Interactive comment on "Investigation of error sources in regional inverse estimates of greenhouse gas emissions in Canada" *by* E. Chan et al.

Anonymous Referee #1

Received and published: 21 September 2015

General comments

Although I see that the authors have put a considerable amount of work into preparing this study, there are a number of major flaws in the method, which make it unacceptable for publication. First, the influence of land biosphere fluxes of CO2 is completely ignored. Therefore, this synthetic case study is of no use in understanding the performance of real observation inversions, since real observations are sensitive to land biosphere fluxes. The gross land biosphere fluxes of CO2 are generally much larger than the fossil fuel fluxes and are the greatest source of uncertainty. By just analyzing an inversion for fossil fuel fluxes, the largest part of the CO2 inversion problem is

ignored. Second, there are flaws in the inversion method. The fact that the scalars optimized, λ, apply to the product of the transport and the fluxes, biases in the transport between sub-regions will be folded into the scalars. I have included further details of these problems under "Specific comments". In addition, I am suspicious about the result in the "no flux and no transport error" experiment, in which large posterior errors were found. This should not be the case with no prior and no observation error, which makes me also skeptical about the other results. Lastly, numerous studies have been previously published analyzing transport and other uncertainties on the retrieved fluxes. Unfortunately, this synthetic study fails to bring any new insights in how to best define these uncertainties or set-up the inversion problem.

Specific comments:

The abstract is difficult to follow and needs to be improved for clarity (see also comments below).

P22717, L10: Could the authors please explain what they mean by the "assumed model-observation mismatch", do they rather mean the uncertainty in the observation space? The model-observation mismatch, obviously is just the difference between the modeled concentrations and the observations, which does not need to be "assumed".

P22717, L11: What is meant by "estimation error" do the authors mean the difference between the "target" and posterior fluxes. Please specify.

P22717, L15-20: Please state what these percentages represent - are they the fraction of posterior-target flux difference compared to the target flux? This is unclear. (Note that the abstract should be understandable without having read the entire paper beforehand).

P22718, L13-14: It is not true that global inversions are unable to resolve fluxes on smaller than sub-continental scales. A number of inversion frameworks based on global Atmospheric Transport Models (ATMs) with, e.g., use of the adjoint model, re-

solve fluxes on the grid cell level, i.e., order of a couple of degrees. Whether or not these inversions are able to independently constrain the grid cell is a matter of the observation constraint.

P22719, L3: There are a number of regional Eulerian models that are used for inversions, e.g. CHIMERE, so please change "typically" to "a number of" or equivalent.

P22719, L4: Generally Lagrangian models are driven by reanalysis data, which are data assimilation products, to say "modeled meteorology" is misleading.

P22720, L16, By "estimation error" do the authors mean the difference between the posterior and target fluxes. "Estimation error" should be defined.

P22721, L5-6: Is it correct that only the fossil fuel emissions of CO_2 were used to simulate CO_2 concentrations? If so, were the very large CO_2 fluxes from the land biosphere ignored? And if this is the case, then the results of this study are of very limited use (and of no use at all for determining fluxes of CO_2), as in order to determine fossil fuel CO_2 fluxes, the land biosphere fluxes also need to be determined, and these have the largest uncertainties.

P22721, L22 – P22722, L5: It is well known that having variables that represent large regions (large with respect to the heterogeneity with the region and the influence this has on the observations) is a source of model representation error, or specifically, aggregation error. This has been shown in numerous previous studies, importantly those of Trampert and Snieder (1996) and Kaminski et al. (2001). The work of Kaminski et al. even provides an algorithm to determine this model representation error in the observation space. On the other hand, while an inversion may not be able to separately constrain the variables at fine resolution, this can be ascertained from the posterior error covariance matrix (seen from negative covariance between variables). In this case, the variables can be aggregated a posteriori to give more robust estimates for the larger regions with smaller uncertainty than for the individual variables since the errors have negative covariance. Therefore, I do not see what can be learned from

performing inversion test cases using differing numbers of regions.

P22723, L13: Why was a height of 100m chosen for the surface layer (or using the author's terminology, footprint layer)? If the height was increased, then there is a greater probability of it containing particles and thus, better statistics, on the other hand, the height of the footprint layer should be within the PBL. What is the influence of changing the footprint layer height?

P22724, L5: "Cost Function Minimization" (CFM) is very generic, as all Bayeisan methods attempt to find the minimum of some cost function, whether it be using conjugate gradient methods, Newton methods, analytical methods or other. Please specify which method was used here.

P22725, L2-4: It is not generally true that the MCMC method requires fewer variables than the method that the authors call CFM. In fact, what the authors call the CFM method is in principle a least squares method.

Eq.1 & 2: The scaling factors should be applied to the fluxes, i.e., the unknown variables, x, and not to the product of Mx. The way that this equation is expressed, one is optimizing the transport as well, which should not be the case, the model M must be assumed to be known within the uncertainties, which are given in the observation space. This means that if transport biases exist between regions, then this will be folded into the scalar estimate.

Fig. 1: related to the above point, I suppose the authors mean by "aggregating the mole fractions to sub-regions" they actually mean the allocation of the transport operator into each of the sub-regions, so that the influence of each sub-region on the mole fractions is separated as shown in Eq. 1?

Eq. 4: Again, the same applies as with the MCMC method, the scaling factors are applied to the product of the transport and fluxes, thus if there are biases in the transport between regions, this is folded into the scaling factors. This is an additional avoidable

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

source of error.

P22728, L9: While it is true that the real error covariances are not known, it is not true that the prior error covariance matrix, Dprior, is typically assumed to be diagonal. A lot of research has been done (and papers published) on defining patterns of error covariance in these matrices.

P22730, L3-5: Because the main fossil fuel sources, e.g., industry, transport, power plants etc. remain largely the same between two consecutive years there will be a strong correlation between the CT2010 and CT2011 fluxes. Therefore, the inversions for the "flux error case" do not represent the reality in which the flux error may have a very complicated structure.

P22730, L20-21: These synthetic experiments do not account for the land biosphere fluxes of CO_2, which strongly influence real CO_2 observations. The co-dependence of transport errors and CO_2 land biosphere fluxes is a very important source of error in CO_2 inversions, e.g. the seasonal and diurnal so-called "rectifier" effects. By ignoring these, this study is of limited use to real observation inversions of CO_2.

P22734, L18: The fact that the posterior error is considerable in the "no prior flux error and no transport error" case makes me suspect that there is a bug in the set-up. The modeled – pseudo-observation differences must be zero, thus the optimal fluxes should be very close to the prior fluxes, which in this case is also the target fluxes.

P22743, L15: I think it is pure coincidence that there is a cancelling effect between the flux and transport errors. This is generally not the case.

Technical comments

English language editing is needed especially in the use of articles and punctuation. I have pointed-out only a few examples here:

P22717, L22: "having" should be "to have"

P22717, L23-24: words missing in "could be unrealistically small that do not cover the target"

P22718, L2: missing article "a real inversion"

P22718, L8: please change to "in the context of national reporting of emissions"

P22718, L12: punctuation "Global inversion systems, such as CarbonTracker, which. . ."

P22718, L14: "at subcontinental scales"

P22719, L8: change "annual" to "1-year" as "annual" is an adjective

P22719, L11: ". . .apply to. . ."

P22720, L25 – P22721, L4: This sentence is very long and difficult to understand. After reading it twice I realized that adding the correct punctuation helps: ". . .atmospheric inversion, using simulations run in a backward (adjoint) mode, are: the synthetic. . ." but please consider revis

---

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 22715, 2015.