Response to Reviewer 1

The reviewer's comments are in black. 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 describes the new development of the coupled weather and atmospheric composition system based on the Environment and Climate Change Canada's (ECCC's) operational Ensemble Kalman Filter (EnKF). While the paper describes this new configuration as an important advance for the ECCC system it misses important points to provide an accurate and complete description that such system should deserves. The first major point that needs to be addressed is that the paper advertises in several places that it is a greenhouse gases (GHG) atmospheric data assimilation and surface flux inversion system. However only CO atmospheric data assimilation is showcased. I would strongly recommend that the authors remove all claims that a flux inversion GHG system has been setup and then use a different terminology such as simply "atmospheric composition data assimilation" or "atmospheric carbon data assimilation" as in the title. The study uses synthetic observation to evaluate the system. Therefore, why the authors did not simulate the HYPNET CO2 and CH4 observations and perform the assimilation of such to at least justify the GHG component of the system? It seems that the added value of the paper is the extension of the ECCC operational system to atmospheric composition using CO assimilation as a proof of concept. While the focus is on CO assimilation, very little importance is given to the meteorology assimilation evaluation in such configuration. How does this compare to the actual operational ECCC system? Almost no references are given to reader to refer to the NWP system and its evaluation. I would recommend the authors to give a short summary on the meteorological data assimilation rather than ascertaining that the meteorological data assimilation is working as expected. The overall presentation of the paper requires strong efforts to improve clarity. Almost all parts of the paper lack clarity. Some parts are over emphasising some aspects that are not relevant for the evolution of the system while other parts that are important are covered very briefly. To give few examples:

- Very little is explained about the simulation of MOPITT synthetic observations, averaging kernels and their errors. It seems that a paragraph is maybe missing.
- Extensive description of the meteorological setup is given but very little is described and showed about the actual meteorological data assimilation results.
- Some of the terminology used is not really common for atmospheric data assimilation, I would encourage the author to revise this throughout the text.
- Several misleading statements about data assimilation and atmospheric composition need to be corrected.

Please refer to the specific comments for details.

Response: We are grateful to the Reviewer for their careful reading of the manuscript and for helpful suggestions. Our original intention was to present our work as the first of a long series of steps to reach our final goal of a greenhouse gas and flux estimation system using an operational weather forecast model. However, both Reviewers felt that the presentation did not sufficiently distinguish the completed work from the context of the desired future work. This led both Reviewers to conclude that the organization of the manuscript was confusing and possibly misleading. We appreciate this feedback and have rewritten the manuscript to focus on only the CO state estimation work which was completed. The overall context and goals of our project are now limited to only the first paragraph in the Introduction, and to a discussion of future work in the Conclusions sections. The term "greenhouse gas" also does not appear anywhere except those two mentioned locations where goals or future work is described. Also, as suggested by the Reviewer, we have provided much more detail about the meteorological assimilation system, both the pre-existing, operational system and our modifications to it in the revised section 2.3. A new supplemental section shows

comparisons of our EC-CAS meteorological estimation with that of the original system (Figures S1-S6) as well as a Table S1 of the types of model perturbations used. Also, the behaviour of meteorological fields in the simulated observation context is now shown with 3 extra panels in Figure 5. In addition, as requested by the Reviewer, the data assimilation terminology was modified as suggested, and all of the specific comments were addressed. Finally, more information about the generation of simulated MOPITT observations was added (line 268-281 in original manuscript and lines 279-297 in the revised manuscript). Overall, we feel that the revised manuscript has greatly benefitted from feedback of both Reviewers. Below, we respond point-by-point to each of the specific comments made by this Reviewer.

Specific comments:

Line 38: Be consistent, so maybe replace by air quality. Or explain that air quality is partly driven by weather.
Response: Good point. "weather" has been changed to "air quality" (line 38 → line 45 in revised manuscript).

Lines 39-41: This sentence has some shortcomings that could mislead the reader. Be consistent with the previous sentence and please develop this statement in more precise information. Air quality is a bit different from tropospheric pollution. Tropospheric atmospheric composition prediction is essential to air quality prediction which is looking at surface levels of pollutants. Tropospheric pollution prediction relates to longer time scales than 5 days, especially for CO. Air quality is driven by emissions variations and synoptic variations of weather regimes.
Response: Thanks for the clarification. "Tropospheric pollution" has been replaced by "Tropospheric atmospheric composition prediction" in this sentence (line 39 → line 45).

Line 41: Which data assimilation systems are we talking about here?
Response: The word "those" has been deleted for clarity (line 41 → line 47).

Line 63 and line 65: Swap years to chronological order
Response: The introduction was rewritten and the paragraph containing these lines was deleted.

Line 78: This system now can estimate emissions using state augmentation as described in Gaubert et al., 2020 (Gaubert, B., Emmons, L. K., Raeder, K., Tilmes, S., Miyazaki, K., Arellano Jr., A. F., Elguindi, N., Granier, C., Tang, W., Barré, J., Worden, H. M., Buchholz, R. R., Edwards, D. P., Franke, P., Anderson, J. L., Saunois, M., Schroeder, J., Woo, J.-H., Simpson, I. J., Blake, D. R., Meinardi, S., Wennberg, P. O., Crounse, J., Teng, A., Kim, M., Dickerson, R. R., He, H., and Ren, X.: Correcting model biases of CO in East Asia: impact on oxidant distributions during KORUS-AQ, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-599, in review, 2020.)
Response: Thanks for the update. The statement on lines 79-80 has been deleted. This reference now appears in the revised section 3.2.

Lines 81-82: Maybe this is a bit misleading as the paper seems to focus on CO (even if CO is important for GHG estimations). Also, the term "estimate GHGs" is a bit vague in my opinion. Maybe replace to something more specific such as "estimate CO atmospheric distribution".
Response: "GHGs" has been changed to "CO atmospheric distribution" (line 82 → line 65).

Lines 88-91: This paragraph is not necessary here as some of it should be moved to the introduction.
Response: This paragraph was deleted.

Line 88: "Trial fields" is quite uncommon data assimilation terminology. Maybe replace by forecast, background, prior or first guess fields depending on what you are meaning by trial here.
Response: "Trial fields" used to be quite common in atmospheric data assimilation. The problem with "background" is that inverse modellers (especially those dealing with CO_2) reserve that term for the large global mean over which local spatial perturbations exist. In addition, inverse modellers use the term "prior" to refer to the flux priors and applying this to the CO state could be confusing. We replaced "trial fields" with "forecast fields" throughout the article. Similarly "trial ensemble" has been replaced by "forecast ensemble" throughout the article.

Lines 88-90: The first and second stage are not explicitly mentioned. I would rewrite those two general sentences with a more traditional way to introduce the general concepts of data assimilation.
Response: As noted above, this paragraph was deleted.

Lines 94-95: The sentence "The model is initialized: : :" is confusing please rephrase.
Response: The statement has been changed to: "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).

Line 95: Please "trial fields" replace with appropriate traditional data assimilation terminology throughout the text.
Response: As indicated in the response to comments for Line 88 above, we have replaced "trial fields" throughout the manuscript with "forecast fields". This paragraph (lines 93-101) has been rewritten for clarity (lines 73-83 in revised manuscript).

Line 97: "Blending" is not really the correct word for the data assimilation procedure. I would recommend the author to use the appropriate vocabulary for data assimilation in the literature that tackles atmospheric data assimilation.
Response: As noted above lines 93-101 have been rewritten.

Lines 99-101: You do not really need to specify what will be the sections to come here. Consider removing.
Response: We have deleted the sentence in lines 99-101.

Line 108: I do not think that "lib" is the appropriate terminology here. Please again replace with, for example: "... 80 vertical levels from the surface to 0.1 hPa."
Response: Done. (line 108 → line 89).

Line 109: what type of hybrid coordinate? There are several of them.

Response: It is a log hybrid pressure and sigma coordinate that is commonly used (with slight variations) in operational weather forecast models. The provided reference (Girard et al., 2014) describes it in detail.

Line 114-155: Please be more specific and add diffusivity in this sentence.
Response: Done. (line 114 → line 95-96).

Line 120: Not correctly written. The atmospheric chemistry scheme is not removed for CO2. You remove the reactive chemistry in a model to increase its performance. Please rephrase.
Response: The sentence was changed to: "In contrast, GEM-MACH-GHG uses a simple parameterized chemistry for CH_4 and CO while CO_2 is treated as a passive tracer." (line 120 → lines 101-102).

Line 130: Start a new paragraph here as you now write about CH4 surface fluxes.
Response: Done. (line 130 → line 112).

Line 135: Start a new paragraph here as you now write about CO emissions.
Response: Done. (line 136 → line 119).

Line 148: Please define xf and xa here. xf and xa are commonly called the prior and posterior state respectively in the EnKF terminology. Alternatively, you could call them forecast (hence the superscript f) and analysis (hence the superscript a). Please consider using the commonly used atmospheric data assimilation vocabulary throughout the text for more clarity.
Response: xf and xa have been defined as the forecast and analysis states. (line 148 → line 145).

Line 148: Consider directly defining the other elements of the equation 1 before going into explanations.
Response: Lines 153-154 were moved to lines 146-147 in the revised manuscript.

Line 152-153: The sentence "Pf is the forecast error: : :" is a bit vague, please be more specific in the definitions.
Response: We have introduced equations 2 and 3 which define P^fH^T and HP^fH^T.

Line 196: I think there are more relevant papers for this statement. In Inness et al., 2015 the system used was a CTM configuration where the meteorological fields are forced by external meteorological fields. In that sense the DA system could not drive any constrain on the meteorology. Please cite instead Barré et al., 2015 and/or Gaubert et al., 2016 and/or Kang et al., 2012 and so on... Those papers are using EnKF with this variable localisation between atmospheric composition and meteorological variables.
Response: Actually, Inness et al. 2015 does refer to a coupled system with chemistry modules embedded in the meteorological model. The older MACC system (Inness et al. 2013) was coupled in an offline way to the IFS, but Inness et al. 2015 point out that the earlier approach was not computationally efficient and that chemical tendencies were held fixed for 1 hour and this caused problems at the day/night boundary (see their Introduction). This was the motivation for a fully online chemistry model (called C-IFS). Furthermore, Inness et al. (2015) state on p3 (section 2.2) that "the error covariance matrix between chemical species or between chemical species and dynamics fields is diagonal". Thus variable localization was done by them. However, it is true that other references could be added here. We added Barré et al., 2015 and Gaubert et al. 2016. (line 196 → line 226).

Line 200-201: The sentence “The spatial correlation: : :” seems to have no link with the previous ones. Please remove or develop in a new paragraph.
Response: The statement was deleted.

Lines 209-210: The sentence “To simulate model : : :” is unclear. Please rephrase and possibly add a reference for this error representation method.
Response: The discussion of the meteorological system was moved to the section 2.3 and rewritten. This section now pertains to only the additional changes needed for CO data assimilation.

Lines 214-215: But this paper is not doing flux estimation. Maybe consider changing to atmospheric composition data assimilation and change to the appropriate references.
Response: As noted in our response above, this section was rewritten and moved to section 2.3. There is no mention of studies related to flux estimation in the revised section 2.3 and section 2.4.

Lines 218-219: “In EC-CAS, for the meteorological assimilation, the same scheme is used, but for GHGs, no such additive error is present.” Is this the configuration used in this paper? If yes, why then bother going though all these details above?
Response: This paragraph was rewritten as noted in our last 2 responses. We moved all discussion of the EnKF configuration for meteorology to section 2.3. This then simplifies and clarifies the additional changes needed for CO assimilation.

Line 224: Then why not using synthetic GHG observation of CO2 and CH4 (amongst other GHGs)?
Response: Although the model does include CO_2 and CH_4 as well as CO, and our ultimate goal is to assimilate all 3 constituents, it is a major undertaking to test and validate the system for each of the three species. Each species is quite different in nature and in fact has a completely different literature. Thus the team involved in the validation exercise for a given species would be different. As an example of the differences, CO_2 has a lifetime of ~200 years and a very large background with primary surface fluxes from the terrestrial biosphere, ocean, fossil fuel emissions, fires, and land use change. However, CH_4 primary surface fluxes include agriculture, wetlands, ocean, anthropogenic emissions, fires and an atmospheric sink and it does not have such a large background value. Thus CO_2 has large positive and negative surface sources whereas CH_4 has mainly positive surface sources. Thus the best way to simulate surface flux uncertainty in the two cases will differ. Also, because CO_2 has a huge background and we are interested in variations of 1-10%, the type of forecast error variance inflation needed will differ from CH_4. CO differs again in having a larger dynamic range in mole fractions than either CO_2 or CH_4 and the shortest lifetime of the three species. Thus, timescales for forecast error variance saturation will differ. Then we come to the observing networks, which are quite different for all 3 species. This is just to name a few of the differences. We do plan to study the assimilation of each species separately, in good time.

Line 236: change to “the use of a surface flux”. I would recommend the authors to be consistent with this terminology as fluxes are not necessarily at the surface in the atmosphere.
Response: Good point. We now refer to surface fluxes here and elsewhere in the paper.

Lines 269-270: ": : : its retrieved profiles are sensitive to CO in the lower troposphere where : : :" MOPITT retrievals are sensitive throughout the entire troposphere. The multispectral retrievals allow an enhanced sensitivity towards the surface over land only when the conditions are favourable. Please correct and amend the text accordingly.

Response: We have revised the sentence to read "retrieved profiles are sensitive to CO in the lower troposphere during daytime over land, where the flux signal from surface emissions is most readily detected." (line 269-270 -→ line 279-281)

Line 271: What are those data assimilation systems? This statement is not true. Number of air quality DAS only assimilate surface stations. Please be more accurate here.

Response: We have changed the sentence to read **"**As a result of this sensitivity to lower tropospheric CO, and the long observational record, MOPITT data are widely used for inverse modelling of CO emissions and for air quality studies**"** . We have added references to indicate some of the specific data assimilation systems for which this statement is true. "(e.g. Arelleno and Hess 2006; Fortems et al. 2011; Barré et al. 2015; Jiang et al. 2015b; Yin et al. 2015; Mizzi et al. 2016; Inness et al. 2019; Gaubert et al. 2020;l Miyazaki et al. 2020). (line 271 → lines 281-283).

Line 273-274: This statement is misleading. You do not use the averaging kernel to construct the observation operator. You feed the observation operator with the averaging kernel to sample the first guess.

Response: We have modified the text to better explain the need to account for the averaging kernel in the observation operator. (Please see lines 288-298 in the revised manuscript).

Line 276: This is unclear. Does this mean you discard all observations that have a retrieval surface pressure below 1000 hPa? I do hope you are not doing this. Please clarify the sentence.

Response: The sentence we wrote does not accurately reflect what we did. The 10 levels are a fixed grid. There are no actual observations below the surface. The lowest retrieved level corresponds to the surface level, which may lie at lower pressures than 1000 hPa. We have deleted this sentence and modified line 275 by replacing "1000 hPa" by "surface". (line 275 → line 286).

Lines 277-278: This is not the proper definition of the averaging kernel matrix. Please use the common definition given by Rodgers 2000. Inverse Methods for Atmospheric Sounding. Theory and Practice. https://doi.org/10.1142/3171 | July 2000. Pages: 256. By (author):; Clive D Rodgers (Oxford).

Response: Please see the modified text from lines 285-298.

Line 278: H is not a forward operator but only an observation operator in the Kalman filter as it does not need to generate a forward model prediction to get a model equivalent quantity. It is true in for example the 4D Var formulation. Please correct.

Response: We have replaced "forward operator" by "observation operator". Please see lines 293-294 in the revised manuscript.

Line 281: The authors do not use the same system as in Jiang et al., 2015a. If they do this needs to be clearer earlier in the paper. If not, please recall a bit more of the methodology or use the appropriate reference to the system used in this paper.
Response: The assimilation systems are different, but those differences are not relevant here. We have removed this sentence since explaining the methodology of the Jiang et al. study would not be helpful for the discussion here.

Line 285: "varied between 10-16%" is this the value that the authors use to set up the observation errors. It seems that few sentences are missing to explain the setup on MOPITT observation errors.
Response: We have used 10% to set up observations errors. We cite Deeter to justify this value. Please see lines 302-304 in the revised manuscript.

Lines 289-290: This sentence is hard to understand. Please rewrite.
Response: We have rewritten these lines.. Please see lines 305-311 in the revised manuscript.

Line 290: "other issues" Please be specific of what other issues.
Response: Please see lines 305-311 in the revised manuscript.

Lines 293-299: So why do the authors bother simulating observations then? Why not testing the DAS in real conditions? Please justify more clearly the choices here and certainly earlier in the paper.
Response: An important stage in the development of any data assimilation algorithm is to prove that it works. We know from data assimilation theory, that in the absence of bias in observations and models and with plentiful, and accurate observations, the system should work. By simulating observations, we can satisfy the constraints of unbiased observations. We have tried to achieve a balance between a highly idealized setup and reality by allowing the transport model to have imperfections and by using (simulated observations from) real networks like ECCC and MOPITT. Adding different observation networks gives us a further qualitative sense of whether the system is behaving properly since we can guess how using more realistic networks with data gaps will behave relative to the uniformly dense network. Assimilating simulated observations helps us to build confidence in the system we have built. This is only the first step. We will be assimilating real observations. Please see lines 305-311 in the revised manuscript.

