Response to Reviewer 2

The Reviewer's comments are in black text. Our responses are in blue text. The modifications and additions to the text are highlighted in yellow in the revised manuscript PDF file. However, to see what was deleted, please see the annotated original manuscript PDF file.

General comments: The paper discusses first developments towards an assimilation system for optimizing greenhouse gas concentrations to analyze the carbon cycle globally, but also for Canada. The new developments comprise an extension of the Environment and Climate Change Canada's operationally used Ensemble Kalman Filter to CO observations. The new systems behavior is analyzed using identical twin experiments.

The paper is of general concern for GMD but lack of clarity in the argumentation. While the system aims at analyzing the carbon cycle on a global scale, but also for Canada, the analyzed time interval from 27 December 2014 to 28 February 2015 is sub-optimal. The papers Fig. 2 suggests that CO fluxes for January 2015 are negligible for Northern Canada. The identical twin experiments should be conducted in the wild fire season to proof the systems ability of analyzing the CO state appropriately even in challenging wild fire episodes. While assimilating synthetic CO observations, the paper claims at several points to be a greenhouse gas assimilation system and that the model state to be optimized is augmented by CO, CO2, and CH4. A clear distinction between future efforts towards the full system, covering also CO2 and CH4, and the current state of the system is not given. In the introduction, it is not made clear how CO assimilation can also improve the concentrations of greenhouse gases. A paragraph about this aspect would be appreciated.

Response: We thank the Reviewer for the many helpful comments and suggestions. Both Reviewers felt that the manuscript did not sufficiently distinguish between the completed work (CO state estimation) and the context of our desired future work of building a full greenhouse gas state and flux estimation system. Thus, we have revised the manuscript to mention our context in only the first paragraph of the Introduction, and our future work in the Conclusions section. The term "greenhouse gas" no longer appears except in these two places.

The focus of this work is on the presentation of the CO state estimation. Thus, the experiments we performed in winter were adequate for our purposes because there is significant wildfire activity in the tropics to generate reasonable CO fields. However, for our future flux estimation work, the Reviewer is absolutely correct that we should test the system during the Canadian fire season (boreal summer). Nevertheless, we carried out some new identical twin experiments in summer 2015. Please see section 3 in the supplementary material. Figure S7 compares our true flux fields in January and July 2015. July 2015 was very active in terms of Canadian forest fires. Therefore the RMSE over North America is much higher in July 2015 than in Jan-Feb 2015. It is seen from the results that the benefit over North America during summer is much larger than that during winter (Figures S8-S10).

We have also provided references to the literature that relate CO and greenhouse gas estimation. Please see revised lines 34-36. For our system, the initial intention is qualitative: simulations where all three species show the same patterns in a given region indicate a fire source of CO_2. It is also generally expected, by the CO_2 flux estimation community, that additional information from other species such as CO will be needed

in the future to attribute surface fluxes to natural or anthropogenic sources. However, the best means for using CO in this fashion has yet to be resolved. Our inclusion of CO prepares for the eventuality of the whole field identifying the best means to constrain CO_2 fields with CO measurements.

Specific comments:

generally, the Grammar of the paper, especially the use of commas and articles, should be reviewed
Response : We have carefully reviewed the manuscript with this focus.

line 17: replace "GHG" by "greenhouse gases (GHG)"
Response: Done

line 24: change "2015 or 2016" to "2015 and 2016"
Response: Done

line 24-25: add a reference for the statement on NIR estimates of anthropogenic CO in Canada
Response: The reference was provided on line 20. It was added again to line 25.

line 29-34: The claim is not clear. If wild fires are important for the carbon cycle, why not conducting the experiments in the wild fire season. Further, EC-CAS v1.0 does not assimilate CO2 and CH4. Thus, the paper should not claim that EC-CAS does include CO2 and CH4 in the assimilation process.
Response: This statement speaks about the goal of EC-CAS and our goal is to assimilate all 3 species, though we do only CO here. The forward model does include all 3 species but the assimilation was tested only for CO, so far. We have rewritten the manuscript to avoid mentioning the other species (CO_2, and CH_4) or surface flux estimation except for the first paragraph which provides motivation and when discussing future plans. In the present work, we are demonstrating that the CO state assimilation is functioning well. Since wildfires occur at all seasons at different locations around the globe, any season is adequate for our present purpose. However, when testing the flux estimation capability it will be necessary to choose the boreal summer when Canadian wildfires tend to occur. We are currently in the process of testing this capability and this extension to our work will be described in a subsequent article. Nevertheless, we also conducted some new experiments for CO state estimation for June 2015 (see the new supplemental material). Indeed, our CO state estimation works just as well in boreal summer when there are larger surface fluxes over Canada from wildfires

