ACP-2019-241 (Editor – Rachel Law)

Response to Reviewer 1

The authors thank the reviewer 1 for a thoughtful review of the manuscript. The responses for the reviewer's specific comments are as follows.

General Comment:

This paper describes an evaluation of different observing locations in Asia to infer Asian surface CO2 fluxes. The authors use the Carbontracker inversion system, with model generated pseudo-data, to assess different observing networks. They compare fluxes estimated with the existing network, with alternative networks based on random addition or relocation of sites, and the choice of sites using sensitivities from the inversion system. This contrasts with previous network design studies that use optimisation to locate the best observing sites, with higher computational cost.

Some aspects of the methodology need improved description, as described below. I have some concerns with the methodology, also described below, however it is possible that I have misunderstood what was done and improved description would give me a better understanding of the methodology and address some of my concerns. There is a need for minor improvements to the English throughout, but this would be addressed with copy-editing and I don't believe it has contributed to any difficulty in my understanding of the methodology or results.

Author's response: Following the reviewer's suggestions, we have tried to improve descriptions. We also have addressed more explanations for the concerns with the methodology. Specific responses to the reviewer's comments and revisions are shown below.

Specific Comments:

1. Self-sensitivity (page 4, line 6 and section 2.2) - it is not completely clear to me what sensitivity is used in this paper. On page 4, line 8 'the relative impact of each CO2 observation for the optimized surface carbon flux can be calculated ... and used as a strategy for selecting potential sites of CO2 mole fraction observations', however on page 8, line 8 'contribution of the observation vector (yo) to the analysis at the observation space (ya)'. From the description in the paper, I understand ya to be the model equivalent of the CO2 mole fraction in air, also described as the predicted observation in Liu et al (2009) that the authors refer to. These are two different quantities (i.e. sensitivity of fluxes or sensitivity of surface mole fraction). Which was used in this study? I can see value in considering the *sensitivity of the optimised flux (or perhaps the scale factor in this study) to each observation, but I am not as clear on the value of the sensitivity of the predicted observation. Of course they are related, but not the same. I am also not clear about how time affects the sensitivities. For example, some information comes from distant sites but with a lag. When is the analysis sensitivity calculated -before lagged information has had a chance to improve an analysis estimate? If so, that would downweight information from other gridcells that arrive after a lag. Thus I have concerns about the methodology, but I admit that it is not clear to me exactly what was done.*

Author's response: The self-sensitivity is calculated in the observation sites. As denoted in Eq. (11), the self-sensitivity is the gradient of the analysis at the observation space ($\boldsymbol{y}^{\mathrm{a}}$) to the observation vector ($\boldsymbol{y}^{\mathrm{o}}$). Here, $\boldsymbol{y}^{\mathrm{a}}$ represents the projection of analysis state vector $\boldsymbol{x}^{\mathrm{a}}$ on the observation space or model analysis equivalent to observations at observation locations. The model analysis $\boldsymbol{x}^{\mathrm{a}}$ (i.e., optimized surface CO_2 flux) is on the model grid point, whereas the model analysis equivalent to observations $\boldsymbol{y}^{\mathrm{a}}$ (i.e., model analysis equivalent CO_2 mole fraction) is on the observation locations (i.e., observation space). As the reviewer denoted, they are not the exactly same although they are closely related as in Eqs. (11), (12), and (13). The self-sensitivity represents the contribution of observations to the model analysis in grid point as well as that in observation locations. Liu et al. (2009) deals with the sensitivity in numerical weather prediction (NWP) problem, thus it considers predicted observation. However, in this study, the analysis equivalent observation is used because the prediction is not much considered in CO_2 data assimilation. Although the self-sensitivity is qualitatively related with both the model analysis in grid point as well as that in observation locations, the quantity used in this study is the sensitivity of "model analysis equivalent CO_2 mole fraction at observation space" to "observed CO_2 mole fractions". Thus, we have revised the text on page 4, line 8 as follows.

"Similar to the numerical weather prediction (NWP), the relative impact of each CO_2 mole fraction observation for the model analysis equivalent CO_2 mole fraction induced by the optimized surface CO_2 flux can be calculated (Kim et al., 2014a, 2017) and used as a strategy for selecting potential sites of CO_2 mole fraction observations."