Lines 308-309: The statement "An ensemble of forecasts: : :" is incomplete as is, I would remove it as this would need couple sentence to make this point clear and this paragraph is not the place for that.
Response: These 2 statements were deleted. We have explained the role of state dependent correlation in spreading observational information in other sections.

Lines 311-312: This was already mentioned earlier. Remove.
Response: Done

Line 320: Regarding the reference to Pires et al., 1996, I think numerical weather prediction and predictability ranges have evolved since the mid-90's. Please use a more recent reference. Also, the time of the DAS RMSE stabilisation is not due to weather predictability but mostly due to the DAS setup, i.e. background error, observation density, type and error and so on... Please rewrite the related statements.
Response: The statement was deleted.

Lines 321-322: The authors could add winds, surface pressure and RH (or another NWP variable of your choice) in a four-panel plot to make your point stronger and avoid such statement.
Response: Figure 5 was expanded to include other meteorological variables.

Line 328: What is the "additive model error term"? Is it inflation? Please refer to the section where it is defined and explained? If not define here and/or add the appropriate reference.
Response: The statement has been rewritten for clarity. It now reads: "The spread in the CO ensemble at any grid point is due the perturbations in the flux and those in winds". (lines 328 → 337-338).

Lines 332-333: This is statistically expected considering Gaussianity and the truth being drawn from the prior distribution itself. Please modify the statement accordingly.
Response: Though one can make sure that at initial time that the Gaussianity is respected (by drawing from Gaussian flux distribution and initial conditions), one cannot control the extent to which forecast distributions are Gaussian. This is due to the nonlinearities in transport model. Therefore Gaussianity is an assumption that can be violated based on the state (time and location).

Line 337: change "establishes" by "is"
Response: Done (line 337 → line 346).

Line 338-339: The authors should stop recalling what would be the next sub-sections at the end of each sub-sections.
Response: These lines were deleted.

Line 341: Please use more common vocabulary; "trial" is not used in atmospheric data assimilation.
Response: As previously noted, "trial" was replaced by "forecast" throughout the manuscript.

Line 347: what is the matrix inverse. Is it the inverse of P?,H?, R? Or K? Please be more specific.
Response: We have rewritten this section to better focus on the illustration of state dependent correlation and localization.

Lines 351-352: The sentence "The scaling factor: : :" is hard to understand. What is the scaling factor of the innovation? Please define.
Response: We have rewritten this section. The scaling factor is (HP^fH+R). We have dropped this sentence since it is not central to the illustration of distance dependent localization.

Lines 356-357: The statement "In theory a given: : :" is misleading statement. In the theoretical case of a perfect ensemble with an infinite number of members, the spurious correlation would not exist, and you would not need to localise. Hence you would apply the filter globally. Please remove or change accordingly.
Response: "In theory" was meant to indicate exactly this situation of an ensemble with infinite members. However, this is made more explicit in the revised text. (line 356-357 –> 361).

Line 359: The statement "small correlations cannot be trusted" is misleading. A GC localisation is not applied to remove small correlations but spurious correlations that are far from the observation location. Small correlations are not necessarily spurious. This also depends on the ensemble size and nature of the state (e.g. lifetime and transport). Please remove or change accordingly.
Response: We agree with the reviewer that both small and large sample correlations can be spurious. We have removed the sentence. However, small correlations are harder to estimate. The sample correlation coefficient has an uncertainty associated with it. The correlation coefficient can be viewed as an estimator. The pdf of this estimator is complex and depends on the sample size and the true correlation (see attached pages from Hoel's Introduction to Mathematical Statistics, Wiley, 1974). For example, with our sample size of 64, for a true correlation of 0.9, the sample correlation coefficient (r) will be estimated as lying with the range $0.84 < r < 0.94$ with 95% confidence. However for the same sample size of 64, a true correlation of 0.1 will be estimated as lying between $-0.15 < r < 0.34$ to 95% confidence. Note that the uncertainty range for a high correlation is smaller than that for a low correlation (here a range of 0.1 versus 0.49). So, it is indeed harder to correctly estimate small correlations for a given sample size.

Line 363: What the meteorological cut off values? Please detail and/or provide reference.
Response: A reference to the values used is included in section 2.3. Please see lines 186-188 in the revised manuscript.

Line 365: Change "has a peak" to "has its maximum".
Response: Done. (line 365 → 366).

Line 407: Change "blob" to "area".
Response: Done. (line 407 → 411).

Lines 409-410: This is incomplete. The transport of corrected concentration plays a major role as well. I would say this is the combination of both in your case. Please update the text accordingly.
Response: The correlations are a result of the flow-dependent transport. This clarification has been made to the text. We also note the role of downstream transport during the forecasts. Please see lines 414-415 in the revised manuscript.

Line 411: If the surface only concentrations and not the 0-5km column were displayed different results might appear as the observation network is at the surface. Also, it is hard to tell in figure 6 that the RMSE is much lower in Western Canada as this is at the edge of the colour scale. The authors should zoom and adjust the plot to make the point clearer.
Response: This is a very good suggestion. We have included two more panels in figure 10. Panel c shows the RMSE and panel d shows the benefit averaged over 0-1 km. The height of the ECCC stations varies from 5 to 707 meters.

Lines 441-442: Again, this is not only the EnKF but also the transport of corrected concentration by the model itself that improves the RMSE. Please correct the text.
Response: The EnKF includes the model forecasts as part of the algorithm. However, we believe the Reviewer's point is to distinguish between the analysis step and the forecast step, and indeed both are important for transporting information of observations downstream. (Lines 441-442 → lines 447-448).

Line 471: I would disagree with that statement. The vertical information content in the MOPITT retrieval as opposed to HYPNET is not precisely located but spreads across the vertical. So, this is not because the degrees of freedom on the vertical are comparable that the vertical information is similar. Please correct the statement.
Response: The reviewer is correct in noting that the MOPITT information is distributed in the vertical. We have modified the text to state that "HYPNET has information at three vertical levels while MOPITT has an information content with one to two degrees of freedom (Deeter et al., 2012) so that limited vertical information is provided by the two networks." (Lines 471 → lines 464-466).

Line 473: The authors do not show this as a not directly GHG gas has been assimilated. I would suggest removing GHG but change to something as "atmospheric composition" as only CO assimilation has been demonstrated in the paper.
Response: We have rewritten the manuscript to avoid discussing GHGs except to mention our future work. We changed "GHG" to "atmospheric composition". (line 473 → 468).

Lines 484: This is true, but this needs to be reformulated correctly. Please mention atmospheric transport.
Response: We have modified the statement to include transport. (Line 484 → 480-481).

Lines 494-495: I am not convinced this is a conclusion from Miyazaki et al., 2012. CO surface flux errors can be correlated with other fields if you consider the co-emission of different species through a given sector. Remove or change the statement accordingly.
Response: We have deleted this paragraph since it deals with flux estimation and is therefore more relevant to a paper which deals with flux estimation.

Line 506: The authors did not show anything about flux inversions. Please remove this statement.
Response: We have removed this statement.

Line 508: Please define smoother. Add a reference. Use the book from Bocquet et al., 2016 for definitions of the smoother.
Response: We have defined a smoother. We have added two references – Liebelt for the definition of smoother and Bocquet, 2016 for the formulation of the smoother we want to develop. See lines 494-495 in the revised manuscript.

Line 509: Future observations? That do not exist unfortunately... Please remove or correct.
Response: The statement has been clarified to indicate that a smoother uses observations later than the time of the analysis as well as those from earlier than the analysis time. See line 494-495.

Figure 7: What do the numbers in the x-axis indicate? Please clarify or change
Response: The x-axis labels refer to the date from 28 Dec. 2014 to 28 Feb 2015. The figure has been revised. Figure 5 has been similarly revised.

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.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

The Environment and Climate Change Canada Carbon Assimilation System (EC-CAS v1.0) : demonstration with simulated CO observations

Vikram Khade[1,2], Saroja M. Polavarapu[1], Michael Neish[1], Pieter L. Houtekamer[1], Dylan B.A. Jones[2], Seung-Jong Baek[1], Tailong He[2], and Sylvie Gravel[1]

[1]Environment and Climate Change Canada, 4905 Dufferin Street, Toronto, Canada, M3H 5T4
[2]Department of Physics, University of Toronto, 60 St. George Street, Toronto, Canada, M5S 1A7

Correspondence: Vikram Khade (vikram.khade@canada.ca)

Abstract. In this study, we present the development of a new coupled weather and greenhouse gas (GHG) data assimilation system based on Environment and Climate Change Canada's (ECCC's) operational Ensemble Kalman Filter (EnKF). The estimated meteorological state is augmented to include three chemical constituents: CO_2, CO and CH_4. Variable localization is used to prevent the direct update of meteorology by the observations of the constituents and vice versa. Physical localization is used to damp spurious analysis increments far from a given observation. Perturbed flux fields are used to account for the uncertainty in CO due to error in the fluxes. The system is demonstrated for the estimation of 3-dimensional CO states using simulated observations from a variety of networks. First, a hypothetically dense uniformly distributed observation network is used to demonstrate that the system is working. More realistic observation networks based on surface hourly observations, and space-based observations provide a demonstration of the complementarity of the different networks and further confirm the reasonable behaviour of the coupled assimilation system. Having demonstrated the ability to estimate CO distributions, this system will be extended to estimate surface fluxes in the future.

1 Introduction

Environment and Climate Change Canada (ECCC) operates a ~~GHG~~ measurement network which has seen rapid expansion during the past decade. ECCC also possesses a GHG inventory reporting division. As required by United Nations Framework

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

on Climate Change (UNFCCC) commitments, Canadian emissions are quantified and reported using bottom-up methods (NIR 2019, http://www.publications.gc.ca/site/eng/9.506002/publication.html). In order to assess the national impact of mitigation efforts, knowledge of the natural sources and sinks is also needed. The challenge is that there are huge uncertainties in the natural carbon budget for Canada. For example, Crowell et al. (2019) find a range of uncertainty estimates of Boreal North American (which is primarily Canada plus Alaska) fluxes from 480 to 700 TgC yr^{-1} for an ensemble of inversion results for 2015 or 2016. This uncertainty in the biospheric uptake is comparable to the NIR estimate of anthropogenic emissions (568 and 559 TgC yr^{-1} for 2015 and 2016 respectively) for Canada. In addition, there is much unknown about the fate of carbon stored in the permafrost under a warming climate (Voigt et al., 2019), and this will have implications for the global as well as the Canadian carbon budget. Thus, ECCC has a need to better understand and quantify GHG sources and sinks on the national scale. The ECCC Carbon Assimilation System (EC-CAS) was proposed to address these needs using the available tools, namely, operational atmospheric modelling and assimilation systems. The goal of EC-CAS is to characterize CO_2, CO and CH_4 distributions and fluxes both globally and over Canada with a focus on the natural carbon cycle. An important aspect of the impact of climate change on boreal forests is the influence of wildfires on the carbon balance in these regions. Over the past several decades there has been an increase in the frequency of wildfires and this trend is expected to continue (Abatzoglou and Williams, 2016; Flannigan et al., 2009), which will have a significant impact on the Canadian carbon budget and on the Canadian economy. It is for this reason that EC-CAS also includes CO alongside the greenhouse gases CO_2 and CH_4.

Carbon Monoxide (CO) plays a role in both tropospheric chemistry and in climate. In terms of air quality, CO is an important precursor of tropospheric ozone, but it is also a by-product of incomplete combustion and thus correlates well with anthropogenic sources of greenhouse gases from fossil fuel and biofuel burning and from forest fires. CO has a lifetime of 1-2 months which is in-between the weather and climate timescales and thus data assimilation systems (DAS) that assimilate CO can focus on either the air quality or the climate problem. Tropospheric pollution prediction concerns short time scales (forecasts up to 5 days) whereas climate problems concern the estimation of surface fluxes over months to years. Those data assimilation systems whose primary objective is to better understand and predict tropospheric pollution typically use a coupled weather and chemistry model with short assimilation windows (e.g. 12 h) to initiate short forecasts. The CO observations are used to estimate CO initial states for the forecasts with either an Ensemble Kalman Filter (EnKF) (Barré et al., 2015; Miyazaki et al., 2012) or a 4-d Variational (4D-Var) approach (Inness et al., 2019, 2015). The chemistry model typically includes the numerous gas phase and aerosol reactions relevant for air quality. On the other hand, systems focused on CO's influence on climate are typically "inversion systems" wherein observations of CO concentrations are used to estimate CO surface fluxes. Here again, both ensemble (Miyazaki et al., 2015, 2012) and variational (Jiang et al., 2017, 2015a, b, 2013, 2011; Fortems-Cheiney et al., 2011) approaches have been used. Typically a chemistry transport model (CTM) driven by offline meteorological analyses is used. Simplified chemistry models with monthly hydroxide (OH) fields (Yin et al., 2015; Fortems-Cheiney et al., 2011; Jiang et al., 2017, 2015a, b, 2013, 2011) or full tropospheric chemistry models (Miyazaki et al., 2015, 2012) may be used.

As noted above, the focus of EC-CAS is ultimately on flux estimation of GHGs thus inverse modelling approaches are appropriate. However, with this approach, the mismatch between model-predicted CO and observed CO is usually used to adjust

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

fluxes only though some systems also adjust CO initial conditions (eg. Fortems-Cheiney et al. (2011)). In reality, the mismatch between modelled and observed CO is due to errors in surface fluxes, meteorological analyses, initial conditions, model formulation, representativeness and measurements. Proper attribution of the model-data mismatch requires accurate characterization of all of these sources of errors. Also, flux estimates from inverse models are sensitive to the choice of meteorological analyses (Jiang et al., 2011), errors in convective mass transfer (Jiang et al., 2013, 2011; Ott et al., 2011; Arellano and Hess, 2006), boundary layer mixing (Arellano and Hess, 2006), biogenic sources (Jiang et al., 2011), aggregation error (Jiang et al., 2011) and OH climatology errors (Jiang et al., 2015a, b, 2011). The presence of these model and meteorological forcing errors confounds the retrieved CO flux estimates. With a coupled weather and chemistry transport model, uncertainty in the meteorological analyses can be accounted for (Barré et al., 2015), and with an ensemble Kalman filter approach, model errors can be directly simulated (Miyazaki et al., 2012) . Flux estimation with an EnKF using a weather/GHG model was demonstrated for CO_2 by Liu et al. (2011) and Kang et al. (2012, 2011).

EC-CAS will adapt the operational Ensemble Kalman Filter (EnKF) (Houtekamer et al., 2014) to perform a coupled meteorology, GHG state and GHG flux estimation using the approach of Liu et al. (2011); Kang et al. (2012, 2011). EC-CAS directly simulates and accounts for all components of transport error (i.e. errors arising from model formulation, meteorological state, GHG initial conditions) as well as observation and flux errors. See Polavarapu et al. (2016) for a detailed discussion of transport errors. EC-CAS will also be able to handle the vast quantities of observations that are anticipated since currently, roughly 10^6 observations are already assimilated every day for weather forecasts. The main drawback is that EC-CAS is computationally expensive. In contrast, inverse modelling techniques are much more expedient especially for multi-year or decadal flux estimates. However, inverse model results are known to be strongly model dependent (Chevallier et al., 2014, 2010; Houweling et al., 2010; Law et al., 1996) due to the inability to adequately characterize transport errors. Since the EC-CAS flux estimation approach has not yet been demonstated with operational weather forecast models, the development of EC-CAS is proceeding in stages. The first stage is to assess the state estimation component since if it works it will give confidence about the ensemble-based CO flux estimation. The second stage is to estimate the fluxes with this ensemble. Since each of the three EC-CAS species has its own assimilation challenges, the testing of the assimilation scheme began with CO because of its shorter lifetime of about 2 months. Thus the computational cost to test the new system could be minimized. The system of Barré et al. (2015) and Gaubert et al. (2016) is similar to EC-CAS in utilizing an EnKF with a coupled atmospheric and chemistry transport model for flux estimation. Those articles demonstrate only CO state estimation (as in this work), as they too are proceeding in stages.