line 81: change "GHG and flux estimation" to "CO estimation". In section 2.4 only consider CO estimation in, not GHG. Further, as this is not the purpose of this paper, do not explain details about flux estimation. This should be attributed to the respective paper
Response: "GHG" was changed to "CO atmospheric distribution". (Lines 81-82 in original manuscript → lines 64-65 in revised manuscript). Section 2.4 was modified to remove any mention of flux estimation. The title of section 2.4 was changed to "EnKF extensions for CO data assimilation").

line 95: please add: …at every grid point "as well as an perturbed CO emission fluxes."

Response: The sentence was rewritten (as requested by Reviewer 1). We also added the point about perturbed fluxes. It now reads: "A number (N=64) of 6 h model forecasts are simultaneously integrated from N meteorological and CO initial conditions with forcing from N perturbed CO surface fluxes". (Lines 94-95 → 73-75 in revised manuscript).

line 99 -101: There is no need for repeating the outline of the section. Please remove

Response: Reviewer 1 made the same comment and these lines were deleted.

line 107: "…and the same lid of 0.1° hPa." Please correct the unit. Do not use "lid", rather use "model top" or equivalent.

Response: Reviewer 1 had the same point and "lid" was removed. Also the spurious "°" symbol was deleted. . . (line 108 → line 89).

line 116: start a new paragraph

Response: Done. See line 98 in the revised manuscript.

line 121: replace "This is because… is used for..." by "Thus, …can be used for"

Response: Done. (line 121 → 102-103).

line 122: replace: "…with an EnKF so the computational expense of complete chemistry is prohibitive and difficult…" by "with an EnKF. The computational expense of the complete chemistry would be prohibitive and difficult…"

Response: Done. (line 122 → line 104).

line 134: add a reference for the statement made in the parenthesis

Response: Our methane simulations have not been published, but we did carry CH_4 in the forward model so we had to define a reasonable initial condition because of conversion to CO. This statement describes how this was done.

line 188, Equation 4: An information about the form of the observation operator, especially for MOPITT-like observations, is missing in the manuscript. Please also consider talking about MOPITT-like observations, rather than MOPITT observations. Further, what is the difference between ρ_m and ρ_o?

Response: The MOPITT observation operator is now better described in section 3.2 . Please see lines 279-298 in revised manuscript. Additionally, we refer to MOPITT-like observations throughout the manuscript. ρ_m and ρ_o are defined immediately after equation 5. See line 177 in the revised manuscript.

line 190: replace "when“ by "if“
Response: Done. (line 190 → line 220).

line 193: replace "For example when the both the row and column: : :“ by "For example, if both, the row and column…“
Response: Done. (Line 193 → Line 223).

line 194: replace "that element“ by "the respective element“
Response: Done. (Line 194 → line 224).

line 200-201: This sentence needs to be linked to the rest of the paragraph
Response: The sentence was deleted as requested by Reviewer 1.

line 213: a table of parameters and the value range would be appreciated
Response: A table was added in the supplementary material (Table S1).

line 235: Do not start a new paragraph
Response: Done.

line 237-238: Do not include the outline of the next paragraphs
Response: Lines 237-8 were deleted.

line 248-250: replace "…one each for January and February 2015" by "one for January and February 2015, respectively". Further, rephrase to following phrase.
Response: Done. (line 248-250 → Line 259).

line 303: check the line breaking
Response: Done.

line 305: replace "This control experiment assimilates…" by "The control experiment (EXP_CNTRL) assimilates…"
Response: Done. (line 305 → line 319).

line 314: …results of assimilating the meteorological variables.
Response: Done. (line 314 → line 326).

line 316: the aspect of area-weighted statistics is not made clear. Please provide a description of the weighting procedure.
Response: This is simply the standard computation in atmospheric modelling for computing global mean quantities on a sphere. The equations are now described in the supplemental material (section 2). Also "area-weighted" was changed to "global mean" in revised line 328.

line 320: A reference for the climatological values of the temperature uncertainty is missing"
Response: These sentences were rewritten as per Reviewer 1's request.

line 331: …RMSE (Figure 6c) and its comparable strength is encouraging…
Response: Done. (line 331 → lines 341-342).

line 334: …over the analysis period is shown by…
Response: Done. (line 334 → line 344).

