Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-535-RC1, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "An atmospheric inversion over the city of Cape Town: sensitivity analyses" *by* Alecia Nickless et al.

Anonymous Referee #1

Received and published: 24 September 2018

Main comments:

The authors present an atmospheric inversion result over Cape Town focusing on sensitivity analyses related to the technical aspects of the inversion method. I can easily see the authors did a lot of work. However, the presentation needs substantial improvement as well as revisions in technical details.

First, the authors definitely need to rewrite the abstract. Simply it is too long and not organized well (please see my specific comments below).

The introduction section also needs lots of changes or rewriting. Please see my comments below. Basically, it is too technical from the beginning of the section, not providing a gentle overview of the study presented. I recommend that the section be

Printer-friendly version

Discussion paper

shortened.

The writing is below average compared to many papers I have reviewed. I understand that the authors did a lot of work but in many places, but the result/discussion presented is not so clear. The paper is too long for the reader to read in current form while there is no exciting scientific findings - this does not mean that the material is not important (it is a different paper). I wonder if the authors can reduce the number of sensitivities cases by (re)moving some of the insignificant results to the supplement.

Please see the detailed comments below and address them before I consider any suggestion for publication.

Detailed comments:

Abstract.

Simply put, the abstract is too long while not conveying useful information in a succinct way. Needs significant improvement in writing (and selecting the most useful pieces of information to be presented here).

Please try to rephrase "A carbon assessment product of natural carbon fluxes, used in place of CABLE, and the Open-source Data Inventory for Anthropogenic CO2 product, in place of the fossil fuel inventory, resulted in prior estimates that were more positive on average than the reference configuration." - a little awkward.

Also, the authors need to divide the following sentences into two (unless made clearer): "For the Cape Town inversion we showed that, where our reference 10 inversion had aggregated prior flux estimates that were made more positive by the inversion, suggesting that the CABLE was overestimating the amount of CO2 uptake by the biota, when the alternative prior information was used, fluxes were made more negative by the inversion. "

Please remove the following (you can state in the results or discussion section): "As the posterior estimates were tending towards the same point, we could deduce that the

Printer-friendly version

Discussion paper

best estimate was located somewhere between these two posterior fluxes. We could therefore restrict the best posterior flux estimate to be bounded between the solutions of these separate inversions. "

What is the main conclusion we can gain from the abstract? The authors need to emphasize it. Currently, I only see many small points and cannot determine which one to take home.

P 2, L12: Please remove "where estimates of CO2 fluxes can be derived from measurements of CO2 concentrations at a point location", which does not represent the general atmospheric inversion.

P2, L17: Not all of inversions do that; depends on the study. It could be fossil fuel only.

Introduction: The authors are more focused on the technical aspects of the inversion method considered here by starting describing what atmospheric inversion means in terms of technique, even in the first paragraph of the introduction! Please reframe the introduction so that the authors approach the problem from the urban greenhouse gas (GHG) perspective. People may be interested in Cape Town GHG emissions (more generally), which I haven't heard much before.

Also, please reduce the introduction section because it includes too many technical details/terms. It should be a gentle "introduction" to the paper.

P2, L25: covariance matrices => uncertainty (or error) covariance matrices

P6, L10: "s" should be the surface fluxes, not including the background (i.e., CO2 concentration at the boundary). This is because "Hs" is from the model, not the measurements. Also, c_mod should be Hs_0 (s_0 is prior fluxes in Eq. 1), right?

P6, L13: Change s to s_0. Is s_0 hourly or weekly? Even if you solve for the weekly mean surface fluxes, for CO2, I would expect that hourly prior fluxes were used. Please clarify.

P6, L17: “The boundary concentrations in s”? Why “s” when you talk about concentration. “s” can only be linked to concentrations via H. When you refer to concentrations, it should be “c”, not “s”; “s” is fluxes. Right?

P6, L20: Change “can be added to the measurement errors contained in C_c” to “can be added to the error covariance matrix C_c that includes measurement errors”. Mathematically, C_c includes all different error sources, but, to be specific/accurate, we want to separate transport errors from those of measurements.