In the present paper we introduce the first version of EC-CAS to demonstrate the extension of the ensemble Kalman filter (Houtekamer et al., 2014) to estimate GHGs. This new coupled meteorological and GHG assimilation system is called EC-CAS v1.0. To demonstrate that the system is working, 3-dimensional CO fields are estimated by assimilating observations from four different networks. The outline of the paper is as follows. Section 2 describes the various components of EC-CAS system. Section 3 presents the experimental design while section 4 describes the data assimilation (DA) experiments and their results. Section 5 presents the conclusions of this work and delineates planned future developments of EC-CAS.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

2 System description and development

~~Any data assimilation system consists of a model which produces the trial fields and a statistical technique which blends these trial fields with the observations. The first stage is referred to as the "forecast step" while the second one is called the "analysis step". This section describes these components for EC-CAS starting with an overview of the whole system in section 2.1. The forecast model, the EnKF and extensions for GHG and flux estimation are presented in sections 2.2, 2.3 and 2.4 respectively.~~

2.1 EC-CAS

The EC-CAS system consists of a coupled weather and GHG transport model as the forecast model and an Ensemble Kalman filter (EnKF) as the data assimilation technique. Figure 1 shows a schematic overview of EC-CAS. ~~The model is initialized with N realizations of meteorological variables at every grid point.~~ The 6 h ensemble forecasts are used as the ~~trial~~ fields at each data assimilation (DA) cycle and the spread of the ensemble about their mean defines the forecast error covariances. ~~The EnKF is used to blend the trial fields with the observations to produce the analysis ensemble at each DA cycle. This blending uses the error in observations, the uncertainty in trial fields along with the correlations within the trial fields. These correlations are estimated by the sample correlations of the trial ensemble. The next section 2.2 describes the coupled meteorological and GHG transport model used in the forecast step while section 2.3 delineates the meteorological component of the EnKF. Section 2.4 describes the extension of the EnKF to include estimation of CO and other GHGs.~~

2.2 The forecast model

The forecast model used in EC-CAS is called GEM-MACH-GHG (Polavarapu et al., 2016; Neish et al., 2019). This model is a variant of GEM (Global Environmental Multiscale), ECCC's operational weather forecast model (Côté et al., 1998a, b; Girard et al., 2014) that was developed for the simulation of greenhouse gases. A detailed description of the GEM-MACH-GHG model is found in Polavarapu et al. (2016), so only a few salient points are mentioned here along with recent model updates. Compared to the operational global weather forecast model, GEM-MACH-GHG uses a lower resolution with 0.9° grid spacing in both latitude and longitude and 80 vertical levels ~~and the same lid of 0.1° hPa~~. The vertical coordinate is a type of hybrid terrain-following coordinate (Girard et al., 2014). The advection scheme uses a semi-Lagrangian approach for both meteorology and tracers. Modifications were implemented to conserve tracer mass on the global scale (see Polavarapu et al. (2016)). This included defining tracer variables as dry mole fractions. In addition, tracers are transported through the Kain-Fritsch deep convection scheme (Kain and Fritsch, 1990; Kain, 2004) but not through a shallow convection scheme. The boundary layer scheme uses a prognostic turbulent kinetic energy (TKE) equation to specify the thermal eddy diffusivity (see McTaggart-Cowan and Zadra (2015)). In Polavarapu et al. (2016, 2018), it was necessary to impose a minimum ~~value~~ of 10 $\mathrm{m^2 s^{-1}}$ in the boundary layer. However, recent model improvements enabled the minimum ~~value~~ to be lowered to 1 $\mathrm{m^2 s^{-1}}$ as

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

in Kim et al. (2020). An operational air quality forecast model based on GEM is used to produce 48 h forecasts of air quality health index on a limited area domain covering most of North America. This model is called GEM-MACH (Anselmo et al., 2010; Gong et al., 2015; Pavlovic et al., 2016) and it employs moderately detailed parameterizations of tropospheric chemistry using 42 gas-phase species, 20 aqueous-phase species, and nine aerosol chemical components. In contrast, GEM-MACH-GHG removes the tropospheric chemistry module entirely for CO_2 and replaces it with simple parameterized chemistry for CH_4 and CO. This is because GEM-MACH-GHG is used for multi-year simulations and flux estimations of long-lived constituents with an EnKF so the computational expense of complete chemistry is prohibitive and difficult to justify for a system focused on GHG fluxes. However, other model processes from GEM-MACH are used in GEM-MACH-GHG, namely, the vertical diffusion and emissions injection. In GEM-MACH-GHG, the methane (CH_4) parameterization involves a single loss rate with a monthly [OH] climatology. The rate constant is specified following the JPL (2011) formulation for bimolecular reaction for methane (see their page 1-12). The parameterized chemistry model used for CO is identical to that used in GEOS-Chem (http://geos-chem.org) in that CO destruction is parameterized following JPL (2011). The same [OH] climatology is used for CH_4 and CO. Specifically, the [OH] monthly climatology is from Spivakovsky et al. (2000) regridded to GEM-MACH-GHG's grid. The production of CO from CH_4 is computed assuming each methane molecule destroyed becomes a CO molecule. For the CH_4 simulation, the fluxes were obtained from CT-CH_4 (Bruhwiler et al., 2014). Since CT-CH_4 fluxes are available from 2000-2010, the last 5-year mean (2006-2010) fluxes were used as the fluxes for the 2015 EC-CAS simulation. The initial condition (IC) for CH_4 for 1 January 2015 was approximated with the CH_4 atmospheric mole fractions from CT-CH_4 at the end of 2010 plus a globally uniform offset to account for the increase in CH_4 from 2010 to 2015 (30 ppb, estimated based on the difference from observations at the South Pole). Even though the initial condition is not correct, the impact of the errors in the CH_4 initial condition (the synoptic spatial patterns) dissipates within weeks. These prescribed CH_4 fluxes and initial conditions appear reasonable as the model simulated CH_4 compares well with surface observations. To define the CO initial state, an inversion constrained by space based observations from MOPITT (Measurement of Pollution in the Troposphere) instrument v7J (Drummond, 1992) was performed with GEOS-Chem on a $4^\circ \times 5^\circ$ grid. The CO combustion emissions are from Hemispheric Transport of Air Pollutants (http://www.htap.org) (Janssens-Maenhout et al., 2015). Biogenic emissions of isoprene, methanol, acetone, and monoterpenes are from a GEM-MACH simulation, with an assumed yield of CO from the oxidation of these hydrocarbons that is based on the GEOS-Chem CO-only simulation employed in Kopacz et al. (2010) and Jiang et al. (2011, 2015a, 2017). The monthly CO posterior fluxes obtained for December 2014 and throughout 2015 were used in EC-CAS EnKF cycles. Since GEOS-Chem is widely used for assimilation of MOPITT CO data, we use a posterior CO distribution from GEOS-Chem for 1 December 2014 18:00:00 UTC as the initial state on 27 December 2014 18:00:00 UTC.

2.3 EnKF equations for meteorology

The Kalman Filter equation (Ghil et al., 1981; Cohn and Parrish, 1991) at a particular DA cycle at time t is given by,

$$\mathbf{x}^a(t) = \mathbf{x}^f(t) + \mathbf{P}^f(t)\mathbf{H}^T(t)\left[\mathbf{H}(t)\mathbf{P}^f(t)\mathbf{H}^T(t)+\mathbf{R}(t)\right]^{-1}\left(\mathbf{y}^o(t)-\mathbf{H}(t)\mathbf{x}^f(t)\right) \quad (1)$$

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

In this equation $\mathbf{x}^f$ and $\mathbf{x}^a$ are state vectors of dimension $d = 400 \times 200 \times (80 \times 4 + 2)$. The dimensionality of the model grid is 400×200. The number of vertical levels is 80. There are four 3-dimensional meteorological variables, namely temperature, two components of winds and humidity. In addition, the state vector includes the 2-dimensional fields of surface pressure and radiative temperature at the surface. $\mathbf{x}^f$ is the ~~trial~~ field produced by a 6 hour forecast of GEM-MACH. $\mathbf{x}^a$ is the analysis produced by combining the information content in the ~~trial~~ fields and the observations. $\mathbf{P}^f$ is the forecast error covariance matrix calculated using the spread of the forecast ensemble about its mean. $\mathbf{y}^o$ is the observation vector of dimension m and matrix $\mathbf{R}$ of dimension $m \times m$ represents the error in $\mathbf{y}^o$. $\mathbf{H}$ is the forward operator which maps the model state to the observation space.

The quantity,

$$\mathbf{K}(t) \quad = \quad \mathbf{P}^f \mathbf{H}^T \; [\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R}]^{-1} \tag{2}$$

in equation 1 is known as the Kalman gain, where the t in the parenthesis for quantities on the right side are dropped for readability.

$(\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R})$ is a huge matrix (order $\approx 10^6$) (Houtekamer et al., 2019) and its inversion is computationally onerous. This problem is circumvented by solving this equation *sequentially* (Cohn and Parrish, 1991; Anderson, 2001; Houtekamer and Mitchell, 2001). In sequential processing, the total number of observations m are subdivided into N_b subsets, known as *batches* containing at most p observations each.

Then, the assimilation proceeds as follows :

$$\begin{aligned}
\mathbf{x}_1^a(t) &= \mathbf{x}^f(t) + \mathbf{P}^f \mathbf{H}_1^T [\mathbf{H}_1 \mathbf{P}^f \mathbf{H}_1^T + \mathbf{R}_1]^{-1} (\mathbf{y}_1^o - \mathbf{H}_1 \mathbf{x}^f) && \textit{Pass 1} \\
\mathbf{x}_2^a(t) &= \mathbf{x}_1^a(t) + \mathbf{P}^f \mathbf{H}_2^T [\mathbf{H}_2 \mathbf{P}^f \mathbf{H}_2^T + \mathbf{R}_2]^{-1} (\mathbf{y}_2^o - \mathbf{H}_2 \mathbf{x}_1^a) && \textit{Pass 2} \\
&\vdots \\
\mathbf{x}_{N_b}^a(t) &= \mathbf{x}_{N_b-1}^a(t) + \mathbf{P}^f \mathbf{H}_{N_b}^T [\mathbf{H}_{N_b} \mathbf{P}^f \mathbf{H}_{N_b}^T + \mathbf{R}_{N_b}]^{-1} (\mathbf{y}_{N_b}^o - \mathbf{H}_{N_b} \mathbf{x}_{N_b-1}^a) && \textit{Pass } N_b
\end{aligned}$$

The subscripts $1, 2, ..., N_b$ represent the pass numbers. $\mathbf{x}_{N_b}^a(t)$ is the updated state (as if all the observations were processed simultaneously). The analysis from a given pass is used as the ~~trial~~ field in the next pass. At each pass at most 600 observations are assimilated.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

Though the covariance estimate $\mathbf{P}^f$ obtained from the ensemble is state dependent, owing to the small size of the ensemble this estimate is noisy. This is remedied by the use of physical localization. The Kalman gain in equation 2 is modified as,

$$\mathbf{K}(t) \quad = \quad (\rho_{\mathbf{m}} \circ (\mathbf{P}^f \mathbf{H}^T)) \; [\rho_{\mathbf{o}} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}]^{-1} \tag{3}$$

where $\rho_{\mathbf{m}}$ and $\rho_{\mathbf{o}}$ constitute the localization in the model space and observation space, respectively and $\circ$ denotes the Hadamard product. These matrices contain weights that smoothly decrease towards zero as the distance from the observation increases. The localization in the model space ($\rho_{\mathbf{m}}$) requires the distance between observations and model coordinates while the localization in the observation space ($\rho_{\mathbf{o}}$) requires the distance between observations (Houtekamer et al., 2016). The covariances in $\mathbf{P}^f \mathbf{H}^T$ are multiplied elementwise by $\rho_{\mathbf{m}}$. Similarly the covariances in $\mathbf{H}\mathbf{P}^f \mathbf{H}^T$ are multiplied elementwise by $\rho_{\mathbf{o}}$. The ensemble size is typically much smaller than the dimensionality of the model. For example, in this work the ensemble size is 64 while the model dimensionality is $\approx 10^7$. Consequently, the correlation estimate calculated from the ensemble can be spurious. Localization is designed to ameliorate this problem of spurious correlations (Hamill et al., 2001). The rate of decrease of the weight is dictated by the Gaspari-Cohn function (Gaspari and Cohn, 1999; Houtekamer and Mitchell, 2001).

2.4 EnKF extensions for ~~GHG and fluxes~~

The state vector discussed in section 2.3 is augmented to include the CO, CO_2, CH_4 fields and their fluxes. This state is referred to as the augmented state. Variable localization (Kang et al., 2011) is implemented in the EnKF code by modifying equation 3 as follows.

$$\mathbf{K}(t) \quad = \quad (\rho^{\mathbf{v}}_{\mathbf{m}} \circ \rho_{\mathbf{m}} \circ (\mathbf{P}^f \mathbf{H}^T)) \; [\rho^{\mathbf{v}}_{\mathbf{o}} \circ \rho_{\mathbf{o}} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}]^{-1} \tag{4}$$

Each element of $\rho^{\mathbf{v}}_{\mathbf{m}}$ and $\rho^{\mathbf{v}}_{\mathbf{o}}$ is either 1 or 0. Unlike the physical localization matrices the elements of variable localization matrices are not distant dependent; they are rather *variable type* dependent. A given element is 1 ~~when~~ the row and column variable is of the same type and 0 otherwise. The (i,j) th element of $\rho^{\mathbf{v}}_{\mathbf{m}}$ and $\rho^{\mathbf{v}}_{\mathbf{o}}$ is set to one if one desires an observation of the jth variable to impact the update of ith variable. Setting the (i,j) th element to zero ensures that the observation of the jth variable does not contribute to the update of the ith variable. For example ~~when~~ the both the row and column of $\mathbf{H}\mathbf{P}^f \mathbf{H}^T$ correspond to a CO observation, ~~that element~~ of $\rho^{\mathbf{v}}_{\mathbf{o}}$ is set to 1. In our initial implementation of EC-CAS, presented in this work, variable localization is implemented such that meteorological observations do not directly update CO state and CO observations do not update the meteorological state as in Inness et al. (2015). Since Miyazaki et al. (2011) and Kang et al. (2012) show that CO_2 updates through wind observations are beneficial, this issue will be considered in future EC-CAS developments. It is worth noting that it is still possible for the CO state to indirectly improve due to the assimilation of wind observations. This can occur because improvement in winds due to meteorological observations leads to an improvement in the

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

spatial distribution of CO. ~~The spatial correlation between the CO at the location of an observation and that at an unobserved location plays a key role in the update.~~

~~While the operational EnKF (Houtekamer et al., 2019) currently uses an ensemble size of 256, EC-CAS uses 64 ensemble members. The main reason for the reduction in ensemble size is to reduce computational cost during the development of EC-CAS. Increases in ensemble size are envisioned in the future. Other EnKF systems used for CO state or flux estimation have used 30 ensemble members (Gaubert et al., 2016; Barré et al., 2015; Miyazaki et al., 2015). In EC-CAS radiosondes, surface stations, ships, aircraft and cloud-drift wind observations are used to constrain the meteorological variables.~~

~~The operational EnKF uses an Incremental Analysis Updating (IAU) scheme (Bloom et al., 1996) to control high frequency waves generated by analysis insertion during the ensemble forecasts. This scheme is also applied in EC-CAS to meteorological and GHG analysis increments. To simulate model errors, the operational EnKF uses different model parameters for different ensemble members. These parameters are associated with the most uncertain parameterized physical processes such as boundary layer turbulence and deep convection. Each ensemble member is assigned a unique combination of optional values for ten such parameters. Since the same sources of meteorological forecast error also impact CO_2 transport error (i.e. synoptical scale signals (Parazoo et al., 2008; Chan et al., 2004)), the same set of parameters are perturbed in EC-CAS. In the case of CO, the flux estimation results are also sensitive to model errors in convective mass transfer (Jiang et al., 2013, 2011; Ott et al., 2011; Arellano and Hess, 2006) and boundary layer mixing (Arellano and Hess, 2006). The operational EnKF also uses an additive homogeneous, isotropic climatological error produced using the so-called NMC method (Bannister, 2008; Parrish and Derber, 1992) for additional model error simulation. In EC-CAS, for the meteorological assimilation, the same scheme is used, but for GHGs, no such additive error is present. It is not needed for the tests with simulated observations which only seek to characterize system behaviour. However, for the flux estimation extensions (currently under development) an inflation scheme will be implemented. The form of the inflation scheme is yet to be determined.~~ The EnKF code with the extensions developed here is available in Khade et al. (2020).

3 Experimental Design

The current paper describes the experiments run for the development and testing of state estimation of CO using simulated observations. The ultimate goal is to develop a system that ingests real GHG observations. However, testing of the system using simulated observations is an important milestone. It not only demonstrates that the system is working properly but also illustrates and defines errors achievable in the best case scenario (under idealized conditions). In this work we use simulated CO observations which are unbiased and have uncorrelated errors.

The experiments in this work are run from 27 December 2014 18:00:00 UTC to 28 February 2015 00:00:00 UTC. At cold start 65 perturbations of meteorological variables are drawn from the same climatological (static) covariance matrix that was

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

used to generate additive model error. These 65 perturbations are added to the meteorological base state valid at 27 December 2014 18:00:00 UTC. This produces 65 ensemble members for the meteorological variables. Out of the 65 ensemble members the 65^{th} ensemble member is designated as the *truth*. The remaining 64 ensemble members are used for the EnKF estimation experiment. With this approach, the truth member is a plausible member of the ensemble having been generated from the same probability density function. The meteorological observations are drawn from the trajectory of the *truth* at 00, 06, 12 and 18Z every day at which time DA is carried out. The observation networks used in this work are described in section 3.2.

An important facet of this work is accounting for flux error in the CO estimate. This is accomplished through the use of a flux ensemble. ~~The following section 3.1 describes the construction of the flux ensemble while the observations used are discussed in section 3.2.~~

3.1 Flux perturbation

The error in flux is an important source of error in the CO estimate. This is especially true close to the surface. For example, biases in CO analyses near the surface in polluted urban areas were attributed to emissions errors (Inness et al., 2013). Here flux error is simulated using perturbed flux fields. The posterior of the 4D-Var based GEOS-Chem inversion constrained by MOPITT observations is used in this work as the *truth*. These posterior flux fields are constant over a period of one month (see Figure 2).

The flux ensemble, of size 64, is generated by using a spectral algorithm (See Appendix A of Mitchell and Houtekamer (2000)). The flux perturbations are generated such that they are spatially correlated over a distance (half width) of 1000 km. The standard deviation of the spread is set to 40% of the value of the true flux as in Barré et al. (2015). The true flux field is used by the 65^{th} ensemble member which generates the truth trajectory. Two sets of 64 flux ensemble members are generated, ~~one each for January and February 2015~~. Each member of the flux ensemble is used with each (distinct) member of the meteorology ensemble.