Since the analysis (i.e., optimized surface CO_2 flux) in this study is calculated considering the time lag, the effect of time lag is already included in the sensitivity calculation. As mentioned in Kim et al. (2014a), CarbonTracker adopts a smoother window to reflect the transport speed of CO_2, which is based on the temporal relationship between the surface CO_2 flux and atmospheric CO_2 observations, as found in Bruhwiler et al. (2005) (Peters et al., 2005). For this reason, the scaling factor is optimized for 5 weeks of lag, which implies that the observations made in the most recent week affect the optimized surface CO_2 flux in the preceding 4 weeks. The optimization of the scaling factor during the data assimilation process is presented in Fig. 1 in Kim et al. (2014a) as shown below. In each assimilation cycle, 5 weeks of analysis scaling factors are estimated by observations from the most recent week. After the fifth cycle, the scaling factor estimated by these 5 weeks of observations is saved as the optimized scaling factor and used to calculate the optimized (i.e., analysis) surface CO_2 flux.

The self-sensitivity is calculated using the analysis produced by the process above. Thus, the analysis surface CO_2 flux already considers the time lag associated with the distant information. Whether the 5 weeks of lag is enough to consider the distant information is fully studied in Peters et al. (2007) and Kim et al. (2018b).

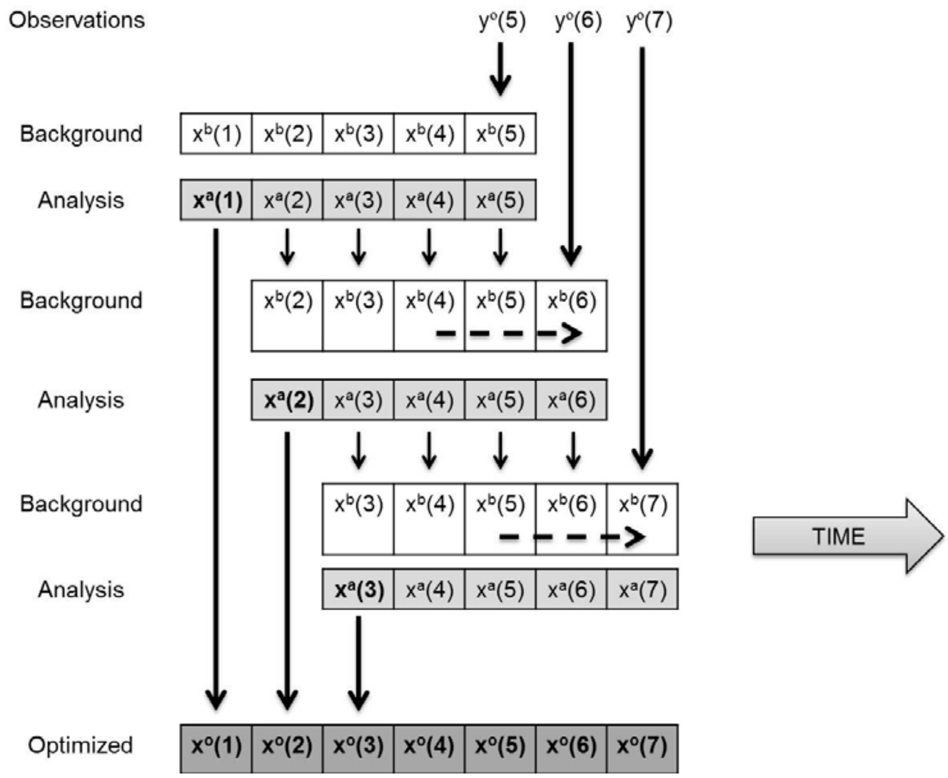


Figure 1. Schematic diagram of the assimilation process employed in CarbonTracker. In each analysis cycle, observations made within one week are used to update the state vectors with a five-week lag. The dashed line indicates how the simple dynamic model uses analysis state vectors from the previous one and two weeks to produce a new background state vector for the current analysis time. The TM5 model is used as the observation operator to calculate the model CO_2 concentration for each corresponding observation location and time. (Kim et al. 2014a, ACP)