P6, L23: Are 4 and 16 ppmˆ2 the total variance (i.e., including transport error, background error, etc.) in the diagonal elements of C_c that the authors actually used in the inversion? Then, do the authors have any scientific/statistical evidence that these numbers really represent the total irreducible variance in the error covariance matrix? In other words, how did the authors come up with these number?

P7, L3: Why is 1-hour assumed for L? It seems too short. After an hour, are the errors uncorrelated? Usually, following synoptic scales of meteorology, it could go hours and days.

P7, L7: Please add a subsection for the transport model because in current form the authors try to combine the Bayesian inversion method with everything (transport, prior flux, etc.) that is part of the inversion system; not convenient for the reader to follow.

P7, L19: Please add information of temporal and spatial resolutions of the prior flux, as a minimum detail.

P7, L32 - 34: Related to this, please add a few sentences about C_s_0 (prior error covariance) including the structure (e.g., dimension, etc.). In this way, the reader should be able to better understand how the authors treated the prior error covariance.

P8, L2-4: Any concern of aggregation of hourly to weekly? If the authors aggregated into monthly, I would be definitely concerned, but weekly aggregation is in the gray area, it seems to me. The way I would do it is that you still use prior predictions in hourly (i.e., Hs_0 in hourly in eq. 1) while solving for weekly mean "s". CABLE is originally 1 x 1 km? If not, please say so.

P8, L29: "in place of" => "in addition to". Both bio prior emissions are used?

P9, L6: Where is this standard deviation coming from?

P9, L9: Please add a few sentences about Figure 1. How are the two bio prior fluxes are different (e.g., in total)? How has the uncertainty in the two priors been estimated?

P10, L4-16: This paragraph can be shortened because it does not include any specifics on the author's work. Does Hetia have anything to do with this work? Except for the product description, I don't see any point here.

P10, L21: Please spare your space more on Figure 2 where you compare the two products for prior fossil fluxes. Are they different? If so, how much, in the bottom-up inventory perspective?

P13, L10: The naming is quite confusing. When I started reading the result section, it was confusing and I had to come back here to check the definition. "an inversion which assumed no temporal error correlation in the specification of Cc" := NEE Corr. But no hint of "NEE" in this definition. In Table 1, it says NEE Corr is defined as "no observation error correlation". I understand this is the case without off-diagonal elements. Right? There is a disconnection between L9 and L10 of P13.

P13, L33: What is the difference between "Simp Obs No Corr" and "No Corr". As written, it is not clear.

P14, L1: Please use "state vectors" instead of "control vectors" because "s" (flux) really means the state, which is commonly used in the timeseries model. In GHG inversion work, I have never heard of control vectors.

P14, L11-12: I don't think I have seen a clear description of the background concentration (or boundary concentration). Why only four corners? Since a Lagrangian approach is used, why not sampling boundary conditions for each of the particles?

Reading “The inversion solved for $4\times2\times4$ = 32 boundary concentrations” I understand that the authors seem to solve (as in “s”) for the a single boundary condition for day or night for each week. 4 corners x 2 (day and night) x 4 weeks? Ideally, each (hourly or sub hourly) CO2 observation has to be associated with the boundary condition. It looks like weekly mean boundary conditions were used, which is not quite okay. Only four corners were used? If so, this is too much simplification. Please clarify how the authors treated the upstream boundary conditions.

Even if the authors used a simple one-valued boundary condition for day and night, I am doubtful about the robustness of the estimation of those 32 values of boundary conditions when solved together with “s”. In a sense, Bayesian inversions use regularization methods via prior assumptions, which means a state vector of 244,824 (huge) can still be solved with a small number of observations. But here because the authors are solving for hundreds thousands of parameters, the posterior is highly dependent on the prior. Related to boundary conditions, what this means is that the posterior boundary conditions (if the authors really estimated the posterior boundary conditions while doing inversions, not pre-subtracting; please clarify) is significantly affected by the prior. If so, what prior did the authors use for the boundary condition?