3.2 Observation networks

The families of meteorological observations used in this work are summarized in Table 1. The location and times of these observations are *real* though the observation values are *simulated*. These meteorological observations are assimilated in all the experiments presented in this work.

The EnKF is tested with five different CO observational networks. These are summarized in Table 2. The first network, HYPNET is a hypothetical network. It has spatially dense coverage of in situ observations. In this network the observations are located every 1000 km on three planes –1 km, 5 km and 9 km. These heights are with respect to the local topography. The

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

Type	0000 UTC	0006 UTC	0012 UTC	0018 UTC
Upper air	54765	3471	51508	2057
Aircraft	78780	54065	56708	78829
Satellite winds	29749	32813	33597	31719
Surface	9927	10314	10271	9984
Scatterometer	18221	17484	19462	16782
GPS-RO	7064	5564	6390	6527

Table 1. Columns shows the typical number of meteorological observations assimilated in each 6 h DA cycle.

Network	Spatial coverage	Temporal coverage
HYPNET	every 1000 km at levels near 1,5,9 km	every 6 hours
ECCC surface	17 stations (Canada only)	hourly
GAW surface	44 stations	hourly
NOAA surface	69 stations	$\sim$ weekly
MOPITT	1 retrieval per 100 km	Global coverage every 3 days.

Table 2. CO observation networks used in this work.

globally averaged topography is 376 m. Therefore, the heights of these planes with respect to mean sea level are roughly 1.376 km, 5.376 km and 9.376 km. The locations of observations are shown in Figure 3a.

The ECCC surface network (Worthy et al., 2005) consists of 17 observing stations in Canada (see Figure 3c, 3d and Table A1). These stations provide measurements at an hourly frequency. Although the ECCC network has expanded rapidly in the past decade to 25 sites in 2020, only the 17 sites providing hourly measurements in 2015 are simulated here. GAW is an acronym of Global Atmospheric Watch (https://gaw.kishou.go.jp). The 44 stations from the GAW network used in the current work are shown in Figure 3c and listed in Table A2. These stations observe at an hourly frequency. The NOAA surface observation network consists of 69 flasks (see Figure 3b). The observations from these flasks are temporally sparse, averaging approximately one per week.

MOPITT (Measurement of pollution in the Troposphere) (Drummond, 1992) is an instrument onboard NASA's Earth Observation Satellite (EOS) Terra that was launched in December 1999. MOPITT is an important component of the global CO observing system because it measures spectra both in the Near InfraRed (NIR) and Thermal InfraRed (TIR) so that its retrieved profiles are sensitive to CO in the lower troposphere where the flux signal from emissions is most readily detected. That is why it is assimilated in almost all CO data assimilation systems whether their focus is on air quality or flux estimation. It has a nadir footprint of 22×22 km and a 612 km cross-track scanning swath. Its orbit repeats every 3 days. We used V7J MOPITT data

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

with locations thinned to one observation per grid box. The coverage on a particular day is shown in Figure 4. The MOPITT averaging kernel is used to construct the MOPITT observation operator. The MOPITT prior denoted by $\mathbf{y}^{pr}$ and the MOPITT CO retrieval denoted by $\mathbf{y}^{obs}$, are both vectors of dimension 10. Both $\mathbf{y}^{obs}$ and $\mathbf{y}^{pr}$ are defined on 10 levels namely, 1000 hPa, 900 hPa, 800 hPa, ... 100 hPa. In case the surface is not at 1000 hPa, the observations below the surface are ignored.

The averaging kernel, denoted by $\mathbf{A}$, is a matrix of dimension 10×10. The averaging kernel is the sensitivity of retrieval at each level to all the levels. The observation operator, also known as the forward operator for MOPITT is given by,

$$\mathbf{Hx} = \mathbf{y}^{pr} + \mathbf{A}(\mathbf{x}^{gem} - \mathbf{y}^{pr}) \quad (5)$$

$\mathbf{x}^{gem}$ is the model profile interpolated to the same levels on which the $\mathbf{y}^{pr}$ is defined. The assimilation of MOPITT NIR/TIR retrievals is performed as described in Jiang et al. (2015a).

In the HYPNET, GAW, ECCC and NOAA networks the observation operator is an interpolation operator. The HYPNET, GAW, ECCC and NOAA network are temporally static whereas the MOPITT observations change locations depending on the particular DA cycle. The observation errors are set to 10% of the observation values for all networks in our work. The validation of the MOPITT V7J data found that the standard deviation of the retrieved profiles varied between 10-16 % relative to independent data (Deeter et al., 2017).

In summary, five different observation networks are simulated out of which four have a significant impact on the CO estimation error. The HYPNET is a hypothetically dense network with uniform spatial coverage and some vertical coverage. This network is useful for identifying coding or other issues with the data assimilation algorithm, since with plentiful, accurate observations and a perfect model a well tuned assimilation scheme should work. The other networks exist in reality and are used to test the assimilation code in a more realistic, but still controlled settings (since no observation biases exist and uncertainty levels are known). Degradation of results relative to the hypothetical network is expected. Our focus is on obtaining a qualitative understanding of the behaviour of the assimilation system in less idealistic settings. It should be noted that although the observations are simulated here, we are not performing Observing System Simulation Experiments (OSSEs) (See wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2018_8.pdf for a discussion of designing OSSEs). OSSEs (Prive et al., 2018) are used to compare the results obtained with different observing systems and require careful configuration and tuning of assimilation system parameters so that conclusions might be quantitatively reasonable. Such tuning is difficult with a system that is just being developed. Thus, to reiterate, our results from using different simulated observation networks serve only as a testbed for understanding the new assimilation system.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

4 Results

This section discusses the improvement in the CO state due to the assimilation of HYPNET, surface observations and MOPITT retrievals, in four seperate experiments. The four CO data assimilation experiments are denoted by *EXP_HYP*, *EXP_GAW*, *EXP_NOAA* and *EXP_MOP*. This improvement is defined with respect to a control experiment which is referred to as *EXP_CNTRL*. *EXP_GAW* assimilates the GAW and ECCC surface observations while *EXP_NOAA* assimilates the NOAA and ECCC surface observations. ~~This control experiment~~ assimilates simulated meteorological observations (see Table 1) but does not assimilate simulated CO observations. The CO data assimilation experiments assimilate the same meteorological observations as assimilated by *EXP_CNTRL* in addition to their CO observations. The results of the *EXP_CNTRL* are discussed in section 4.1. Section 4.2 illustrates the role of dynamically changing spatial correlations in an EnKF update. ~~An ensemble forecast contains state-dependent correlation information, which to a large extent is dictated by the wind field. In an EnKF this state-dependent correlation is used to spread observational information to unobserved locations.~~ The results of the four CO data assimilation experiments are described in section 4.3. ~~All experiments are run from 27 December 2014 18:00:00 UTC to 28 February 2015 00:00:00 UTC.~~

4.1 Control experiment

Before delving into the results of the CO data assimilation it is important to examine the results of the meteorological variables. CO is advected by the winds and hence it is critical to ensure that assimilation of meterological observations is working well. Figure 5 shows the timeseries of ~~area weighted~~ temperature ~~trial~~ and analysis root mean square error (RMSE) from 27 December 2014 18:00:00 UTC to 28 February 2015 00:00:00 UTC in *EXP_CNTRL* experiment. The RMSE is calculated based on the error between the ensemble mean and the truth which is available at every grid point. As observations are assimilated the RMSE decreases and stabilizes to $0.5°C$ in about 7 days. ~~This timescale is expected since the predictability of weather is about 7-10 days (Pires et al., 1996). After day 7, the errors in synoptic scale motions have saturated at their climatological values. The results for other meteorological variables (not shown) were also examined and it is ascertained that the meteorological data assimilation is working as expected~~.

The column mean of the ensemble mean of CO averaged over the 7 week assimilation period is shown in Figure 6a. The values of CO are clearly higher in regions of high flux (see Figure 2). The CO from central Africa is advected to equatorial Atlantic by the easterly winds. Figure 6b shows the ensemble spread of the CO analysis which is estimated by the standard deviation about the analysis mean. This quantifies the EnKF expected error in the CO mean. The spread in the CO ensemble at any grid point is due to perturbations in the flux~~,~~ spread in the winds ~~and spread obtained by using different realizations of the physics parametrizations within the ensemble and the additive model error term.~~ Figure 6c shows the CO analysis RMSE which quantifies the actual error between the ensemble mean and truth. Clearly, comparing Figure 6a and 6c the RMSE is higher in regions of higher values of CO. As noted earlier, the regions with high CO correspond to regions of large flux.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

The similarity of the spatial pattern of CO ensemble spread (Figure 6b) and the RMSE (Figure 6c) is encouraging because it indicates that the DA system is simulating the actual error well with 64 ensemble members.

The ~~trial~~ and analysis RMSE are identical because CO observations are not assimilated in *EXP_CNTRL*. The time series of RMSE over the ~~period of experimentation~~ is shown by the blue curve in Figure 7a. The RMSE in January 2015 stabilizes to about 16 ppb. The flux field changes in February 2015 and hence the RMSE enters a different regime starting on $\sim$ 1 February 2015.

The RMSE of the control experiment ~~establishes~~ a baseline against which the RMSEs from CO data assimilations are compared. ~~These comparisons are presented in the section 4.3. The sample error correlations estimated by the ensemble play an important role in ensemble based data assimilation. This is explained in the next section (4.2).~~

4.2 Role of estimated correlations

The correlations estimated using the ~~trial~~ ensemble plays a key role in spreading the information from a given CO observation to other grid points for any observational network. The correlation estimate changes dynamically depending on the flux perturbations and winds. This state dependence of sample correlation is an important characteristic of ensemble-based filters. The role of the sample correlation and physical localization is illustrated for an observation located at the University of Toronto.

Figure 8 shows the spatial correlation structure for the University of Toronto location at two different times from *EXP_CNTRL*. In equation 1 the term $(\mathbf{y}^o(t) - \mathbf{H}(t)\mathbf{x}^f(t))$ is the *innovation*. This quantity is in the observational space and is thus a scalar for the case of a single observation located at Toronto. Similarly, the matrix inverse in equation 1 is also a scalar. The sample correlation $\mathbf{P}^f(t)\mathbf{H}^T(t)$ is used to map the innovation into model space. This is the correlation between the ensemble of $\mathbf{x}^f(t)$ and $\mathbf{H}(t)\mathbf{x}^f(t)$. For simplicity, we take the nearest gridpoint to Toronto as its actual location so that $\mathbf{H}(t)$ becomes a column vector of the identity matrix with the column index corresponding to the location of the Toronto gridpoint. This renders $\mathbf{P}^f(t)\mathbf{H}^T(t)$ to be the column of $\mathbf{P}^f(t)$ corresponding to the Toronto gridpoint. The scaling factor of the innovation and its uncertainty is neglected since magnitudes are of no concern. The innovation in CO at the University of Toronto location updates the CO at all the other grid points in proportion to the correlation as estimated by the ~~trial~~ ensemble. The regions of high correlation change significantly from 15 January 2015 06:00:00 UTC to 22 January 2015 12:00:00 UTC. Consequently, the impact of the CO observation at University of Toronto on other grid points is different on 15 January 2015 06:00:00 UTC and 22 January 2015 12:00:00 UTC. In theory a given CO observation should update the CO estimate at all other grid points globally. However, due to the physical localization function a given CO observation updates the CO state only within a limited region defined by the horizontal (hlr) and vertical localization radius (vlr). Localization is necessary for EnKF because the small ensemble size (64) will generate sampling noise in correlations. ~~In other words, small correlations cannot be trusted.~~ Correlations at large physical distance must be filtered because they are most likely spurious and would harm the analysis if retained. The values employed in this work are $hlr = 2000$ km and $vlr = 4$ km. Different values of hlr and vlr were tested and

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

it was found that the best results were obtained for $hlr = 2000$ km and $vlr = 4$ km. Note that these values of hlr and vlr are used for CO data assimilation only. The assimilation of meteorological observations uses different values for hlr and vlr. The red circle in Figure 8 has radius of $hlr = 2000$ km. The Gaspari-Cohn function (Houtekamer and Mitchell, 1998) (not shown in the Figure) used for physical localization has ~~a peak~~ at University of Toronto and decays moving away from the observation's location. This function is used as a weight to modulate the correlation values. As a result, the impact of the observation decays with distance from the observation. Distance-dependent localization assumes that the sample correlation given by the ensemble is less trustworthy (that is more spurious) as one moves away from the observation. As noted in section 2.4 variable localization ensures that CO observations do not update meteorological variables. Therefore, the estimates of meteorological variables are the same in all the experiments.

4.3 CO DA experiments

EXP_HYP assimilates HYPNET observations (see section 3.2) in addition to the same meteorological observations assimilated in *EXP_CNTRL*. The HYPNET observations are assimilated starting on 10 January 2015 18:00:00 UTC after a spin up from 27 December 2014 18:00:00 UTC to 10 January 2015 18:00:00 UTC. This spin up period allows time for the meteorological assimilation to stabilize (Figure 5) before the CO data assimilation begins. This spin up also helps the development of correlations within the CO field.

Figure 6d shows the column averaged CO RMSE for the *EXP_HYP* experiment. Compared to the RMSE for the *EXP_CNTRL* experiment the RMSE decreases substantially because HYPNET observations effectively constrain the CO state. The time series of RMSE for the *EXP_HYP* is shown by the red curve in Figure 7a. The blue and red curves overlap from 27 December 2014 18:00:00 UTC to 10 January 2015 18:00:00 UTC during the spin up period. As soon as CO observations are assimilated starting on 10 January 2015 18:00:00 UTC, the RMSE decreases. The reduction in RMSE due to assimilation of HYPNET observations is ~ 7 ppb. This reduction is defined as the *benefit*,

$$benefit \quad = \quad RMSE(control) - RMSE(DA) \tag{6}$$

The *relative benefit* is defined as,

$$relative\_benefit \quad = \quad 100 \times \frac{benefit}{RMSE(control)} \tag{7}$$

The second term in equation 6 is the RMSE of the experiment which assimilates CO observations. Since the *EXP_CNTRL* does not assimilate CO observations, *benefit* measures the value of assimilating CO observations from a particular network. This metric quantifies the extent to which CO observations constrain the CO state. Figure 9a shows the spatial structure of

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

benefit in the *EXP_HYP* experiment. This figure is basically the difference between Figures 6c and 6d. The *benefit* is positive in most parts of the globe except in parts of Tibet and eastern China. A negative value of *benefit* means that assimilation of CO observations increased the RMSE compared to the *EXP_CNTRL*. Negative values can occur because of the statistical nature of data assimilation. However, if the data assimilation system is well tuned, such regions of negative *benefits* should be few and small, as seen here. By comparing the spatial structure of RMSE in Figure 6c and *benefit* in Figure 9a it is clear that the *benefit* is proportional to the RMSE. This makes sense because where the RMSE is large, the observations have a larger scope for improving the CO state estimate. The relative *benefit* (equation 7) in this experiment is 41%. This means that the assimilation of HYPNET observations decreases the control RMSE by 41%.

The time series of RMSE in *EXP_MOP* is shown by the red curve in Figure 7b. The similarity of the amplitudes of the red curves in Figures 7a and 7b indicates, surprisingly, that the global *benefit* of MOPITT data is only a little worse than that due to the hypotetically dense in situ network (HYPNET). The spatial structure of the *benefit* due to the assimilation of MOPITT retrievals (*EXP_MOP*) is shown in Figure 9b. Comparing Figure 9a and Figure 9b, it is evident that the *benefit* due to assimilation of HYPNET observations and MOPITT retrievals, is also quite similar in the column mean, in spite of the different spatio-temporal distribution of observations. The relative *benefit* due to assimilation of MOPITT retrievals is 38%.

Figure 10a shows the *benefit* ~~in~~ *EXP_GAW*. Both the GAW and ECCC networks are assimilated in this experiment. The combined network is temporally dense with observations every hour but is spatially sparse except in Canada and western Europe (Figure 3c and 3d). The relative *benefit* in this experiment is 8%. Figure 10b shows the *benefit* over North America. The bulk of the *benefit* is in the eastern part of this domain though the ECCC stations are located both in eastern and western Canada. The ~~blob~~ of highest *benefit* of about 10-20 ppb is centered on ECCC stations located in Ontario. ~~Though USA does not have any stations in this experiment~~, the *benefit* of observations in Ontario reaches as far as Florida spreading throughout eastern USA. This is because of the spatial correlation between locations of observations in Ontario and the eastern part of USA which is evident in Figure 8. The western part of Canada has a weak *benefit* inspite of having several observation sites in this region. This is because the RMSE in the western region is substantially lower than that in the eastern region (see Figure 6c). Thus there is little scope for the observations to improve upon the control RMSE. The relative *benefit* over North America is 38%.

Figure 10a shows that assimilation results in significant *benefit* over Europe and parts of central Africa. The *benefit* in Europe is due to high spatial density of GAW stations there. However, with only 5 stations in Africa, a *benefit* of 5-20 ppb in central Africa is produced. Some *benefit* is seen over north eastern China and Malaysia due to GAW stations located in these regions.

