Response to Reviewer 2 for Assimilating Compact Phase Space Retrievals of Atmospheric Composition with WRF-Chem/DART: A Regional Chemical Transport/Ensemble Kalman Filter Data Assimilation System

Response to the Minor Comments:

1. p. 7698 l.2: truncation is an arbitrary process and requires clarification. At what thresholds are singular values/vectors set to zero?

Text will be changed. On page 7698, at line 3, change "trailing singular values" to "trailing singular values (i.e., the singular values that were less than or equal to 1.0 x 10^{-4})".

2. p. 7701 l. 16: it is not clear why increased vertical resolution would make the assimilation more sensitive to vertical localization. Do authors claim that the vertical lenghtscales fall within the grid spacing? Can that be elaborated on?

Text will be clarified. On page 7701, at line 16 after "vertical localization may increase" insert the following: "(because as the model's vertical resolution increases: (i) the vertical solution becomes less smooth and may exhibit greater vertical variability and (ii) the fidelity of vertical localization becomes greater)".

3. p. 7705 description of Fig. 1: comparison of forecasts and increments does not lead to conclusions as optimistic as the authors claim. E.g. difference MPO – MET DA is negative SW of Lakes Huron and Michigan and in Ohio Valley while the increment over this area is positive suggesting that the forecast issued from MET DA might have been superior to MPO there. Even over Bay area the sign of the difference is not as consistent with the increment as the authors claim. To somewhat lesser extent the same applies to Fig. 5

Text will be clarified. On page 7705, at line 5 after the end of the paragraph insert the following: "However, that is a general statement because the MOP QOR – MET DA differences are partially related to the impact of assimilating CO observations during the preceding assimilation cycle and partially related to the impact of assimilating all the CO observations since the beginning of the cycling experiment (~100 cycles). Consequently, there are locations where the signs of the MOP QOR – MET DA differences are different from the signs of the increments (e.g. southwest of Lakes Michigan and Huron and over the Ohio River Valley and San Francisco Bay). The sense of those sign differences is not an indication of relative forecast accuracy but that the: (i) impact from assimilating CO during the preceding cycle was similar to that from assimilating CO throughout the cycling experiment (same signs), and (ii) impact from assimilating CO during the preceding cycle was different to that from assimilating CO throughout the cycling experiment (different signs)."

4. p. 7705 description of Fig. 2: it would be much easier to see biases and correlations if scatter plots were shown rather than time series of dots that are somewhat difficult to follow.

Regarding the caption to Fig. 2, we appreciate the reviewer's comments. We had chosen not to use scatter plots to show the bias and correlation because we wanted to show the temporal trends. To address the reviewer's concerns, we propose to revise the figure caption. Replace the caption to Fig. 2 with the following: " Time series of the domain average CO from the MOP QOR and MET DA experiments. The red and magenta dots show the domain average CO in retrieval space for the MOP QOR and MET DA analyses respectively (denoted in the legend by "A"). The blue and black dots show the domain average CO in retrieval space for the MOP QOR and MET DA forecasts respectively (denoted in the legend by "F"). The green dots show the domain average MOPITT CO retrievals and are the same in both panels. The solid lines show the domain average CO in model space with the same color scheme as used for the analyses and forecasts in retrieval space. The solid lines are also the same in both panels."

5. p. 7706 l.11-19: difference between two results is not likely to be exactly zero. Can authors specify what hypothesis testing involved and how much would the means need to differ compared to what they were to reject/accept the hypothesis?

Text will be clarified. On page 7706, at line 15, after "two means from a normal distribution." insert the following: "The test statistic was $Z = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} - \frac{\sigma_2^2}{n_2}}}$ where $\bar{Y}_1, \sigma_1^2,$ and n_1 denote the sample mean, sample variance, and number of samples for the MOP QOR experiment; $\bar{Y}_2, \sigma_2^2,$ and n_2 denote the analogous sample statistics for the MET DA experiment; and $n_1 = n_2 = 104$. The rejection criteria was $|Z| > z_{\alpha/2}$ where $\alpha = 0.05$ and $z_{\alpha/2} = 1.96$ for a two-tailed test at the 95% confidence level."

Also on page 7706, at line 16 delete "at the 95% confidence level".

6. p. 7709 l. 1-13: authors talk about positive/negative sensitivities of singular vectors but later note that the sign of the vectors is arbitrary because left/right singular vectors can jointly by negative one which is true. However, do the considerations on the positive/negative sensitivities make any sense in this case. Can that be discussed?

Text will be clarified. On page 7709, at line 11, after "as a singular vector." insert the following: "However, when multiplied by one sign the singular may have physical meaning and when multiplied by the other it may not. For our application the sign that made the vertical structure of the singular vector most similar to that of the averaging kernel had physical meaning while the other did not."

Also on page 7709, at line 11, replace "In Fig. 4b" with "Therefore, in Fig. 4b".