2. Simulated hypothetical observations (section 2.3) - there are no details or references given about the EXTASI experiment - are the EXTASI fluxes based on the same flux modules as used in this study but with different scale factors? Therefore, are there differences in the spatial distribution of fluxes within the ecoregions that are used to generate the simulated hypothetical observations compared to the flux modules, as there would be between modelled *fluxes and real world fluxes in an inversion of real observations? This is perhaps most relevant for the two regions that each account for close to 20% of the domain. If the distribution is the same, that's probably ok, but it should be mentioned, as model error in the spatial distribution within each region is not considered.*

Author's response: In EXTASI experiment, the surface CO_2 fluxes are optimized by the inverse modeling using the real observation data (i.e., observed CO_2 mole fractions). Thus, EXTASI produced optimized surface CO_2 flux (from the inverse modeling using real CO_2 mole fraction observations). In contrast, SF1 experiment produced another estimated surface CO_2 flux (by setting scaling factor as 1). Thus, as the reviewer mentioned, EXTASI and SF1 are based on the same prior flux modules with different scale factors. The EXTASI produces fluxes that is closer to real fluxes (although the real fluxes are not exactly known), compared to SF1.

Using the above two CO_2 fluxes, two model CO_2 mole fractions were generated and averaged to have the hypothetical true CO_2 mole fraction observations. We made hypothetical true CO_2 mole fractions this way because we liked to produce hypothetical "true data" close to real data but not the same. If the model CO_2 mole fractions produced by EXTASI are used as "true data", then they may be similar to the real observed CO_2 mole fractions, but they are constrained much by the real observation network. This configuration causes that, when we choose observation sites using several strategies, the experiment using the current observation network (i.e., CNTL in this study) has more benefits compared to other network designs. To be fairly compared the results from several network configurations, we have made hypothetical true CO_2 fractions that are somewhat similar to the real feature but still hypothetical.

Following the reviewer's suggestion, we have clarified and added details of the EXTASI experiment as follows. The revised parts are underlined.

"In this paper, simulated hypothetical observations were created and used to design the observation network. Simulated hypothetical observations with similar values and seasonal variations compared to real CO_2 observations were generated by <u>averaging</u> model CO_2 mole fractions from the experiment conducted with real NOAA observation data (EXTASI) and <u>model CO_2 mole fractions from</u> the experiment with a fixed scaling factor of 1 (SF1). <u>Both EXTASI and SF1 experiments were done for the year of 2008. In EXTASI experiment, the real CO_2 mole fraction data were used to update the scaling factors in Eq. (1) to estimate the surface CO_2 fluxes. In contrast, in SF1 experiment, the scaling factors were fixed as 1.</u>

Figure 2 shows the station-averaged time series of CO_2 mole fractions from real observations (OBS), EXTASI, SF1, and an average <u>(i.e., simulated hypothetical observations: TRUE, hereafter)</u> of EXTASI and SF1."

3. Average of random redistribution (page 11, line 1) - My understanding from the text is that REDIST is created by averaging the fluxes from three random redistribution experiments of 7 sites. Firstly, is this correct? And if it is, I am concerned that this may lead to a better solution than you would expect from just 7 sites, as 7x3=21 sites were actually used to generate the average. Errors in the individual results may cancel in the average. The statistics of the average may not reflect the statistics of individual experiments, and therefore it would be an unfair comparison. The ADD case is also an average of three experiments, so would potentially have the same issue. Perhaps it would be a fairer comparison to instead calculate the PC, BIAS, RMSD and UR statistics for the individual experiments then average these statistics?

Author's response: As the reviewer mentioned, we did three random redistribution experiments and averaged the results. In each experiment, 7 sites were used and statistics (i.e., PC, BIAS, and RMSD) were calculated. This experiment was done three times independently. In each experiment, 7 sites were selected randomly. We calculated average statistics for three experiments rather than statistics for individual experiment since the statistics of individual experiment can be easily skewed by specific configuration or selection of sites by only one experiment. To avoid the sampling error that can possibly be caused by only one sampling, we did three experiments with different configurations to get more general results. This experimental configuration is used in previous observation network studies as Yang et al. (2014). Thus, we have revised the text as follows. The added parts are underlined.