The last experiment, *EXP_NOAA*, assimilates the NOAA flask stations in addition to the ECCC surface stations. It is seen (figure not shown) that *benefit* over Canada is same as that seen in *EXP_GAW*. However, globally the NOAA flask stations do not result in any significant *benefit* over any other region. This is because the flask observations are available, on an average, only once a week. This experiment is not discussed further in this work.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

In the case of the assimilation of HYPNET and MOPITT observations (Figures 9a and 9b) many parts of Atlantic, Pacific, Indian and other oceans show significant *benefit*. The CO flux over oceans is practically zero compared to that on land. The HYPNET and MOPITT observations over oceans contribute to the *benefit* on oceans. However, the improvement of the CO state over land also contributes to be *benefit* over oceans. For example any improvement in CO state over central Africa improves the state over tropical Atlantic ocean due to the downwind transport.

In the discussion so far the horizontal and temporal structure of *benefit* was explored. Figure 11 examines the vertical structure of *benefit*. Figure 11a shows the globally averaged profile of the control RMSE and *benefits* from the three CO DA experiments. The average RMSE of the control experiment (*EXP_CNTRL*) peaks close to the surface with a value of 21 ppb. The average *benefit* in the bottom 4 km is ~ 7 ppb in the *EXP_HYP* and *EXP_MOP* experiments, but is only ~ 1 ppb for the *EXP_GAW* experiment. Comparing the shapes of the blue and red curve to the black curve, the *benefit* is proportional to the control RMSE except in the bottom ~ 1 km. The shape of the *benefit* profile is dictated both by the shape of the control RMSE and the location of observations in the particular network. The *EXP_HYP* profile shows a local peak at 1 km. This is because HYPNET observations are located at 1 km. HYPNET observations are also located at 5 km. However the control RMSE decreases by a factor of 2 from 1 km to 5 km. Consequently the *benefit* also decreases. The peak in the blue profile at ~ 3 km is due to a combination of ~~value of~~ RMSE and information content in the MOPITT retrievals.

The profiles averaged over Africa (Figure 11b) have similar shapes to those in Figure 11a. This is because both the RMSE and *benefit* in Africa are high compared to other parts of the globe (see Figures 6c and 9). Hence the global average is dominated by values over Africa. ~~The planes of HYPNET observations are located at roughly 1.3, 5.3 and 9.3 km.~~ The average height of the GAW observations is 2.2 km. The *benefit* in the *EXP_GAW* experiment has a peak value of ~ 4 ppb at 3.5 km. This is because the GAW station located at Mount Kenya (0.06° S,37.29° E) has an altitude of 3678 meters. Additionally the site at Assekrem, Algeria located at 23.26° N,5.63° E is situated at an altitude of 2715 meters. The EnKF spreads the information content from these observations to the surrounding regions.

The profiles for North America are shown in Figure 11c. ~~The planes of HYPNET observations are located at approximately 1.3, 5.3 and 9.3 km.~~ The average height of the ECCC observations is 0.38 km. The *benefit* due to the ECCC observations close to the surface is ~ 4.1 ppb. The *benefit* in both HYPNET and MOPITT experiments is ~ 2 ppb. The *benefit* due to assimilation of ECCC observations decreases monotonically with height because the RMSE decreases monotonically and also because the ECCC observations cannot constrain the CO state beyond the vertical localization radius. The average height of stations in eastern Canada is ~~212 meters~~. These stations make a major contribution to the *benefit* over North America. Temporally HYPNET observes every 6 hours while ECCC stations observe every hour. Both, the higher temporal frequency and the lower altitude contribute to the higher *benefit* close to the surface in case of the ECCC observations compared to HYPNET ~~which is located at 1 km~~.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

Figure 11d shows the profiles for Europe. ~~The planes of HYPNET observations are located approximately at 1.19, 5.19 and 9.19 km.~~ The average height of the GAW observations is 1.13 km. ~~In the case of GAW network, the *benefit* is 5.9 ppb close to the surface, The HYPNET *benefit* is much smaller (1.7 ppb).~~ As in the case of North America the better performance of GAW stations in the bottom 500 meters is a due to both the lower altitude of the stations and higher temporal frequency of observations compared to the HYPNET.

~~Figure 11e shows the profiles for South America. The planes of HYPNET observations are located at 1.18, 5.18 and 9.18 km. GAW stations are not present in South America and hence the *benefit* is practically zero. The HYPNET and MOPITT *benefits* are approximately proportional to the control RMSE which peaks at 3 km. The South American domain also contains a large part of the Pacific ocean and some part of Atlantic ocean. The average *benefit* profiles for the *EXP_HYP* and *EXP_MOP* represent the improvement in the oceanic CO state due to the assimilation of observations over the oceans and also the *benefit* due to downwind transport from land regions.~~

~~Figures 11f and 11g show an improvement of about 6 ppb close to the surface for *EXP_HYP* and *EXP_MOP* in South and east Asia. The planes of HYPNET observations, in east Asia are located approximately at 1.34, 5.34 and 9.34 km while those in south Asia are located at 1.45, 5.45 and 9.45 km. The average height of the GAW observations is approximately 0.55 km.~~

It is evident that the spatio-temporal structure of *benefit* is similar between HYPNET and MOPITT, both horizontally and vertically. This suggests that, in the idealistic setting of unbiased observations and precisely known observation and model error covariances, the performance of MOPITT retrievals is similar to insitu observations. It is should be noted that MOPITT observations are spatially more dense than our HYPNET observation network. ~~In addition,~~ HYPNET has information at three vertical levels while MOPITT has an information content with one to two degrees of freedom (Deeter et al., 2012) so that ~~vertical information is similar in the two networks~~.

5 Conclusions and further work

A new ~~greenhouse gas~~ data assimilation system based on an operational weather forecast model (EC-CAS v1.0) was developed and validated for the estimation of the 3-dimensional state of CO using simulated observations from HYPNET, ECCC, GAW and MOPITT networks. The spread in CO is obtained by perturbing the winds, flux fields and physics parametrizations. The CO spread approximately matches the RMSE suggesting that an ensemble size of 64 is acceptable for CO estimation. However, these conclusions are based on the assimilation of simulated observations which are unbiased.

These experiments lead to a qualitative understanding of the decrease in RMSE due to the assimilation of CO observations from realistic networks. With all networks it is seen that the benefit due to assimilation of ~~observations~~ is proportional to the CO RMSE. Another factor which controls the pattern of benefit is the locations of observations. For example, the GAW network has only one station in central Africa. The observations from this station are able to effectively constrain the CO state within

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

2000 km. The benefit is the highest in the plane at which this observation is located. The CO state close to the surface is better constrained by observations in the lowermost 500 m than ~~the~~ observations at 1 km. This is suggested by the results in North America and Europe. The CO state over the ocean is constrained partly due to the improvement of the CO state over flux-rich land regions. In the case of MOPITT assimilation, the benefit in central Africa (which is the region with strongest flux) ranges from 10 to over 40 ppb. The downwind transport results in a benefit of 5 to 40 ppb over the tropical ~~Pacific~~. The benefits over south and east Asia range from 2 to 20 ppb. These quantitative findings are expected to change when real observations are assimilated. Biases in observations and correlations in the observational errors along with unaccounted model errors make assimilation of real observations more challenging so that the error reductions are expected to be smaller.

~~In the current state of our EnKF, CO prediction (transport) errors due to uncertain meteorological analyses and model formulation errors (e.g. boundary layer transport, convective transport) are encompassed in our forecast ensemble. However, errors in the simplified chemistry model are not addressed. Since we use simulated observations generated by the same model which later performs the CO data assimilation, this type of error is not simulated. However, CO flux estimates are impacted by the choice of OH field (Yin et al., 2015; Jiang et al., 2017, 2015a, b, 2011). On the other hand, Miyazaki et al. (2012) found that in their forecast ensemble, CO flux errors were correlated only with CO fields (in the lower troposphere) and not with other species. In addition, Miyazaki et al. (2015) note that in their multi-species assimilation system, the assimilation of other species had little impact on CO through OH adjustments. This suggests that for flux estimation, simplified CO reactions with OH climatologies may be sufficient, particularly, if the uncertainty in the chemistry can be accounted for. Thus we plan to allow for the uncertainty in the chemistry module through either using an ensemble of OH climatologies or by directly perturbing the reaction rates for CO loss and for the production of CO by methane oxidation.~~

This work has presented only the very first step in the development of EC-CAS. There are many further stages of development because the goal of EC-CAS is to estimate the 3-dimensional fields of CO, CO_2 and CH_4 and their fluxes along with meteorological fields by assimilating all available observations of meteorological variables and chemical species using an ensemble smoother. These include both in situ and remotely sensed measurements. The immediate next step is to modify EC-CAS 1.0 to allow the update of the CO flux by CO observations since we demonstrated here that the CO state estimation is working well. ~~Preliminary work suggests that HYPNET observational network is able to estimate the CO flux field after about a week of data assimilation.~~ After the ability to estimate fluxes is demonstrated EC-CAS will be tested for estimation of CO and CO_2 3-dimensional field and their fluxes using real observations. The estimates of flux can be improved by using a smoother rather than a filter since a smoother assimilates future observations too. Ultimately, an ensemble Kalman smoother will be developed.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

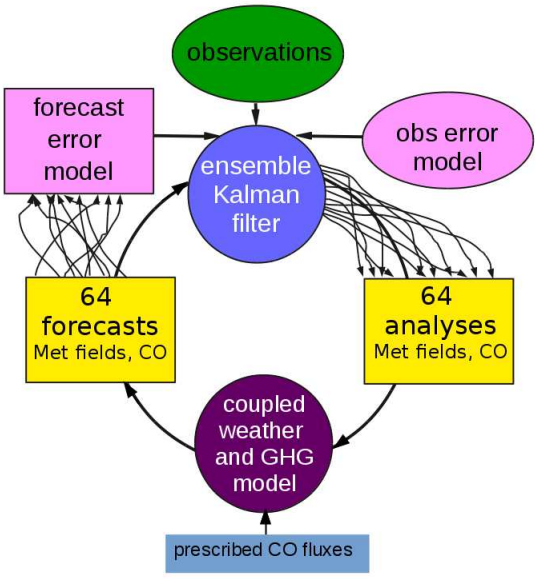

Figure 1. EC-CAS v1.0. There ~~are~~ 64 prescribed CO ~~flux ensemble members~~. In EC-CAS v1.0 flux fields are not estimated. Instead, the perturbed flux field is used to account for uncertainty in CO due to error in the flux. The 64-member forecast ensemble is used along with the observations and the statistics of observation errors as inputs to the EnKF. The 64 analyses of meteorology and CO generated by the EnKF are used as initial conditions for the next 6 hour forecast. This cycle repeats every 6 h.

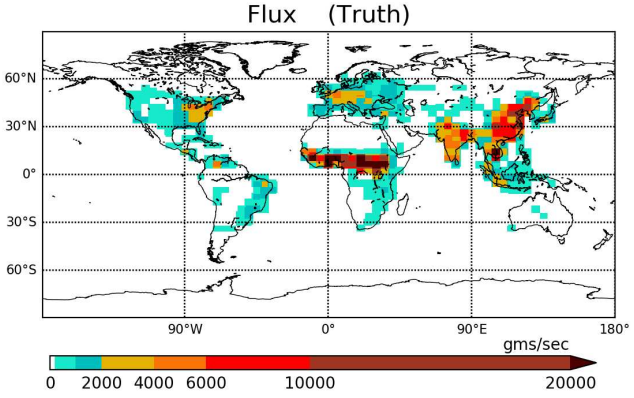

Figure 2. The true CO flux field for January 2015.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

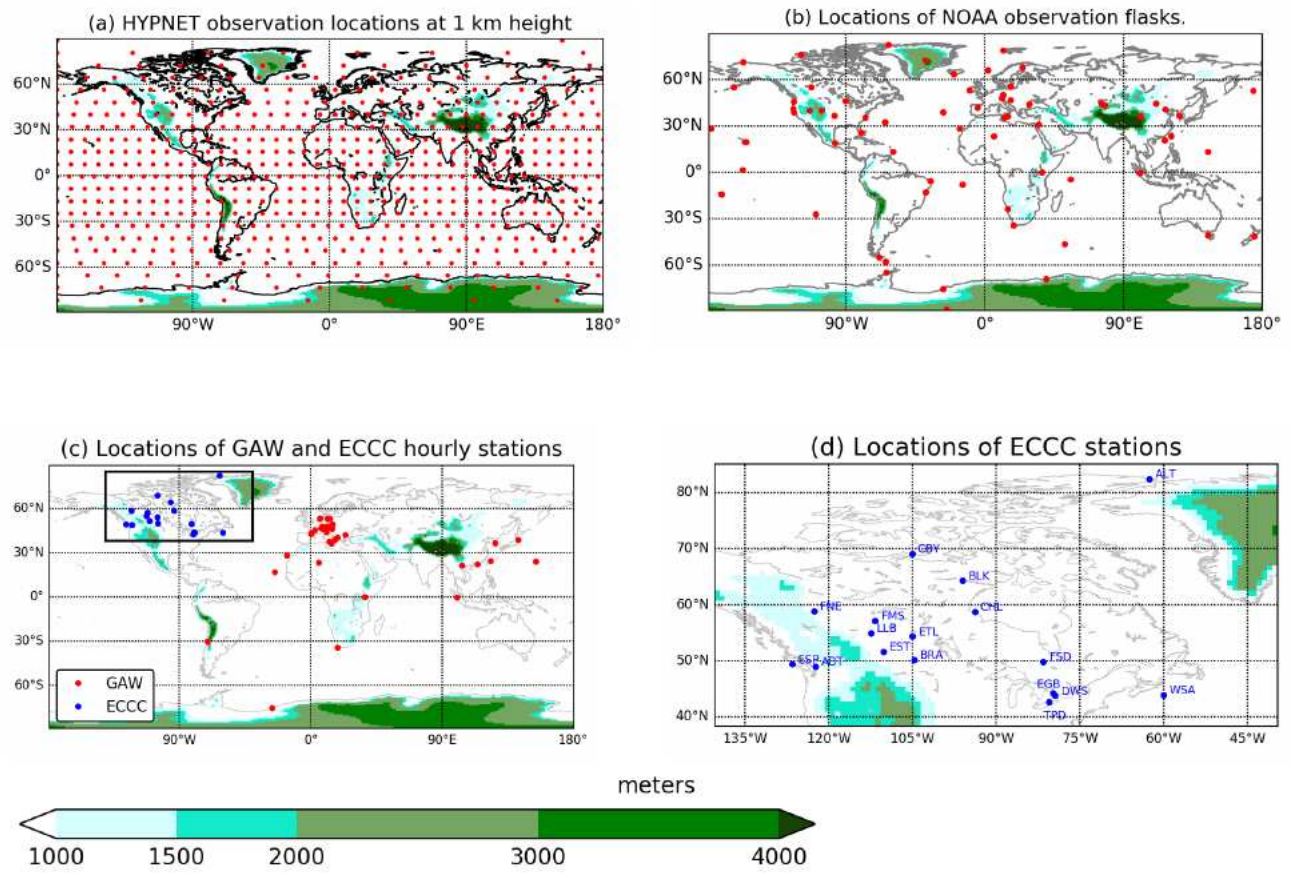

Figure 3. In each panel, topography is shown by the color. (a) HYPNET exists at 1 km, 5 km and 9 km. At each of these levels the observations are located 1000 km apart in the horizontal. There are 622 observations locations at each height and total of $622 \times 3 = 1866$ observations every 6 hours. (b) The locations of flask observations from the NOAA network. (c) The GAW and ECCC station locations. (d) The ECCC station locations also shown in panel (c).

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

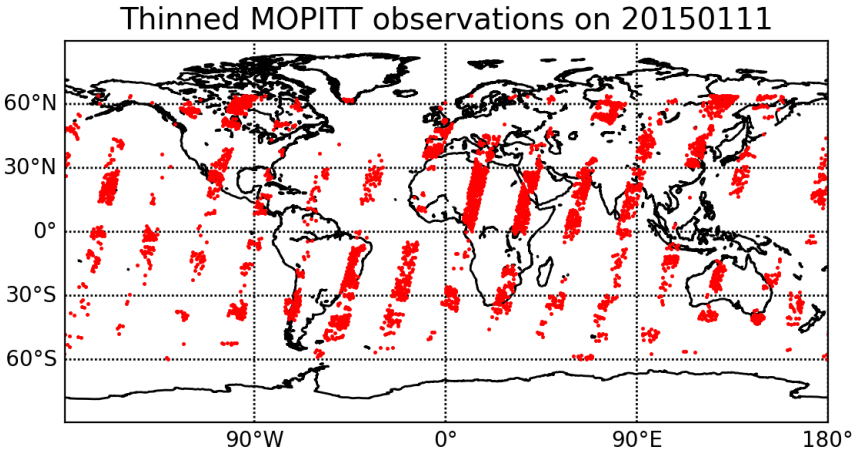

Figure 4. An example of the distribution of MOPITT satellite observations. Thinned MOPITT orbits on 11 January 2015 00:00:00 UTC are shown.