line 337-339: Do not include an outline of the next sections
Response: These lines were deleted as requested by Reviewer 1, also.

line 340 (section 4.2): This paragraph lacks on focus. The spatial correlation on two specific days is given. How does this supports the analysis in the subsequent sections? No investigation about the influence of different localization radii is done. The influence of the localization radius on the assimilation results is not investigated. Please consider removing this section or expand it to a more detailed investigation on the localization radii.
Response: This section is meant to be pedagogical. We assume that our paper will be read by researchers who specialize in ensemble techniques but also by researchers who work in inverse techniques but are not familiar with EnKF. This section is meant to illustrate the concept of state dependent covariance estimate and the related issue of localization. Some concepts are easier to explain with the help of pictures in conjunction with equations. In this section we have tried to explain the motivation behind physical localization along with state dependent sample spatial correlation. We did investigate effect of localization radius value on the results and concluded that 2000 km is a good choice. This is mentioned in the revised section 2.4.

line 371 (section 4.3): throughout this section please be careful in the description of the results. E. g., refer to HYPNET observations but to the EXP_HYP experiment. The same for all other experiments/observation types. This has been mixed up several times.
Response: We have reviewed the manuscript and ensured the appropriate usage of the terms.

line 395: The mean relative benefit…
Response: Done. (line 395 → line 396).

line 397 – 402: Please make the description more specific by, e. g., including specific values of the benefit.
Response: We added the values of relative benefit in the last sentence of the paragraph. See lines 405 in the revised manuscript.

line 400: By comparing Fig. 9a and 9b, it is evident that the benefit due to the assimilation of HYPNET observation and MOPITT-like retrievals is also comparable in the column mean,…
Response: Done. (line 400 → line 402)

line 403: Fig. 10a shows the benefit of the EXP_GAW experiment.
Response: Done. (Line 403 → 406)

line 407: replace: "Though USA does not have any stations in this experiment…" by "Even though no stations are located in the USA in this experiment,…"
Response: Done.(line 407-409 → line 412).

line 414: Fig. 10a shows that assimilation of GAW observations results in…
Response: Done. (line 414 → line 419).

line 448: please use km as unit for consistency
Response: Done (line 448 → line 453).

line 450: …ECCC observations compared to HYPNET observations, which are located at about 1 km.
Response: Done. (line 450 → line 455).

Description of Fig. 11: This description is tedious and have to be condensed. For the results of this analysis, the precision of the given height of the observations is irrelevant. Please consider summarizing the mean height information of the different station types ad experiments in a table
Response: We agree with the reviewer that broadly speaking the exact height of observations is irrelevant for the analysis. However there are exceptions: eg. the observation at Mt. Kenya spreading observational information in Central Africa and the observations in eastern Canada spreading information to eastern USA. We agree with the reviewer that the discussion is tedious and needs to be condensed. We have shortened the discussion. We have deleted sentences mentioning the heights of HYPNET since it is repetitive. We have retained the discussion about panel (11a) with minor modifications. We decided against including a table. This is because there are qualitative aspects to the connection between heights of observations and the benefit. If a table is included we will have to point the reader to the table for the height and then describe its connection with the benefit. Therefore leaving the heights in the text is better.

line 473: no greenhouse gas assimilation system was presented. Please be more precise
Response: "greenhouse gas" was changed to "atmospheric composition". (line 473 → line 468).

line 479: .. due to the assimilation of observed CO is proportional...
Response: Done. (line 479 → line 473).

line 480: Another factor, which controls the pattern of the benefit, is the location of observations.
Response: Done. (line 480 → line 475).

line 482: …2000 km, which is the localization radius used in these experiments.
Response: Done. (line 482 → line 477).

line 483: …lowermost 500 m than observations at 1 km.
Response: Done. (line 483 → 478).

line 486: replace "Pacific" by "Atlantic".
Response: Done. (line 486 → line 483).

Figure 1: change "prescribed CO fluxes" to "ensemble of prescribed CO fluxes"
Response: Done.

Figs. 4 and 5: please increase the resolution of the figures title and axes annotations
Response: We have maximized them.

Figure 5: please change the x-axis annotations to dates, same for Fig. 7
Response: Done.

Figure 6: The vertical range of the averaged column (0-5km) is not consistent with other figures, where the range is 0-10 km. Please verify. Further: … (d) RMSE of the EXP_HYP experiment.
Response: Figure 7 was changed to show average from 0-5 km. The Figure caption was corrected.

Figure 8: Spatial correlation of CO between Toronto…
Response: Done.