"Figures 3b, c, and d show the distribution of three observation networks, in which the seven observation sites in Asia are randomly redistributed. <u>To obtain general results without sampling error, each random redistribution experiment was performed three times with different sets of randomly distributed observation sites, as denoted in previous observation network studies (e.g., Yang et al. 2014).</u> The average of three random redistribution experiments was denoted as REDIST, to check the impact of the reallocation of the existing observation network."

4. What affects the self-sensitivity of an individual gridcell in the ALL case? In Fig 6, most of the gridcells with high self-sensitivity are near the boundaries of the regions used for this calculation. Presumably this is because they contain some information not available from neighboring gridcells in the ALL case. But for gridcells with many neighbors that contain similar information to each other, the information from any one of those gridcells may not be needed when all of the others are available, as in the ALL case. But that does not mean that at least some of these gridcells that rank low in the ALL case are unimportant in a case with a much lower number of observing sites. The authors point out on page 17, line 18 that self-sensitivity is generally inversely proportional to the number of assimilated observations in an ecoregion, and that makes sense, but within a region, does the self-sensitivity pick out some sites that will give most value in a network with few sites, or just those with most sensitivity in *a case with many sites (ALL)? In network design studies that use optimisation, the value of observation sites is determined for a network that is closer to the expected size of the potential network. I am not yet convinced of the value of determining the worth of any single site from the self-sensitivity in the ALL case when many more sites than would be practical are included. This is my greatest concern about the methodology, and I believe this would need to be addressed for the paper to be published. Of course, exactly what the self-sensitivity is (sensitivity of fluxes or surface mole fraction) is also important here (see above comment).*

Author's response: As mentioned in page 17, four influential regions with high sensitivities are located in western Siberia, the southern part of the Tibetan Plateau, and southeastern and northeastern Asia. Except the western Siberia, the other regions do not coincide much with the boundaries of the model domain. As defined, the self-sensitivity represents how CO_2 mole fraction observations affect the model analysis equivalent CO_2 mole fraction observations at observation sites. Since the model grid points at 2° intervals are the observation sites in ALL experiment, if the self-sensitivity value at some grid point is large, then the observation at that grid point will affect highly the model analysis equivalent of CO_2 mole fraction observations. All grid points in ALL experiment are in same condition: 1) the observation sites at every 2° intervals on the land are used, 2) at every sites, only one simulated CO_2 mole fraction values around afternoon (i.e., 13 local standard time (LST)) are assimilated per day for one year), and 3) the self-sensitivities at every sites are calculated and averaged as shown in Fig. 6. Thus, the self-sensitivity based on these same conditions can be a measure to determine which observation sites should be used for assimilation to have a large effect on the model analysis equivalent at observation sites.

In Page 17 line 18, we mention that the self-sensitivity is generally inversely proportional to the number of assimilated observations, as shown in Kim et al. (2014a, 2017). This inversely proportional relationship is for observation numbers and self-sensitivity at one site, not for many observation sites case vs few or no observation sites case.

The genetic algorithm (GA) method which considers many sets of observation networks and finds out the network with minimum error with much computations, can be considered as the forward approach. In contrast, the method in this study uses the backward approach that can calculate the contribution of observations to the analysis equivalent mole fractions with much smaller computation. Using the self-sensitivity, we can select the potential observation site one by one. Practically, redistributing all observation sites at once is not easy or is even impossible. Adding or redistributing some sites given existing observation sites may be a more practical way to design the observation network. Once we have self-sensitivity value, we can use the value to determine the observation sites that would affect much on the analysis results. Using the self-sensitivity to determine the potential sites, we did forward calculation to verify whether the sites by the strategy are good or not.

The definition of the self-sensitivity is already explained in detail in the response to the specific comment 1 above.

Minor points:

5. Page 1, line 10 - "Inverse modeling .... derives estimated CO2 mole fractions in the air from calculated surface carbon fluxes using model and observed CO2 mole fraction data" - No, forward modelling derives CO2 mole fractions in the air from surface fluxes. Inverse modeling derives surface fluxes from CO2 mole fractions in the air.

Author's response: We have revised the confusing description of the inverse modeling as follows. The revised parts are underlined.