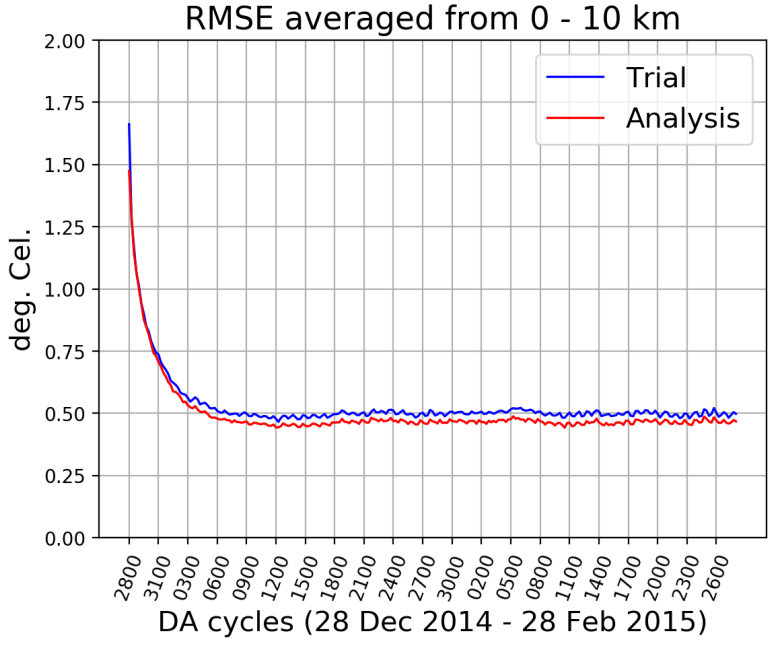

Figure 5. Evolution of the root mean squared error of the temperature analyses (red curve) and ~~trial~~ fields (blue curve) from the control experiment which assimilated meteorological observations only.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

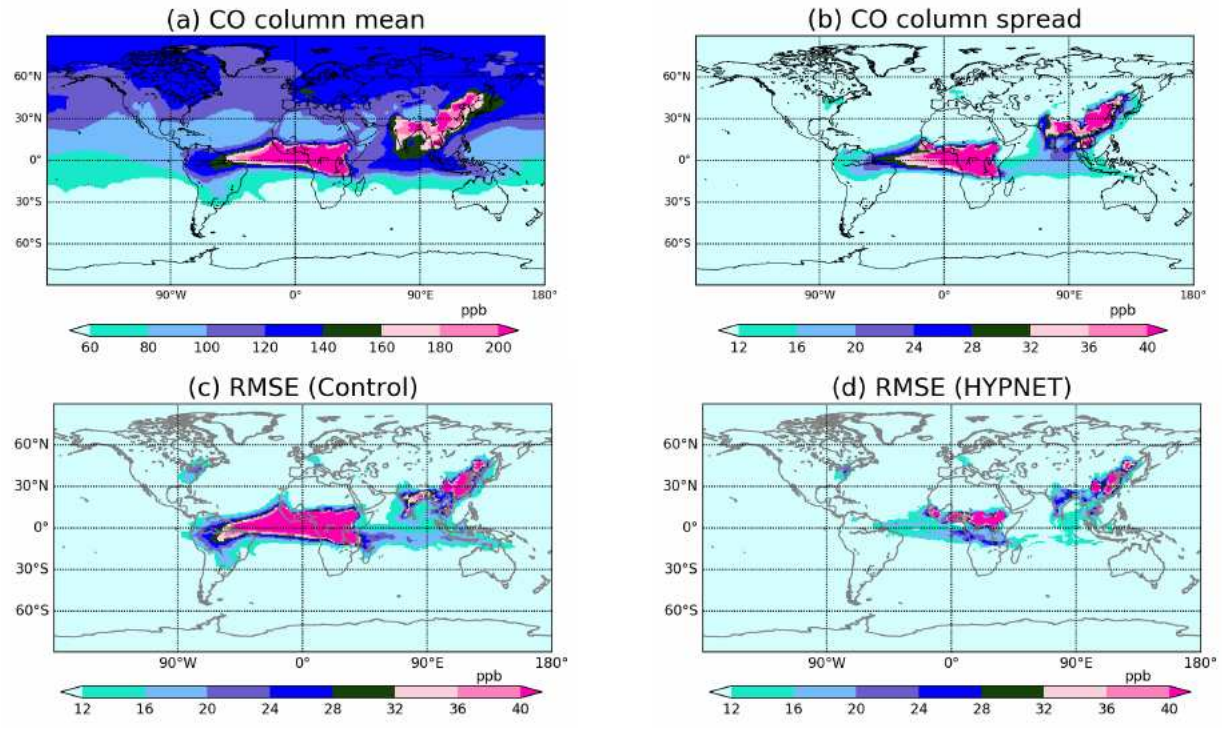

Figure 6. In each panel, column means (0-5 km) averaged from 10 January 2015 18:00:00 UTC to 28 February 2015 00:00:00 UTC are shown. (a) CO ensemble mean of the control experiment. (b) CO ensemble spread of the control experiment. (c) RMSE of the control experiment. (d) RMSE of the HYPNET DA experiment.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

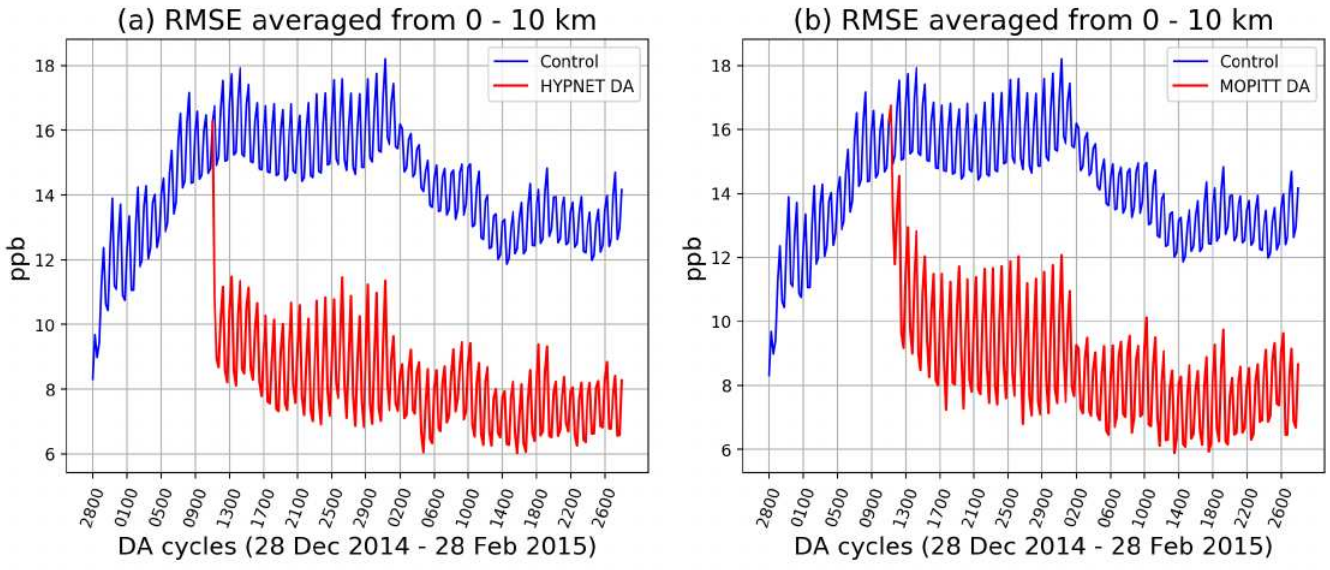

Figure 7. Column (0-10 km) mean RMSE of CO analyses from various experiments : (a) The control, *EXP_CNTRL* (blue curve) and the DA experiment assimilating HYPNET observations, *EXP_HYP* (red curve). (b) The control, *EXP_CNTRL* (blue curve) and the DA experiment assimilating MOPITT observations, *EXP_MOP* (red curve). The blue curves in panel(a) and (b) are identical. The 24 h oscillations in the curves are meteorology induced.

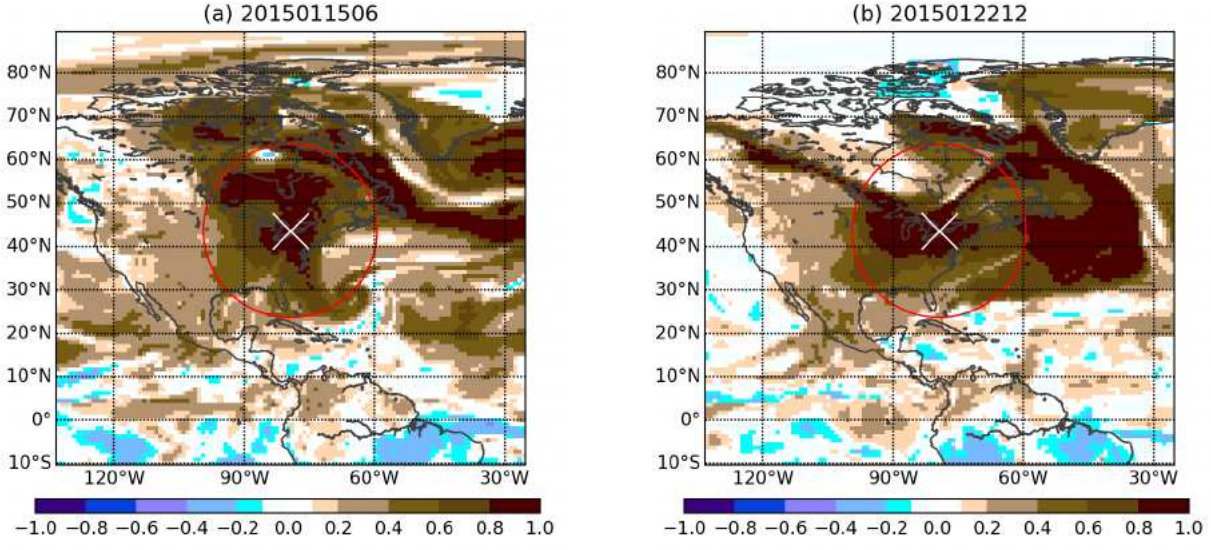

Figure 8. The spatial correlation of CO field between ~~University of~~ Toronto (shown by the white cross) and other locations in the horizontal plane defined by the lowest model level. The red circle of radius 2000 km shows the horizontal localization. The correlation is estimated by using the 64 ~~trial~~ ensemble members.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

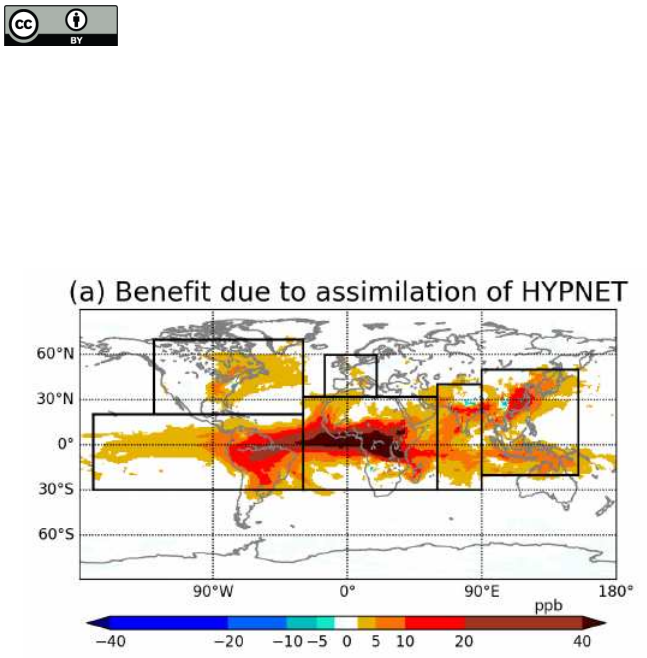

Figure 9. Column (0-5 km) mean benefit averaged from 10 January 2015 18:00:00 UTC to 28 February 2015 00:00:00 UTC. In panel (a), the marked boxes show the domains of North America, South America, Europe, Africa, South Asia and East Asia. These domains are used in Figure 11. Note that all the domains contain both ocean and land. The South American domain includes a large part of the Pacific ocean.

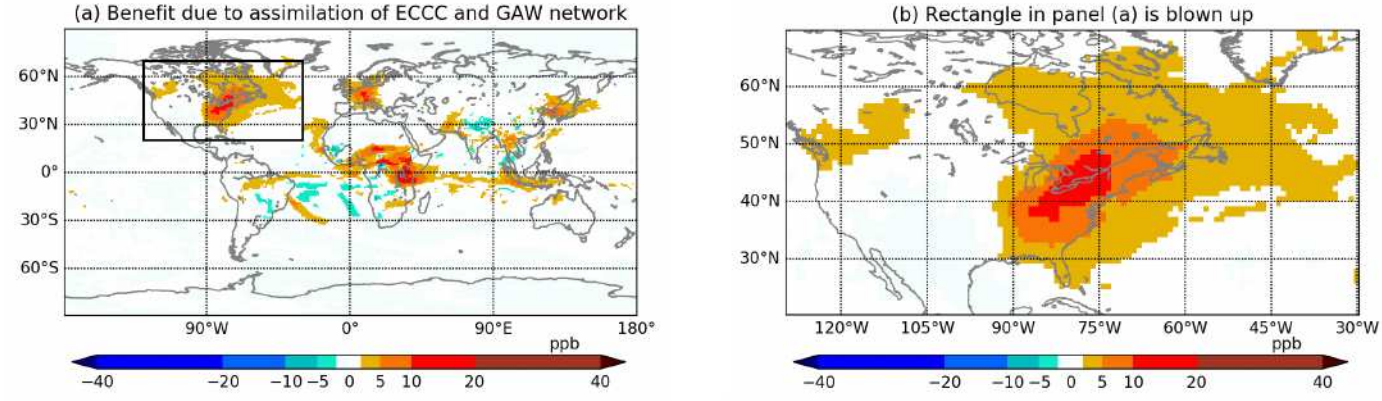

Figure 10. (a) Column (0-5 km) mean benefit averaged from 10 January 2015 18:00:00 UTC to 28 February 2015 00:00:00 UTC for the experiment which assimilates near surface measurements from ECCC and GAW networks (*EXP_GAW*). (b) The North American domain from panel (a).

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

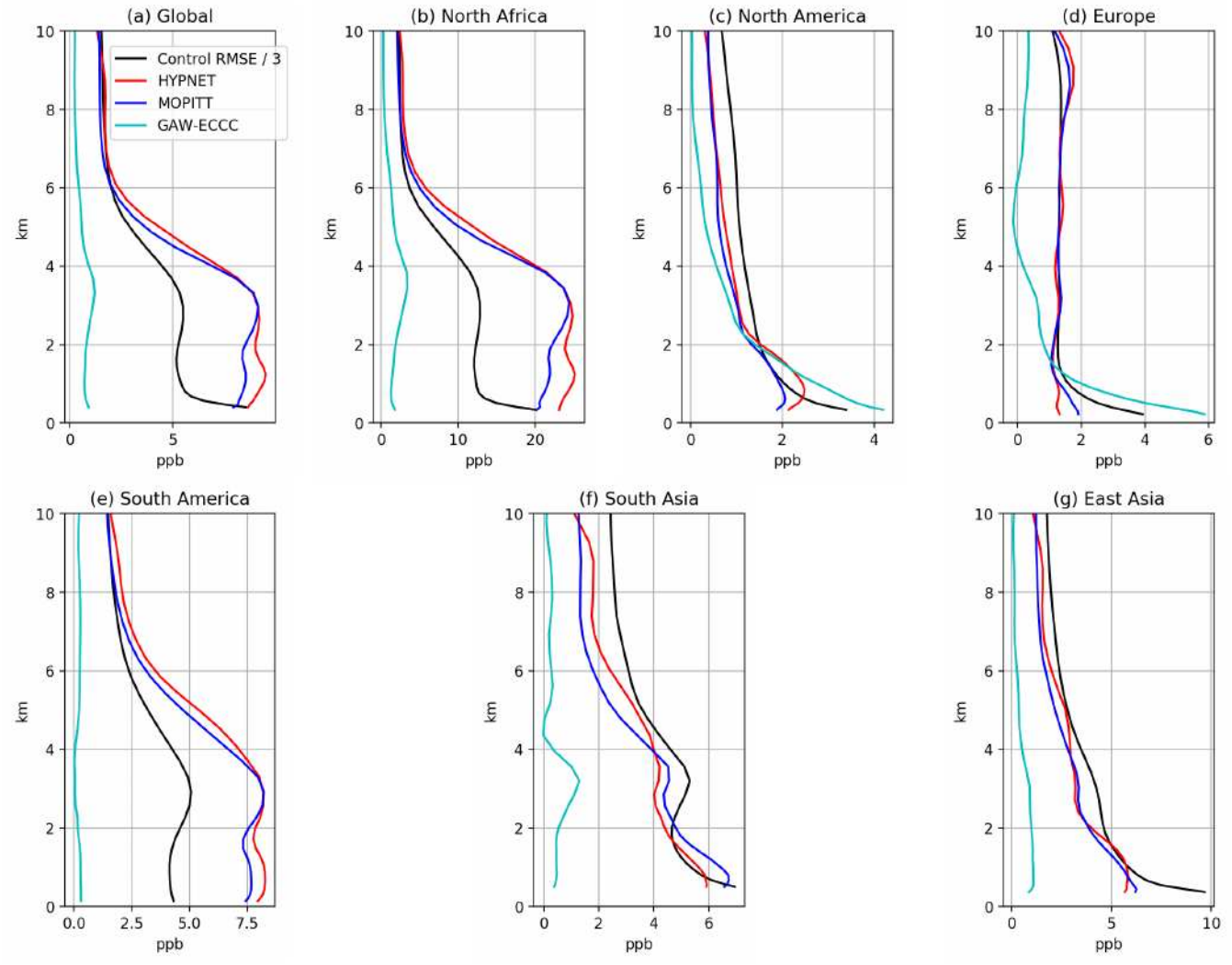

Figure 11. Profiles of benefit and control RMSE averaged over different domains. In each panel the black curve shows the control RMSE scaled down by a factor of 3. The colored curves show the benefits due to the assimilation of three different observational networks (*EXP_HYP*, *EXP_MOP* and *EXP_GAW*). Panel (a) shows the globally averaged profiles. The other panels show profiles averaged over domains shown in figure 9a. The limits on the x-axis are different in each panel.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

