If we consider just using the background directly from the TM5 optimized fields, i.e. not including these in the optimization of the state vector in the Lagrangian inversion, then there is little circularity, but rather this can be seen as a two-step optimization. In this case, the observations are used to optimize the background in the first step and the fluxes in the second step. On

the other hand, if the background variables are also included in the state vector, then there is a degree of circularity. However, the two inversions, i.e. the TM5 and the Lagrangian inversion, do not use the exactly same set of observations nor at the same time frequency, uncertainty estimate or transport, therefore there is still new information when optimizing the background in the Lagrangian inversion. Generally, global models, such as TM5 are optimized using a standard observation dataset based on e.g. the NOAA GMD network, while a regional inversion (in this case the Lagrangian inversion) makes use of other observations instead of or in addition to those used in the global model thus minimizing or avoiding the problem of circularity. It should also be noted that the background calculated for the Lagrangian model is an extremely smoothed version of the Eulerian mixing ratio fields, as the sensitivity to the background even for a single measurement is distributed over large parts of at least one hemisphere. In the global model, this background is constrained mainly by measurements from outside the region of interest. In contrast, the influence of the stations used for the regional inversion (which typically are concentrated in the area of interest) on the background as sampled by the Lagrangian model is quite small. This minimizes the degree of circularity inherent to the use of optimized concentration fields.

2) using the lower quartile of observations. We tested a number of methods for calculating the baseline using the observations themselves, including also higher temporally resolved baselines. We found from these tests that having step jumps in the baseline can cause problems in the optimization of the fluxes given the imperfect transport and often weak observational constraint. Instead, we found a smoothed baseline to be more robust although this does not take into account instances where the background has considerably different mixing ratios, which as the reviewer states, is an important problem and was also mentioned by Stohl et al. (2009). In their study, Australian emissions were overestimated, as they were constrained mainly by measurements from a single station (Cape Grim), which is located in the south of the emission area and for which a background was derived that is mainly representative for air masses coming from the west. The derived background, however, was systematically too low for

air flows that are coming from the north (where the Australian emissions are located). This was compensated in their inversion with unrealistically high emissions in Australia. Such problems are much less severe for emission regions surrounded by several measurement stations, for which background biases will be in the opposite direction and will partly cancel each other. Future work, could involve combining information from the Lagrangian model about where the air masses originate (especially the latitude) to help better define the observation-based background. In any case, the degree of smoothing, i.e. the window length over which to calculate the lower quartiles and the averaging is a parameter that can be changed to best represent the variability in the background. Concerning the circularity of using the observations to determine the prior background, the observations that correspond to background are those that have very little influence from fluxes within the domain. Thus the corresponding rows of the transport operator, H (or the Source-Receptor Relationships = SSRs), are close to zero, thus these observations will not contribute much to the constraint of the fluxes within the domain.

We have now expanded our discussion in section 3.4 to include both the concern of circularity and the temporal resolution of the background calculated from the observations.

Specific comments

Page 3754, line 13: I'm not sure what this sentence means, or whether "smearing" is the best word to use.

We mean that a Lagrangian model can have infinitesimal resolution and thus is able to better resolve measurements at a site while an Eulerian model has fixed resolution and the measurement site is represented by the value of the corresponding grid cell (or in some cases the interpolation between grid cells). We have changed this sentence to:

"A further advantage of LPDMs is that they can be run backward exactly from a measurement site, unlike Eulerian models, where site measurements are represented by

the averaged value of the corresponding grid cell."

Page 3761, line 11 and Equation 6: I'm not entirely clear why this is necessarily an aggregation error. I think you can formulate this problem so that ymod for the variable grid is identical to that obtained using the full grid, by using the emissions-weighted footprint. In that case, the aggregation error would only come in during the inversion.

It should be noted that Hnestvg*fnestvg is not equivalent to Hnest*fnest. The difference between these is the aggregation error.

Equation 7: I think Fout should be lower case

In Eq. 7 we have used Fout as this is a matrix consisting of row vectors of fout. Thus Fout and fout are not equivalent.

Page 3762, Line 26: I think this should be a numeric "1", rather than "one"

We have changed "one" to "1".

Page 3766, Line 18: This idea has been used elsewhere. Perhaps a reference or two should be given.

We have included a reference to Bergamaschi et al. 2009.

Page 3768, Line 2: superscript "T" for transpose in the line below the equation.

This has been corrected.

Section 2.9: If I read this correctly, it appears that this "non-negativity" correction only applies to those grid cells where negative emissions were obtained? I reality, if the first inversion could "see" this constraint, wouldn't its effects be felt further away than the individual grid cells where negative emissions were derived? Perhaps a line or two of clarification, could be provided.

In fact, the non-negative constraint (given in Eq. 20) does not only affect the grid cells with negative values but there is also some adjustment to other cells according to the

correlations described by the posterior error covariance matrix, Aflux. This is now also stated in Section 2.9:

"The inequality constraint does not only affect the grid cells with negative values but there is also some adjustment to other cells according to the correlations described by the posterior error covariance matrix, Aflux."

Page 3768, Line 26: Appendix C

We actually refer to the description of the full code and not just to the inequality constraints. This sentence was supposed to be a new paragraph. Appendix B is the correct reference.

Page 3770, Line 22: What about the representation error at the flask sampling sites?

Since flask measurements were assimilated without averaging (and by choosing the closest modelled back trajectory to the flask sampling time) no representation error was included for these, i.e. only the model and measurement errors were used.

Page 3774, Line 24: This assertion could be tested by running the model at a higher release height.

We tested this by running the Lagrangian model, FLEXPART, at a height of 2000 masl to represent PUY. The correlation coefficient and normalized standard deviation for the observations and the simulation using the prior emissions with 2000 masl was 0.25 and 2.09, respectively, compared with 0.18 and 2.24 with 1475 masl, i.e., the actual height of PUY. It is likely that at 1475 m, the influence of local emissions at PUY is overestimated owing to a too deep BL in FLEXPART, meaning that PUY is too often in BL compared to the reality.

Interactive comment on Geosci. Model Dev. Discuss., 7, 3751, 2014.