"Continuous efforts have been made to monitor atmospheric CO_2 mole fractions as it is one of the most influential greenhouse gases in Earth's atmosphere. The atmospheric CO_2 mole fractions are mostly determined by CO_2 exchanges at the Earth's surface (i.e., surface CO_2 flux). Inverse modeling, which is a method to estimate the CO_2 exchanges at the Earth's surface, derives surface CO_2 fluxes using model and observed atmospheric CO_2 mole fraction data."

6. Page 2, line 7 - "Inverse modeling .... uses observation data and transport models to estimate the sources and sinks of surface carbon flux and associated atmospheric CO2 mole fractions" - better than the previous description, but doesn't specify what observation data are used (should be CO2 mole fractions in air). The associated modelled atmospheric CO2 mole fractions can be estimated from the inferred fluxes (or perhaps during the inversion), but I don't consider that part of the inverse calculation.

Author's response: We have revised the text as follows. The revised parts are underlined.

"Inverse modeling, one of the methods to complete this mission, uses observed atmospheric CO_2 mole fraction data and transport models to estimate the sources and sinks of surface CO_2 flux (Enting, 2002; Gurney et al., 2002)."

7. Page 3, line 14 - Add 'alone' after 'data' i.e. Assimilating XCO2 data alone ...

Author's response: We have added the text following the reviewer's suggestion.

8. Page 3, line 22 and many other locations - OSSEs (with an 's' at the end) is often used for the plural of OSSE. I.e. We conducted one OSSE, and they conducted many OSSEs.

Author's response: As many other references (e.g., Wang et al. 2018), we have used

“OSSEs” for the plural of “OSSE”. Considering the reviewer’s suggestion, in the last paragraph of Section 1, we have replaced “OSSEs” by “many OSSEs” as follows.

“In this study, <u>many OSSEs</u> were conducted using CarbonTracker (CT) to~”

9. Page 4, line 12 - "which does not seem feasible in the near future" - what is meant here? Is the 43 site network not feasible? Or the 233 site network (is this not like the ALL case considered here, to see what would be possible with observations everywhere)? Or are the authors referring to the computation of the network design calculation for many sites?

Author’s response: We meant that many $^{14}CO_2$ sites may not be feasible in the near future in Asia. But we found that the original meaning does not fit in the paragraph well. Thus, we have revised the text as follows.

“Although Wang et al. (2018) showed the potential impact of adding observation sites on the existing $^{14}CO_2$ sites in Europe using OSSEs, the potential $^{14}CO_2$ observation sites were not chosen based on specific selection strategies.”

10. Page 4, line 24 - I would add at the end of the sentence ‘, as an alternative to optimisation that has been used in previous studies’ to make it clear that optimisation is not used in this study. Alternatively (or perhaps in addition), point out clearly elsewhere in the introduction that optimisation of the network is not part of this study, as that point was initially not clear to me. (At page 4, line 3, problems with IO and GA are discussed, but that doesn’t mean another optimisation method wasn’t going to be used).

Author’s response: In my knowledge, the IO and GA are methods to select observation sites for observation network design until now. We could not find other methods used for determining surface CO_2 observation network. The IO and GA are strategies selecting observation sites to minimize the error in their own framework. The IO and GA are called as the optimization method, but one of them has lower error than another in specific cases (Nickless et al. 2015).

Instead of using the term “optimization”, we proposed a selection strategy based on influence matrix. If the observation network is designed in the region without observation sites, then the IO and GA methods would be useful. When adding observation sites over the region with existing observation sites, redistributing all observation sites at once may not be easy or may be even impossible. Adding or redistributing some sites given existing observation sites may be a more practical way to design the observation network.

Considering the reviewer's suggestion, we have revised the text as follows. The added parts are underlined.

“In the case of addition experiments, random addition and addition based on influence matrix (self-sensitivity) as well as ecoregion information of the model were considered as strategies, as alternatives to IO and GA that have been used in previous studies."

"Due to time and computing restraints, the IO and GA methods seem ineffective or unfeasible for designing the observation network on continental scales like Asia. In addition, determining and redistributing all observation sites at once using the IO and GA methods may not be practical for most regions with existing observation sites. Adding or redistributing some sites given existing observation sites may be a more practical way to design the observation network."