Appendix

	Code	Station name	Latitude	Longitude	Altitude
1	ABT	Abbotsford	49.01	−122.34	60.3
2	ALT	Alert	82.45	−62.52	200.0
3	BLK	Baker Lake	64.33	−96.01	94.8
4	BRA	Bratts Lake	50.20	−104.71	595.0
5	CBY	Cambridge Bay	69.13	−105.06	35.0
6	CHL	Churchill	58.74	−93.82	29.0
7	DWS	Downsview	43.78	−79.47	198.0
8	EGB	Egbert	44.23	−79.78	251.0
9	ESP	Estevan Point	49.38	−126.54	7.0
10	EST	Esther Alberta	51.67	−110.21	707.0
11	ETL	East Trout Lake	54.35	−104.99	493.0
12	FMS	Fort McKay	57.15	−111.64	250.0
13	FNE	Fort Nelson	58.84	−122.57	361.0
14	FSD	Fraserdale	49.88	−81.57	210.0
15	LLB	Lac LaBiche	54.95	−112.47	540.0
16	TPD	Turkey Point	42.64	−80.55	231.0
17	WSA	Sable Island	43.93	−60.01	5.0

Table A1. Information about ECCC surface sites used in this study. The altitude is in meters above sea level.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions
Open Access
EGU

	Code	Station name	Latitude	Longitude	Altitude		Code	Station name	Latitude	Longitude	Altitude
1	AMY	Anmyeon-do	36.5383	126.3300	71.0	23	IZO	Izaña	28.3090	−16.4993	2403.0
2	ASK	Assekrem	23.2667	5.6333	2715.0	24	JFJ	Jungfraujoch	46.5475	7.9851	3580.0
3	BEO	Moussala	42.1792	23.5856	2931.0	25	KMW	Kollumerwaard	53.3333	6.2667	3.5
4	BKT	Bukit Kototabang	−0.2019	100.3180	874.0	26	KOS	Kosetice	49.5833	15.0833	535.0
5	CGR	Capo Granitola	37.6667	12.6500	9.0	27	KTB	Kloosterburen	53.4000	6.4200	0.0
6	CMN	Monte Cimone	44.1667	10.6833	2172.0	28	KVV	Krvavec	46.2973	14.5333	1750.0
7	CPT	Cape Point	−34.3534	18.4897	260.0	29	LMT	Lamezia Terme	38.8763	16.2322	14.0
8	CUR	Monte Curcio	39.3160	16.4232	1800.9	30	MKN	Mt. Kenya	−0.0622	37.2972	3682.5
9	CVO	Cape Verde	16.8640	−24.8675	20.0	31	MNM	Minamitorishima	24.2883	153.9833	27.1
10	ECO	Lecce	40.3358	18.1245	86.0	32	NGL	Neuglobsow	53.1428	13.0333	62.0
11	GAT	Gartow	53.0657	11.4429	99.0	33	PAY	Payerne	46.8129	6.9435	494.5
12	GAT	Gartow	53.0657	11.4429	129.0	34	PDI	Pha Din	21.5731	103.5157	1478.0
13	GAT	Gartow	53.0657	11.4429	201.0	35	PDM	Pic du Midi	42.9372	0.1411	2881.0
14	GAT	Gartow	53.0657	11.4429	285.0	36	PUY	Puy de Dome	45.7723	2.9658	1467.0
15	GAT	Gartow	53.0657	11.4429	410.0	37	RIG	Rigi	47.0674	8.4633	1036.0
16	GLH	Giordan	36.0700	14.2200	167.0	38	RYO	Ryori	39.0319	141.8222	280.0
17	HBA	Halley	−75.3500	−26.3900	38.0	39	SNB	Sonnblick	12.9578	47.0542	3111.0
18	HKG	Hok Tsui	22.2095	114.2578	60.0	40	SSL	Schauinsland	47.9000	7.9167	1205.0
19	HPB	Hohenpeissenberg	47.8000	11.0200	1003.0	41	TLL	El Tololo	−30.1683	−70.8036	2159.0
20	HPB	Hohenpeissenberg	47.8011	11.0246	1035.0	42	YON	Yonagunijima	24.4667	123.0106	50.0
21	HPB	Hohenpeissenberg	47.8011	11.0246	1078.0	43	ZSF	Zugspitz	47.4165	10.9796	2670.0
22	HPB	Hohenpeissenberg	47.8011	11.0246	1116.0	44	ZUG	Zugspitz-Gipfel	47.4211	10.9859	2965.5

Table A2. Information about GAW surface sites used in this study. The altitude is in meters above sea level.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020


Geoscientific Model Development Discussions

Code and data availability. The source code is publicly available at https://doi.org/10.5281/zenodo.3908545 under the GNU Lesser General Public License version 2.1 (LGPL v2.1) or ECCC's Atmospheric Sciences and Technology licence version 3. The model data output are available at http://crd-data-donnees-rdc.ec.gc.ca/CCMR/pub/2020_Khade_ECCAS_all_data/.

Author contributions. Vikram Khade developed the code related to the extension of EnKF to GHGs and their fluxes. Vikram Khade and Saroja Polavarapu designed, carried out the experiments and interpreted them. Michael Neish prepared the input data required for the experiments. This includes but is not limited to the regridding of flux fields used as the truth. Pieter Houtekamer supervised the development of the code. Seung-Jong Baek helped with code debugging and optimization of the performance of the code. Dylan Jones played an important role in the designing of the experiments and interpretation of the results. Tailong He carried out the 4D-Var inversion using GEOS-Chem to produce flux fields used as truth. Sylvie Gravel helped with the development needed to include CO as a species in the forecast model, GEM-MACH-GHG.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We are grateful to the following people for their advice or assistance. Douglas Chan helped with the understanding of the surface observations. Doug Worthy and his team provided the ECCC surface observations. Feng Deng helped with interpretation of results. Jinwong Kim ran few experiments. Michael Sitwell provided the code to convert CO observations into BURP format. Yves Rochon provided the MOPITT data and also the code for MOPITT forward operator. Jim Drummond helped with the understanding of the MOPITT data. We thank Michael Sitwell and Yves Rochon for helpful comments on a previous version of the manuscript.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Geoscientific Model Development Discussions

References

Abatzoglou, J.T., and Williams, A.P. : Impact of anthropogenic climate change on wildfires across western US forests. Proceedings of the National Academy of Sciences of the United States of America,113(42), 11770-11775,2016.

Anderson,J.L.: An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev., 129, 2884-2903, 2001.

Anselmo, D., M.D. Moran, S. Menard, V. Bouchet, P. Makar, W. Gong, A. Kallaur, P.-A. Beaulieu, H. Landry, C. Stroud, P. Huang, S. Gong, and D. Talbot, 2010.A new Canadian air quality forecast model: GEM-MACH15. Proc. 12th AMS Conf. on Atmos. Chem., Jan. 17-21, Atlanta, GA, American Meteorological Society, Boston, MA, 6 pp. [see http://ams.confex.com/ams/pdfpapers/165388.pdf]

Arellano, A. F. and Hess, P. G. : Sensitivity of top-down estimates of CO sources to GCTM transport, Geophys. Res. Lett., 33, L21807, http://doi:10.1029/2006GL027371, 2006.

Baker, D.F., Doney, S.C. and Schimel, D.S.: Variational data assimilation for atmospheric CO2, Tellus, 58B, 359-365, 2006.

Bannister, R.N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I : characteristics and measurements of forecast error covariances. QJRMS, 134(637),1951-1970,2008.

Barré, J., Gaubert B., Arrellano A.J., Worden H.M., Edwards, D.P., Deeter M.N., Anderson J.L., Raeder K., Collins N., Tilmes S., Clerbaux C., Emmons L.K., Pfister G., Coheur P-F., Hurtmans D. : Assessing the impacts of assimilating IASI and MOPITT CO retrievals using CESM-CAM-chem and DART,J.Geophy. Res. Atmos., 120. doi:10.1002/2015JD023467.

Bloom, S., Takacs, L., DaSilva, A., and Ledvina, D.: Data assimilation using incremental analysis updates, Mon. Wea. Rev., 124: 1256-1271, 1996.

~~Bruhwiler, L. M. P., Michalak, A. M., and Tans, P. P.: Spatial and temporal resolution of carbon flux estimates for 1983-2002, Biogeosciences, 8, 1309-1331, doi:10.5194/bg-8-1309-2011, 2011~~.

Bruhwiler, L., Dlugokencky, E., Masarie, K., Ishizawa, M., Andrews, A., Miller, J., Sweeney, C., Tans, P., and Worthy, D.: CarbonTracker-CH4: an assimilation system for estimating emissions of atmospheric methane, Atmos. Chem. Phys., 14, 8269-8293, https://doi.org/10.5194/acp-14-8269-2014, 2014.

~~Chan, D., Yuen, C. W., Higuchi, K., Shashkov, A., Liu, J., Chen, J., Worthy, D.: On the CO2 exchange between the atmosphere and the biosphere: the role of synoptic and mesoscale processe s. Tellus B, 56. doi:http://dx.doi.org/10.3402/tellusb.v56i3.16424, 2004.~~

~~Chevallier, F., Feng, L., Bosch, H., Palmer, P.I. and Rayner: On the impact of transport model errors for the estimation of CO2 surface fluxes from GOSAT observations, Geophys. Res. Lett., 37, L21803, doi:10.1029/2010GL044652, 2010~~.

~~Chevallier, F., Palmer, P.I., Feng, L., Boesch, H., O'Dell, C.W. and Bousquet, P.: Toward robust and consistent regional CO2 flux estimates from in situ and spaceborne measurements of atmospheric CO2, Geophys. Res. Lett., 41, 1065-1070, doi:10.1002/2013GL058772, 2014.~~

Cohn, S.E., and Parrish,D.F.: The behavior of forecast error covariances for a Kalman filter in two dimensions. Mon. Wea. Rev., 119, 1757-1785, 1991.

Côté, J., Gravel, S., Methot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation, Mon. Wea. Rev., 126, 1373-1395, doi:10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2., 1998a.

Côté, J., Desmarais, J.-G., Gravel, S., Methot, A., Patoine, A., Roch, M. and Staniforth, A.: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part II: Results, Mon. Wea. Rev., 126, 1397-1418, doi:10.1175/1520-0493(1998)126<1397:TOCMGE>2.0.CO;2,1998b.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015-2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmos. Chem. Phys., 19, 9797-9831, https://doi.org/10.5194/acp-19-9797-2019, 2019.

Deeter, M., Worden, H., Edwards, D., Gille, J., and Andrews, A.: Evaluation of MOPITT retrievals of lower-tropospheric carbon monoxide over the United States, J. Geophys. Res., 117, D13306, doi:10.1029/2012JD017553,2012.

Deeter, M., Edwards, D., Francis, G., Gille, J., Martinez-Alonso, S., Worden, H., and Sweeeney, C.: A climate-scale satellite record for Carbon monoxide : the MOPITT version 7 product, Atmos. Meas. Tech., 10, 2533-2555, https://doi.org/10.5194/amt-10-2533-2017, 2017.

Drummond, J.R.: Measurement of Pollution in the Troposhere (MOPITT). The use of EOS for studies of Atmospheric Physics,77-101,1992.

Flannigan,M.D., Krawchuk, M.A., de Groot,W.J., and Growman,L.M : Implications of changing climate for global wildland fire, International Journal of Wildfire,18,483-507,2009

Fortems-Cheiney, A., F. Chevallier, I. Pison, P. Bousquet, S. Szopa, M. N. Deeter, and C. Clerbaux : Ten years of CO emissions as seen from Measurements of Pollution in the Troposphere (MOPITT), J. Geophys. Res., 116, D05304, doi:10.1029/2010JD014416, 2011.

Gaspari, G., and Cohn, S.E.: Construction of correlation functions in two and three dimensions. Quart. J. Roy. Meteor. Soc.,125,723-757, 1999.

Gaubert, B., Arellano Jr., A., Barre, J., Worden, H., Emmons, L., Tilmes, S., Buchholz, R., Vitt, F., Raeder, K., Collins, N., Anderson, J., Wiedinmyer, C., Martinez Alonso, S., Edwards, D., Andreae, M., Hannigan, J., Petri, C., Strong, K., and Jones, N.: Toward a chemical reanalysis in a coupled chemistry-climate model : An evaluation of MOPITT CO assimilation and its impact on tropospheric compostion. Quart. J. Geophys. Res. Atmos.,121,7310-7343,2016.

Ghil, M., Cohn, S., Tavantzis, J., Bube, K., and Isaacson, E. : Applications of estimation theory to numerical weather prediction. Dynamic meteorology - data assimilation methods. Eds. Bengtsson,L., Ghil,M., and Kallen, E.,Springer-verlag, Heidelberg, 134-224, 1981.

Girard, C., Plante, A., Desgagne, M., McTaggart-Cowan, R., Cote, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P. A., Zadra, A.: Staggered Vertical Discretization of the Canadian Environmental Multiscale (GEM) Model Using a Coordinate of the Log-Hydrostatic-Pressure Type. Mon. Wea. Rev., 142, 113, DOI: 10.1175/MWR-D-13-00255.1, 2014.

Gong, W., P.A. Makar, J. Zhang, J. Milbrandt, S. Gravel, K.L. Hayden, A.M. Macdonald, and W.R. Leaitch, 2015. Modelling aerosol-cloud-meteorology interaction: A case study with a fully coupled air quality model (GEM-MACH).Atmos. Environ., http://dx.doi.org/10.1016/j.atmosenv.2015.05.062, 115, 695-715,2015.

Hamill, T.M., Whitaker, J.S. and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon. Wea. Rev., 129, 2776-2790, 2001.

Houweling, S., Aben, I., Breon, F.-M., Chevallier, F., Deutscher, N., Engelen, R., Gerbig, C., Griffith, D., Hungershoefer, K., Macatangay, R., Marshall, J., Notholt, J., Peters, W., and Serrar, S.: The importance of transport model uncertainties for the estimation of CO2 sources and sinks using satellite measurements, Atmos. Chem. Phys., 10, 9981-9992, doi:10.5194/acp-10-9981-2010, 2010.

Houtekamer, P.L., and Zhang, F.: Review of the Ensemble Kalman Filter for atmospheric data assimilation. Mon. Wea. Rev., 144(12), 4489-4532, 2016.

Houtekamer, P.L., and Mitchell, H.L.: Data assimilation using an Ensemble Kalman Filter technique. Mon. Wea. Rev., 126(3), 796-811, 1998.

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmos. Chem. Phys., 13, 4073-4109, https://doi.org/10.5194/acp-13-4073-2013, 2013.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, Atmos. Chem. Phys., 15, 5275-5303, https://doi.org/10.5194/acp-15-5275-2015 , 2015.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515-3556, https://doi.org/10.5194/acp-19-3515-2019 , 2019.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmos. Chem. Phys., 15, 11411-11432, https://doi.org/10.5194/acp-15-11411-2015, 2015.

Mitchell, H.L. and Houtekamer, P.L.: An adaptive ensemble Kalman filter. Mon. Wea. Rev., 128, 416-433, 2000.

Houtekamer, P.L., and Mitchell, H.L.: A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev., 129, 123-137, 2001.

Houtekamer, P.L., Deng, X., Mitchell, H.L., Baek, S.-J. and Gagnon, N.: Higher Resolution in an Operational Ensemble Kalman Filter, Mon. Wea. Rev., 142, 1143-1162, 2014.

Houtekamer, P.L., Buehner, M., and Chevrotiere, M.: Using the hybrid gain algorithm to sample data assimilation uncertainty, QJRMS., 145, 35-56, 2019.

Jiang, Z., Jones, D. B. A., Kopacz, M., Liu, J., Henze, D. K., and C. Heald : Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations, J. Geophys. Res., 116, D15306, doi:10.1029/2010JD015282, 2011.

Jiang, Z., Jones, D. B. A., Worden,H. M.,Deeter, M. N., Henze,D. K., Worden, J., Bowman, K. W., Brenninkmeijer, C. A. M., and Schuck, T. J. : Impact of model errors in convective transport on CO source estimates inferred from MOPITT CO retrievals, J. Geophys. Res. Atmos., 118, 2073-2083, https://doi:10.1002/jgrd.50216, 2013.

Jiang, Z., Jones, D. B. A., Worden, H. M., and Henze, D. K.: Sensitivity of top-down CO source estimates to the modeled vertical structure in atmospheric CO, Atmos. Chem. Phys., 15, 1521-1537, https://doi.org/10.5194/acp-15-1521-2015 , 2015a.

