“Estimation of volatile organic compound emissions for Europe using data assimilation”
M. R. Koohkan, M. Bocquet, Y. Roustan, Y. Kim, and C. Seigneur
acp-2012-786, J. West Ed.

Response to Reviewer #2

We thank the reviewer for his/her useful comments and suggestions.

The authors present an adjoint inversion of European VOC emissions based on an annual cycle of concentration measurements at a network of sites. The mathematical framework for the analysis is thoughtfully presented, and the paper provides an interesting comparison of results obtained using varying statistical assumptions in the inversion. The application is fairly novel, in the sense that there have not been many inverse studies of speciated VOC emissions based on this type of surface network. The topic is appropriate to ACP.

While the statistical assumptions going into the solution are well thought out, the physical assumptions are less so. My major comments relate to this issue. If these are addressed I recommend publication in ACP.

Scientific comments.

1. The authors examine how their inferred fluxes change with different assumptions regarding the error covariance matrices, etc., which is great. However, they do not test their sensitivity to forward model error, which can often be a larger source of error. How sensitive are your inferred fluxes to uncertainty in a) model transport, b) model OH, c) boundary layer height, and d) boundary conditions? Currently I think this is the biggest shortcoming of the paper.

2. A related point is that the authors do not provide any uncertainty range on their inferred fluxes. That should be fixed. Even if it is difficult to fully characterize the a posteriori uncertainties, some attempt should be made. For instance, do you regard the range between the Case B1, B2 and C fluxes in Table 6 to provide a reasonable estimate of the true uncertainty in the solution?

We treat questions 1 and 2 altogether since we consider these questions to be interconnected.

We agree that a posteriori uncertainty is a key issue. Note that (i) the cross-validation study and (ii) the forecast study of the original manuscript are to a large extend sufficient to validate the interest of the retrieved emissions because of the objective gain in the scores. That is why we deemed the a posteriori uncertainty computation unnecessary. Nonetheless, we agree it would be better (and interesting *per se*) to compute them.

One way to do it is to perform a sensitivity study, as suggested in the first comment. However the impact on the retrieved emissions will depend on the estimation of the prior uncertainty of the identified error sources, such as model errors, boundary conditions, representativeness errors, which is a very difficult task in this context, since so little is quantitatively known about most VOC. For instance, with an ensemble, it would all depends on its calibration whose satisfying reliability would imply using observations, that are not numerous enough... Nevertheless, we give ideas on the uncertainty in the boundary conditions in response to your comment and in one of the specific answers to Reviewer 1, because we believe they are the first source of possible errors for long-lived species emission inversions.

We have chosen to compute an objective (i.e. following a statistical model of the errors) estimation of the errors (as a whole) in the retrieved emissions. In the revised manuscript, we have added a dedicated subsection “Uncertainty of the retrieved emissions” that replaces the “Information and DFS section”. Similarly to Winiarek et al. 2012, we have performed an extensive Monte Carlo study that samples posterior emission fields, using the prior errors whose magnitudes have been objectively assessed with the optimal hyperparameters. This is meant to account for the impact of any source of errors. But, as downsides, it is only accounted for via two coarse scaling parameters, and the method does not assume biases which are possible for some of the VOC boundary conditions.

We do not believe that the range of variability of cases B1, B2 and C can provide a reasonable estimate of the true uncertainty. However, in a restrictive sense, the variability between B2 and C, as two different reasonable inverse modelling methods, is likely to provide an estimate of the fraction of posterior uncertainty due to choice of the inverse modelling method. This remark has been included in the new posterior uncertainty section.

3. Figure 5 shows that the B2 optimization does not significantly change the OH fields. What about the B1 optimization, does that change OH to any significant degree? Since this is the only scenario where isoprene emissions change appreciably, this is where one might expect to see a notable OH change. (also, just because the OH concentration doesn't change much with the inversion doesn't mean it was accurate in the first place, see comment 1).

B1 optimisation is in the manuscript because it is a very simple approach and for the sake of comparison with the B2 and C optimisations. However it leads to unphysical (negative) emission fields. This is especially true for isoprene. It turns out that the B1 optimisation scheme does not lead to a very different picture as seen in Fig.5 because they are average values over the whole domain. But again since the negative fluxes are unphysical it is not wise to analyse too much the results of the B1 optimisation.

In connection to your question and one of Revewer 1's, we have significantly changed subsection 4.1.1 “A posteriori verification of the model linearisation”. It was supported by new numerical simulations and new checks.

4. On page 33238, you make the point that the inversion does not work very well for isoprene because of its short lifetime. Does the surface network provide measurements of MVK and MACR, and if it did would you expect this to help to any significant degree?

No, routine monitoring networks do not provide measurements of methyl vinyl ketone (MVK) and methacrolein (MACR). A limited field campaign is planned to measure oxidised volatile organic compounds and we hope to take those new measurements into account to further constrain the inverse modelling of emission fluxes. However, this future work is out of the scope of the present work.

5. What do we learn from the inversions that gets at some improved knowledge of process? Are there specific source sectors or regions that appear to warrant specific investigation? What recommendations would you make to inventory developers?

We agree that the interpretation of the results can appear a bit frustrating. However going beyond our current conclusions appears difficult. On the basis of a Positive Matrix Factorisation (PMF) approach we have tried to disaggregate the global inversion to propose an information detailed by activity sectors. However this attempt has failed to provide robust results due to the relative collinearity of the chemical profiles of the emissions sectors of the species. The use of source specific tracers, or a larger number of VOC species, might overcome this issue, but such data are not available to our knowledge for this study. The current study should be seen as a feasibility demonstration rather than a fully operational tool to correct emission inventory. Our text has been modified to report this discussion (see response to Reviewer 1 too).

Technical / editorial comments.

1. The first several pages are taken up with mathematical underpinnings of chemical transport modeling and adjoint analysis, much of which is not new. As a suggestion, this section could be shortened by referring to previously-published work.

In the acp final format, sections 2.2, 2.3 and 2.4 only take less than two pages. The purpose of this introduction is to present the notations to the reader and (very briefly) describe the methods and the key concepts. It is very difficult to pursue without introducing them at the risk of witholding crucial information to the reader. On the contrary, Reviewer 3 would prefer us to comment more on the methodological concepts.

Note that section 2.5 (one additional page) is not standard. It had only been introduced in Winiarek et al, 2012, in a different context.

2. Top of page 33235, ug/m3 is an odd unit for OH, O3, and NO3.

The concentrations are now provided in # cm^{-3} for radicals and in molec cm^{-3} for ozone.

3. Throughout, suggest spelling out the compound names (and/or formulae) rather than using abbreviations, which are not defined in the text. While they are defined in a table, readers might just assume "ACE" refers to acetone or acetaldehyde (for example) rather than acetylene.

Ok. In addition to the table, we have systematically introduced the abbreviation at (at least) the first occurrence of the species in the text.

Figure 7, consider putting the species names inset in the panels rather than "a", "b", etc.? It would make for easier reading.

Done. Thank you for the suggestion.