We have already mentioned that our purpose is to identify "a better" in situ observation network for optimizing surface CO_2 flux estimation in Asia. The text is shown in page 4, line 20, as follows.

"In this study, many OSSEs were conducted using CarbonTracker (CT) to identify a better in-situ observation network for the purpose of optimizing surface CO_2 flux estimation in Asia."

11. Page 5, section 2.1 - There are many details of the inversion that are not clear: Does the inversion run globally with a focus on Asia, or just run over Asia as a regional inversion (i.e. are fluxes outside the Asian domain estimated)? How many ecoregions are used in this study? (Is 156 regions a global number or for Asia? What are the 240 ecoregions? There are 40 lines in Table 3, is that the number for Asia? Could say 'We estimate x scale factors for y times'.) Is it possible to include a map of the ecoregions for Asia? How contiguous are the ecoregions?

Author's response: The inversion run is done globally with a focus on Asia using a nesting domain over Asia. The ecoregions used is 156 regions globally, 40 in the verification region (black dashed box in Fig. 1). As mentioned in the manuscript, 240 is the number of total ecoregions of the earth including ocean and unused vegetation types. The number of effective ecoregions globally is 156, and the 40 is the number of ecoregions in the verification region. The 40 ecoregions include mostly ecoregions of Asia and a very small portion of ecoregions of Europe. Since the portion of ecoregions (i.e., ecoregion indices of 191, 193, 194, 197, 201 in Table 3) of Europe is approximately 0.5% of the verification region, including ecoregions of Europe does not affect much on the verification results. To clarify, we have included Transcom region as well as Land ecosystem type in Table 3. In addition, we have revised the text as follows. The added parts are underlined. We also have included the map of ecoregions for Asia in Fig. 1b.

"This means that the optimization of the scaling factors that were assigned to the ecoregions of the earth is crucial for the estimation of simulated surface CO_2 fluxes. The ecoregions are defined as the mix of the modified 19 vegetation types from Olson et al. (1992) and 11 Transcom regions (Gurney et al., 2002) on land, with 30 ocean regions. As all 19 vegetation types are not used for the 11 Transcom regions, the number of effective ecoregions of the earth is 156 (Peters et al., 2010)."

"The horizontal resolution of TM5 is 3° x 2° globally and the nested horizontal grid is 1° x 1° over Asia, with verification region inside of the nested domain over Asia (Fig. 1). The number of ecoregions of the verification region is 40, in which 36 are the Asian ecoregions and 4 are the ecoregions of Europe. Since the proportion of the 4 European ecoregions is approximately 0.5% of the verification region (Table 3), the verification region was considered to be located over Asia. A two-way nested grid was used to optimize surface CO_2 fluxes in Asia. The model run including both forward and inversion runs was done globally with nesting over Asia and verification was done over the verification region located in Asia."

12. Page 5, line 9 - I would mention up front that the fluxes from the flux modules are scaled, and not wait until line 19. e.g. at line 9 'The estimated surface CO2 fluxes are mainly calculated by scaling fluxes from the flux modules composed ...'

Author's response: We have revised the text as the reviewer suggested.

13. Page 5, line 28 - the sentence that begins 'In addition, also ....' is not clear. It does not say what the model counterparts are. I would replace that sentence with something like 'From this spatiotemporal CO2 distribution, the model equivalents of atmospheric CO2 at the times and locations of the observation data can be calculated, and these are used in the data assimilation process.'

Author's response: To clarify, we have revised the text considering the reviewer's suggestion.

"In addition, from this spatiotemporal CO_2 distribution, the model atmospheric CO_2 concentrations at the times and locations of the observation data are calculated, and these are used for the data assimilation process."

14. Page 7, line 17 - I would say 'A statistical method' rather than 'The statistical method', otherwise a reader would wonder which method is 'the' method. I would replace 'feasible' with 'meaningful'.

Author's response: We have revised the paragraph including the text as follows. We kept "feasible" since what we meant is "possible". The revised parts are underlined.