Jiang, Z., Jones, D. B. A., Worden, J., Worden, H. M., Henze, D. K., and Wang, Y. X.: Regional data assimilation of multi-spectral MOPITT observations of CO over North America, Atmos. Chem. Phys., 15, 6801-6814, https://doi.org/10.5194/acp-15-6801-2015 , 2015b.

Jiang, Z., Worden, J. R., Worden, H., Deeter, M., Jones, D. B. A., Arellano, A. F., and Henze, D. K.: A 15-year record of CO emissions constrained by MOPITT CO observations, Atmos. Chem. Phys., 17, 4565-4583, https://doi.org/10.5194/acp-17-4565-2017, 2017.

JPL: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies. Evaluation Number 17. NASA Panel for Data Evaluation. June 10, 2011. Available from: https://jpldataeval.jpl.nasa.gov/previous_evaluations.html

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Kain, J.S., and Fritsch, J.M.: A one-dimensional entraining/detraining plume model and its application in convective parameterizations, J. Atmos. Sci. 47, 2784-2802, 1990.

Kain, J.S.: The Kain-Fritsch convective parameterization: an update, J. Appl. Meteorol. 43, 170-181, 2004.

Kang, J.-S., Kalnay, E., Liu, J., Fung, I., Miyoshi, T. and Ide ,K.: "Variable localization" in an ensemble Kalman filter: Application to the carbon cycle data assimilation, J. Geophys. Res., 116, D09110, doi:10.1029/2010JD014673, 2011.

Kang, J. S., Kalnay, E., Miyoshi, T., Liu, J. and Fung, I.: Estimation of surface carbon fluxes with an advanced data assimilation methodology, J. Geophys. Res., 117, D24101, doi:10.1029/2012JD018259, 2012.

Kim, J., Polavarapu, S., Chan, D., and Neish, M.: The Canadian atmospheric transport model for simulating greenhouse gas evolution on regional scales: GEM-MACH-GHG v.137-reg, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-123, 2020.

Khade, V., Neish, M., Houtekamer, P.L., Polavarapu, S., Baek, S-J., and Jones. D.B.A. (2020, June 29). EC-CAS v1.0. Zenodo. https://doi.org/10.5281/zenodo.3908545

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), Atmos. Chem. Phys., 10, 855-876, https://doi.org/10.5194/acp-10-855-2010, 2010.

Law, R. M., Rayner, P. J., Denning, A. S., Erickson, D., Fung, I. Y., Heimann, M., Piper, S. C., Romonet, M., Taguchi, S., Taylor, J. A., Trudinger, C. M., and Watterson, I. G.: Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO2 inversions, Global Biogeochem. Cy., 10(4), 783-796, 1996.

Liu,J., Fung, I., Kalnay, E. and Kang. J : CO2 transport uncertainties from the uncertainties in meteorological fields, Geophysical Research Letters,38(12), https://doi.org/10.1029/2011GL047213, 2011.

Liu, Y., Kalnay, E., Zeng, N., Asrar, G., Chen, Z., and Jia, B.: Estimating surface carbon fluxes based on a local ensemble transform Kalman filter with a short assimilation window and a long observation window: an observing system simulation experiment test in GEOS-Chem 10.1, Geosci. Model Dev., 12, 2899-2914, https://doi.org/10.5194/gmd-12-2899-2019, 2019.

Miyazaki,K., Maki T., Patra P. and Nakazawa, T. : Assessing the impact of satellite, aircraft and surface observations on CO2 flux estimation using an ensemble based 4-D data assimilation system, J. Geophys. Res., 116, D16306,doi:10.1029/2010JD015366, 2011.

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.: Simultaneous assimilation of satellite NO2, O3, CO, and HNO3 data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545-9579, https://doi.org/10.5194/acp-12-9545-2012 , 2012.

Miyazaki, K., Eskes, H. J., and Sudo, K.: A tropospheric chemistry reanalysis for the years 2005-2012 based on an assimilation of OMI, MLS, TES, and MOPITT satellite data, Atmos. Chem. Phys., 15, 8315-8348, https://doi.org/10.5194/acp-15-8315-2015, 2015.

McTaggart-Cowan, R., and Zadra, A.: Representing Richardson Number Hysteresis in the NWP Boundary Layer, Mon. Wea. Rev., 143, 1232-1258, 2015.

Neish, M., Tanguay, M., Semeniuk, K., Polavarapu, S., DeGrandpre, J., Girard, C., Qaddouri, A., Gravel, S., Chan, D., Ren, S., GEM-MACH development team. (2019, June 19). GEM-MACH-GHG revision 137 (Version 137). Zenodo. http://doi.org/10.5281/zenodo.3246556

Ott, L., Pawson, S., and Bacmeister, J. : An analysis of the impact of convective parameter sensitivity on simulated global atmospheric CO distributions, J. Geophys. Res., 116, D21310, doi:10.1029/2011JD016077,2011

https://doi.org/10.5194/gmd-2020-219
Preprint. Discussion started: 18 August 2020

Parazoo, N. C., Denning, A. S., Kawa, S. R., Corbin, K. D., Lokupitiya, R. S., and Baker, I. T.: Mechanisms for synoptic variations of atmospheric CO2 in North America, South America and Europe, Atmos. Chem. Phys., 8, 7239-7254, https://doi.org/10.5194/acp-8-7239-2008, 2008.

Parrish, D.F. and Derber, J.C.: The National Meteorological Center's spectral statistical interpolation analysis system. Mon. Wea. Rev., 120, 1747-1763,1992.

Pavlovic, R., J. Chen, K. Anderson, M.D. Moran, P.-A. Beaulieu, D. Davignon, and S. Cousineau, 2016. The FireWork air quality forecast system with near-real-time biomass burning emissions: Recent developments and evaluation of performance for the 2015 North American wildfire season. J. Air & Waste Manage. Assoc., 66, 819-841, https://doi.org/10.1080/10962247.2016.1158214, 2016.

Pires, C., R. Vautard and Talagrand, O.: On extending the limits of variational assimilation in nonlinear chaotic systems. Tellus A. 10.3402/tellusa.v48i1.11634,1996.

Polavarapu, S. M., Neish, M., Tanguay, M., Girard, C., de Grandpre, J., Semeniuk, K., Gravel, S., Ren, S., Roche, S., Chan, D., and Strong, K.: Greenhouse gas simulations with a coupled meteorological and transport model: the predictability of CO2, Atmos. Chem. Phys., 16, 12005-12038, https://doi.org/10.5194/acp-16-12005-2016, 2016.

Polavarapu, S. M., Deng, F., Byrne, B., Jones, D. B. A., and Neish, M.: A comparison of posterior atmospheric CO2 adjustments obtained from in situ and GOSAT constrained flux inversions, Atmos. Chem. Phys., 18, 12011-12044, https://doi.org/10.5194/acp-18-12011-2018, 2018.

Prive, N., Errico, R.M., and Carvalho, D. : Performance and evaluation of the Global Modeling and Assimilation Office observing system simulation experiment, 98th AMS annual Meeting, Austin, TX, 2018.

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, Journal of Geophysical Research-Atmospheres, 105, 8931-8980, 2000.

Voigt, C., Marushchak, M.E., Mastepanov, M., Lamprecht, R. E., Christensen, T. R., Dorodnikov, M., Jackowicz-Korczynski, M., Lindgren, A., Lohila, A., Nykanen, H., Oinonen, M., Oksanen, T., Palonen, V., Treat, C. C., Martikainen, P. J., and Biasi, C.: Ecosystem carbon response of an Arctic peatland to simulated permafrost thaw. Glob Change Biol. 2019.

Worthy, D E., Platt, J. A., Kessler, R., Ernst, M., Audette , C., and Racki, S.: An update on the Canadian GHG measurement program, In: Report of the 12th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, Toronto, Canada, September 2003, D. Worthy and L. Huang (editors), World Meteorological Organization Global Atmosphere Watch, Report 162, 220-231, 2005.

Yin, Y., F. Chevallier, P. Ciais, G. Broquet, A. Fortems-Cheiney, I. Pison, and M. Saunois : Decadal trends in global CO emissions as seen by MOPITT, Atmos. Chem. Phys., 15, 13,433-13,451, doi:10.5194/acp-15-13433-2015, 2015.

A WILEY PUBLICATION
IN MATHEMATICAL STATISTICS

Introduction to Mathematical Statistics

FOURTH EDITION

PAUL G. HOEL
Professor of Mathematics
University of California
Los Angeles

JOHN WILEY & SONS, INC.
New York London Sydney Toronto

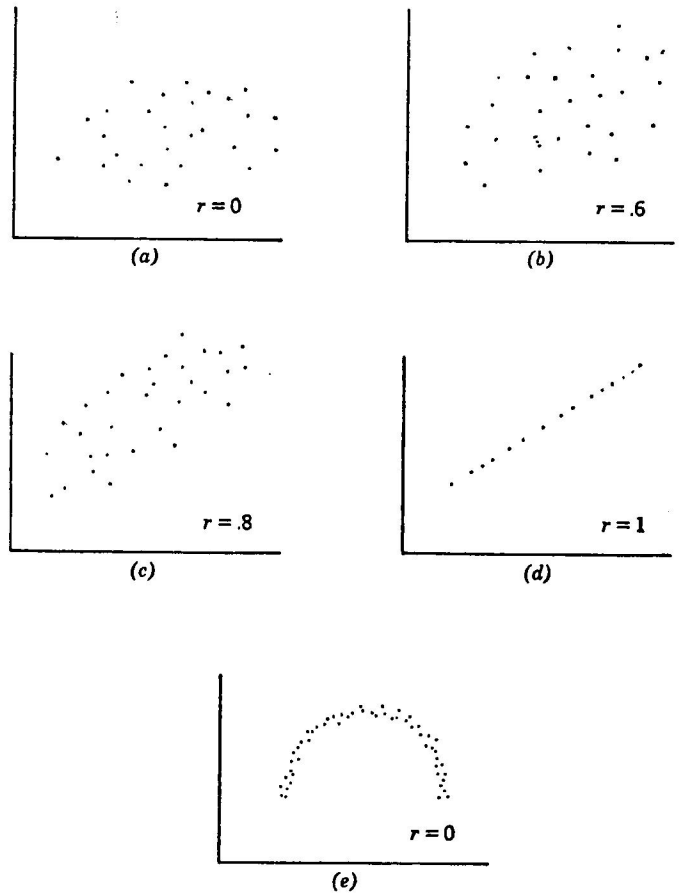

Fig. 2. Scatter diagrams and their associated values of r.

2 THE RELIABILITY OF r

If X and Y are jointly normally distributed the pairs of random sample variables X_i, Y_i, $i = 1, \cdots, n$, will be independently distributed each with the same distribution as X and Y. In terms of these random variables the sample means and variances are denoted by $\bar{X}$, $\bar{Y}$, S_X, and S_Y, and the sample correlation coefficient by

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{nS_XS_Y}.$$

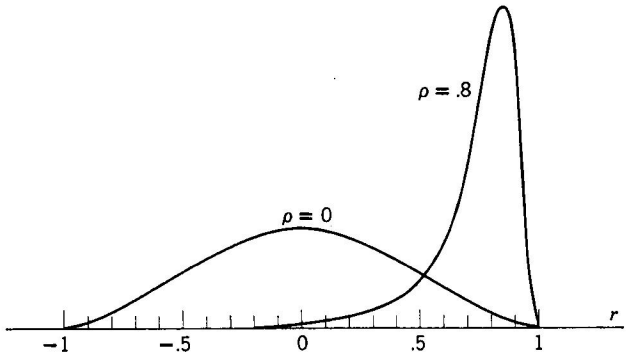

Fig. 3. Distribution for r for $\rho = 0$ and $\rho = .8$ when $n = 9$.

After a sample of size n has been taken and the observational values $(x_1, y_1), \cdots, (x_n, y_n)$ made available, r as given by formula (1) can be computed and is merely a number. However, in discussing how the function r behaves in repeated sampling experiments, r is a random variable which is a function of the n pairs of random variables X_i, Y_i, $i = 1, \cdots, n$. It is theoretically possible to derive the probability density function of r from the density function of those n pairs of variables; however, both the form and the derivation of this density are too complicated to be considered here. It turns out that the density function of r depends only on the parameters ρ and n, where n is the number of points in the scatter diagram. Graphs of the density function of r for $\rho = 0$ and for $\rho = .8$ when $n = 9$ are shown in Fig. 3.

It is clear from Fig. 3 that the distribution of r is decidedly non-normal for large values of ρ; consequently it will not suffice to obtain the standard deviation of r and use it to determine the accuracy of r as an estimate of ρ. Fortunately, there exists a simple change of variable which transforms the complicated distribution of r into an approximately normal distribution. The resulting normal distribution may then be used to determine the accuracy of r as an estimate of ρ in the same way that the normal distribution of $\bar{X}$ was used to determine the accuracy of $\bar{X}$ as an estimate of μ. This change of variable is from r to z, where

$$z = \tfrac{1}{2}\log_e \frac{1 + r}{1 - r}. \tag{3}$$

It can be shown that when the preceding assumptions are satisfied, the random variable z will be approximately normally distributed with mean

$$\mu_z = \tfrac{1}{2}\log_e \frac{1 + \rho}{1 - \rho}$$

and standard deviation

$$\sigma_z = \frac{1}{\sqrt{n-3}}.$$

As an illustration of how this transformation is used, consider the problem of determining an interval of values within which r could reasonably be expected to fall if $\rho = .8$ and if r is based on a sample of size 28. Let reasonably be understood to mean with a probability of .95. The construction of such an interval can be accomplished by first constructing such an interval for z and then transforming it into an interval for r. The simplest interval for z that possesses the desired property is the interval with end points $z_1 = \mu_z - 2\sigma_z$ and $z_2 = \mu_z + 2\sigma_z$. For $\rho = .8$ and $n = 28$, it follows from (3) that those end points are

$$z_1 = \tfrac{1}{2}\log 9 - \frac{2}{\sqrt{25}} = .70$$

and

$$z_2 = \tfrac{1}{2}\log 9 + \frac{2}{\sqrt{25}} = 1.50.$$

From tables of the exponential function it will be found that values of r that correspond to these values of z are $r_1 = .60$ and $r_2 = .91$. Thus it can be stated that the probability is approximately .95 that the sample correlation coefficient will satisfy the inequality $.60 < r < .91$ when r is based on a sample of 28 and $\rho = .80$. This example illustrates how unreliable r is as an estimate of ρ unless one has a very large sample.

Although the preceding relationship simplifies the problem of determining the accuracy of r as an estimate of ρ, it has the disadvantage of being unreliable if X and Y do not have a joint normal distribution; consequently unless one is quite certain that these variables possess such a distribution, at least to a good approximation, the results should not be relied upon.

3 INTERPRETATION OF r

Given any two random variables X and Y one can ask the question whether those variables are independent. Since two variables are independent if, and only if, $f(x, y) = g(x)h(y)$ where $g(x)$ and $h(y)$ are the marginal densities of X and Y and since an extremely large sample would be required to determine whether this relationship is being satisfied, it is clear that some other method is needed to solve this problem. One approach is to introduce some measure of the relationship between two variables, whose value is zero for independent variables, and use it to determine whether the variables are independent.

In view of the definition of $f(x, y)$, it is seen that the normal variables X and Y are independent if, and only if, they are uncorrelated. Thus, the parameter ρ completely determines whether or not two normal variables are independently distributed. As a result, it suffices to determine whether $\rho = 0$ for such a pair of variables to ascertain independence. Since the sample correlation coefficient r serves as an estimate of ρ, it can be used to determine whether it is reasonable to assume that $\rho = 0$,

If X and Y cannot be assumed to be normally distributed, even approximately, then ρ can no longer be used as a basis for determining the extent to which X and Y are related. Figure 2 and Fig. 5, Chapter 6, both indicate the inefficiency of ρ and r for measuring the extent of the relationship when X and Y are not normally distributed.

Even though two variables may possess a joint normal distribution, and therefore that ρ may be used as a measure of the strength of the relationship of the two variables, it does not follow that the relationship as measured by ρ is meaningful in a practical sense. The fact that two variables tend to increase or decrease together does not imply that one has any direct or indirect effect on the other. Both may be influenced by other variables in a manner that will give rise to a strong mathematical relationship. The favorite example to illustrate this fact is the one concerned with teachers' salaries. Over a period of years the correlation coefficient between teachers' salaries and the consumption of liquor turned out to be .98. During that period of time there was a steady rise in wages and salaries of all types and a general upward trend of good times. Under such conditions teachers' salaries would also increase. Moreover, the general upward trend in wages and buying power would be reflected in increased purchases of liquor. Thus, the high correlation merely reflected the common effect of the upward trend on the two variables. This is a type of correlation that has received the name of *spurious correlation*. The preceding discussion should make it clear that success with correlation coefficients requires familiarity with the field of application as well as with their mathematical properties and that both the reliability and interpretation of r depend heavily upon the extent to which X and Y are jointly normally distributed.

4 LINEAR REGRESSION

As has been observed, empirical correlation methods are often useful in studying how two variables are related. It frequently happens, however, that one studies the relationship between the variables in the hope that any relationship that is discovered can be used to assist in making estimates or predictions of one of the variables. Thus if the two variables are the high