P15, L27: It is okay to use Xˆ2 for assessing the goodness-of-fit, but please state the assumption related to this test and whether the data used in the inversion meet the test assumptions. Also, state that what Xˆ2 results mean. Xˆ2 itself does not guarantee the accuracy of the results.

P18, 3. Results: Please add a subsection here; it looks like an introduction to the Results section but it is a mix of many things. I strongly recommend that the authors remove some to other sections or rewrite it. Basically, what is the main topic for this whole page?

P21, L2: Please define bias (obs - model?) if it has not been done somewhere else.

P21, L11: Then what does it suggest? The model (Gaussian here) and data using ODIAC are more consistent . . .?

P21, L14: That's because the prior uncertainty was extremely small. Is it a correct prior assumption? It is over-confident!

P22, L2: Which uncertainty? Please be specific.

P22, L7: Typically, biospheric fluxes are much more uncertain. This near-zero uncertainty on the prior suggests to me that the prior assumption is wrong.

P22, L9: Before moving to spatial distribution, do we have any conclusion from this time series comparison? What does all this comparison mean?

P27, L12: How small is the Xˆ2 value? Ideally Xˆ1 should be close to 1. Is it good or bad? This sounds like ignoring temporal correlation is okay?

P27, L13-15: This needs some clarification. What is the difference between Ref with positive covariance (L13) and just Ref (L15). Which one is compared with which one here. This result suggests "no correlation" has a minimal impact on the posterior?

P27, L17 - L22: The author should be able to explain why there is a such a big difference between weekly and monthly. I don't quite understand why.

P27, L23 - 27: The paragraph starts with Ref and NEE Corr and then mixed up with Obs Corr and No Corr. It is really hard to follow; this happens in many places throughout the paper. Not a smooth reading at all.

P27, L 27: This result seems to be important in terms of error reduction. Please add a couple of sentences for this. From Figure 7, I see the central estimates between No Corr and Ref are similar while the error reductions are different.

Section 3.3 & 3.4: I don't have much comment except for the fact that it is somewhat boring to read - please try to convey in a clearer and succinct way!

Printer-friendly version

Discussion paper

P38, L9: Please clarify what “and could not react to local climate conditions” means.

P38, L13 - 15: Not clear what the authors mean by “The ODIAC product extended the fossil fuel fluxes much further a field from the CBD region than the reference inventory. This led to aggregated estimates that were much larger under the ODIAC inversion than the reference inversion.” How is the first sentence is related to the second sentence? What do the authors mean by the statement in the first sentence?

P38, L15: “The inversion attempted to reduce the aggregated flux” means when the model tries to match the observations?

P38, L18-20: Please provide estimates (in numbers) for both in the text so that the reader can clearly see the likely true emission estimates. Each inversion should have a uncertainty bound and then I don’t understand what it means by “a much narrower uncertainty region than for either inversions.”

P38, L26-28: 1 hour is too short. It should be useful to see the results based on 6 hours or 24 hours. I expect the length scale would be hours or even a couple of days.

P39, L17: This is not correct. Prior is just prior. Your sampling from a prior distribution with a fixed mean and a fixed covariance is still a priori info. It does not require the prior sample to be accurate.

P39, L19 - 20: This is because your data points are too small compared to the number of parameters to be solved. In other words, your inversion system is more dependent on the prior rather than observations. In this case, the posterior estimate for the individual pixels won’t have much constraint; only the regional total emission may be estimated more or less independently, in the best case. From the Bayesian perspective, the only thing you can do is to report what your assumption was, what model was used and what the result is.

P39, L23 - 27: Not a Bayesian way of thinking, subject to criticism from frequentists.

P42, L14 - 15: Since the authors are using an analytical solutions for a Gaussian likelihood function, they could use a simple maximum likelihood estimator for the length scale.

P42, L25 - 29: Please correct the sentences. Also, I don't know what the authors are trying to say here, except for the fact that a hierarchical approach may be better.

#paperreview

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-535, 2018.