"In the EnSRF, the covariance localization method is necessary to reduce the impact of the sampling error due to the limited size of the ensemble and to avoid filter divergence due to the underestimation of the background error covariance (Houtekamer and Mitchell, 2001). Because calculating the physical distance between scaling factors is not feasible, instead of the covariance localization method, a statistical method is applied in this study."

15. Page 8, line 8 - define ya (e.g. = Hxa) and give some information about what it is (e.g. model equivalent of observations, or predicted observation).

Author's response: We have revised the text as follows. The added parts are underlined.

"The analysis of the state vector and the influence matrix ($\boldsymbol{S}^{\mathrm{o}}$) that shows the contribution of the observation vector ($\boldsymbol{y}^{\mathrm{o}}$) to the analysis at the observation space ($\boldsymbol{y}^{\mathrm{a}}$) (i.e., the projection of analysis state vector $\boldsymbol{x}^{\mathrm{a}}$ on the observation space or model analysis equivalent to observations at observation locations) can be defined as:"

16. Page 8, line 12 - replace 'size of observation' with either 'size of the observation vector, n' or 'number of observations, n'. Is that the number of observations at only one time or all times?

Author's response: The original text was wrong. The dimension of $\mathbf{I}_n$ corresponds to the dimension of the analysis state vector x^{a}. Thus, we have revised the text as follows. The revised parts are underlined.

"where, $\mathbf{I}_n$ is the identity matrix with the size of *n*-dimensional analysis state vector."

17. Page 8, line 25 - do you (and should you) assume no correlations between observation errors? It seems to me that the errors in your simulated data would be correlated, and also likely in the real world.

Author's response: The observation errors in data assimilation are usually assumed to have no correlations. Although the observation errors in real world would be correlated, this no correlation assumption is very common and may be only way in this ensemble sensitivity study and data assimilation. Using this assumption, $\mathbf{R}^{-1}$ in Eq. (12) can be simplified as $\frac{1}{\sigma_j{}^2}$ in Eq. (14). Please note Liu et al. (2009).

18. Page 9, line 6 - please give more explanation of what So is, e.g. 'In our case, this would be the contribution of a CO2 observation to the inferred CO2 at that model gridcell/time' - is that the correct explanation?

Author's response: We have revised the paragraph including the line as follows. The revised parts are underlined.

"According to Liu et al. (2009) and Kim et al. (2014a), $\mathbf{S}^{o}$ represents the sensitivity of the analysis state vector $\boldsymbol{y}^{a}$ to the observation state vector $\boldsymbol{y}^{o}$ in the observation space (i.e., location). $\mathbf{S}^{o}$ has a value between 0 and 1, which shows the contribution of a CO_2 observation to the analyzed CO_2 at the observation site. If $\mathbf{S}^{o}$ is close to 0, the analysis is mainly derived from the background. In contrast, the influence of observation data to the analysis increases as $\mathbf{S}^{o}$ goes to 1. The self-sensitivity was used as a criterion for selecting the observation locations in designing the observation network."

19. Page 10, line 2 - please explain 'On the basis of the nautical time zone'. Also explain '13 LST'.

Author's response: The text means the 13 local standard time (LST) (i.e., afternoon in local standard time) at each time zone. To clarify, we have revised the text as follows. The revised parts are underlined.

"The simulated values around afternoon (i.e., 13 local standard time (LST)) in the mid-latitudes in the northern hemisphere are averaged and utilized as TRUE data."

20. Page 10, line 9 - 'Model-data mismatch (MDM) was set to 3' - what does the setting of 3 mean? Is it a setting within Carbontracker, in which case it should be explained.

Author's response: MDM corresponds to the observation error covariance in data assimilation. The observation error for CO_2 mole fraction observations at continuous surface observation sites is set to 3 ppm in CarbonTracker. The number 3 ppm for continuous surface observation sites is usually used in other inversion modeling system, either. To clarify, we have added text (underlined) as follows.

"Model-data-mismatch (MDM) (i.e., observation error) for CO_2 observation was set to 3 ppm, consistent with the previous setting of 3 ppm for continuous observation site types (Peters et al., 2007; Kim et al., 2014b, 2017)."

21. Page 11, line 25 - add 'for observation j' to 'The normalized self-sensitivity for observation j is defined...'

Author's response: For consistency, we have revised the text as follows.

