Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-213-RC1, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Objective evaluation of surface- and satellite-driven CO_2 atmospheric inversions" *by* Frédéric Chevallier et al.

Anonymous Referee #1

Received and published: 5 June 2019

General comment:

The article presents global inversion results for an ensemble of 6 inversion set-ups, comprising observations from OCO-2, GOSAT and in situ measurements, and using two versions of the transport model. The presented results show that the OCO-2 XCO2 retrievals have matured sufficiently to yield inversion results based on satellite observations that are of similar quality as inversions based on (sparse) in situ measurements. This is a very promising result, paving the way for future satellite inversions. However, the main point the authors want to make consists of "defining quality measures for global inversion systems in order to evaluate the current skill of global inversions". This effort, in my opinion, falls short. Whereas the manuscript presents a nice analysis of an inversion ensemble, a clearly defined set of quality measures is hard to extract

from the text, nor is the main verification approach (comparison against aircraft observations) very novel. It is unclear to this reader what exactly these proposed quality measures look like, or how they should be applied to objectively benchmark the quality of inversion results. Overall this is a nice article, but the focus of it is not very clear: is it the definition of quality measures, or the analysis of a set of global inversions in view of the effect of the transport model and assimilated observations? This ambiguity unfortunately causes neither of these points to come across sufficiently strong.

Specific comments:

- The effect of the transport model on the inversion result is included in the analysis by using two different versions of the same model. The results indicate that this difference has very little effect on the obtained surface fluxes, from which the conclusion seems to be drawn that transport errors have a much lesser influence on the quality of the inversion than do the assimilated observations. Although this seems to be the case for the analyzed inversions, it is well known that in general the transport model is one of the largest sources of difference between inversions (as for example shown again recently in Schuh 2019). This analysis is therefore misleading: the chosen variations in transport model are likely not representative for the actual transport variations typically observed between different inversion set-ups. In order to include these results, they should therefore be situated in the context of an analysis of the decrease in model error between both versions, and preferably also compared against the typical spread between transport models.

- The OCO-2 MIP (Crowell 2019) is touched upon rather briefly, but it would make the analysis stronger if more of the obtained results are put into perspective with respect to the inversion ensemble analyzed in the MIP.

- One aspect which might be given some more attention in the manuscript is the presence of spatial patterns in the optimized surface fluxes due to the used prior, another aspect that typically causes large differences between inversion results. It is men-

tioned that the OCO-2 and in situ based fluxes look similar, can this be due to the prior pattern being dominant over observation patters in regions with sparse observational coverage?

- The aircraft verification data has especially large coverage over North America. Is the coverage over the other continents sufficient to make statements about the inversion quality? How is this considered when analyzing the root-mean-square difference (Line 28 p7)? Are these values a representative quality measure, or biased towards quality above North America?

- Similar to figure 1, it would be nice to include maps with the observational coverage (and if possible quality) of the different sets of assimilated observations, and to refer to this information in the discussion of the inversion results.

- Line 32, page 4: please include a little more information about the used transport uncertainty.

- On page 4 (lines 37-39) it is mentioned that the bias of OCO-2 ocean observations is still too large for these observations to be included in the inversion, based on initial testing. It would be interesting to include a little more information here: how significant is the bias and the effect of it on optimized fluxes?

- Line 2 on page 5 states that outliers are rejected: are these considered outliers in the aircraft data, or in the simulation result? If the latter, is there a pattern to be discovered? And are the same observations removed for all inversions in the ensemble?

- Line 4 p6: this is not entirely clear: Do you mean that the grid cell value is used as simulated value for the verification data?

- Line 33 p9: if the aircraft data does not allow to distinguish between the quality of OCO-2 and in situ-based fluxes, what does this imply for future quality assessments of inversions? One might expect that the quality of inversion results will increase in the coming years, both through the availability of high-quality satellite data, and through

improvements in inversion set-ups. Would this mean that the presented aircraft verification approach will no longer be useful to distinguish between their quality?

Small comments / technical corrections:

- Please use ppm/yr instead of ppm/a

- Lines 47-51 p7: it would be clearer if you move these after eq(1), especially since the equation is now on the next page.

- Figure 6 might be clearer if you remove the edges of the contourplot, as they now look messy in combination with the edges of the continents.

- Figures 7&8: please start the y-axis of both plots from zero. It would be insightful to include the 0.15ppm benchmark to show which biases fall within the measurement uncertainty.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-213, 2019.