"The normalized self-sensitivity for jth observation ~"

22. Page 12, section 2.4 - define n for equations 16-18.

Author's response: "*n*" is the total number of model grid-point in the verification domain shown in Fig. 1. Thus, we have revised the text as follows. The added parts are underlined.

"where EXP_i and $TRUE_i$ are the surface CO_2 fluxes at the ith model grid-point of an experiment and TRUE, respectively, <u>and *n* is the total number of model grid-point in the verification domain shown in Fig. 1.</u>"

23. Page 15, line 10 - 'the three experiments show increasing trends' - be careful that this in not misinterpreted as a trend with time. I assume you mean that for RMSD in the summer, CTRL>ADD>ALL? Please clarify what is meant here.

Author's response: We have revised the text as follows. The revised parts are underlined.

"In terms of the RMSD, the three experiments show <u>larger values</u> in the summer compared to other seasons (Fig. 5c)"

24. Page 16, line 9 - what does 'enabled in the CT2013B framework' mean? There may be a better way to express this.

Author's response: We have revised the text as follows. The revised parts are underlined.

"In particular, the ALL experiment, which added many observation sites <u>under the given modeling framework</u>, shows a high level of reproducibility of TRUE."

25. Page 16, line 16 - 'showing the impact of each observation site for the model simulation results' - could you be more specific here about what quantity the impact of the observation sites is calculated for.

Author's response: As denoted in Section 2.2, the self-sensitivity is the sensitivity of inverse model results with respect to the observations assimilated in EnKF data assimilation system in CarbonTracker. For each observation at each observation site, the self-sensitivity is calculated. For each observation site, all self-sensitivities are added up to have self-sensitivity at that site. To clarify, we have revised the text as follows. The revised parts are underlined.

"Since the self-sensitivity is the metric showing the impact of <u>observations at</u> each observation site for the model simulation results, as stated in Sect. 2.2, ~"

26. Page 17, line 13 and Table 3 - could the ecoregions be described in terms of vegetation types rather than just as a number which may not mean anything to the reader?

Author's response: We have included ecosystem types in text and tables (Tables 3, 4, and 5). We also have added Transcom region information in Tables 3, 4, and 5.

27. Page 18, line 8-9 - These sentences are difficult to follow, consider rephrasing without the 'this is in contrast' beginning to each new sentence.

Author's response: We have revised the text as follows. The revised parts are underlined.

"Nevertheless, the ECOSS experiment that considered both self-sensitivity and ecoregion information maintains lower RMSD than the ADD experiment over the experimental period. Additionally, except in the period from April to late-August, the RMSD of SS is lower than that of ADD, which differs from ADD that is mainly better than CNTL in summer, as shown in Fig. 5. Thus, compared to ADD and CNTL, the SS (ECOSS) experiment demonstrates improvement in the other seasons except summer (over the experimental period)."

28. Page 19, line 5 - 'because they were derived from an uneven distribution of observation sites' - do you mean an uneven number of sites for each ecoregion?

Author's response: We have revised the text as the reviewer indicated.

29. Page 19, line 13 - add 'each' after 'one observation site'

Author's response: We have added the word following the reviewer's suggestion.

30. Page 20, line 5 - I don't think 'and this is in contrast' is the appropriate wording here.

Author's response: We have revised the text as follows. The revised parts are underlined.

"The NSS, NECOSS1, and NECOSS2 experiments show lower RMSDs compared to the ADD experiment (Fig. 8c). The RMSD of NSS is lower than that of ADD for most of the time, which is different from SS that showed a degradation in summer and little improvement in other seasons compared to ADD in Fig. 7c."

31. Page 21, line 18 - replace 'below 50oN' with 'north of 50oN'

Author's response: We have changed "below 50° N" with "south of 50° N" since "south" is what we meant.

32. Page 21, line 20 - replace 'slight increases in UR' with 'slightly more UR'

Author's response: We have revised the text following the reviewer's suggestion.

33. Page 21, line 21 - add 'than REDIST' after 'including China and India'.

Author's response: We have added the text following the reviewer's suggestion.

34. Table 6 - Bias in Fig 7 looks like it is lower for ADD than SS and ECOSS – is this consistent with the numbers in Table 6? Is the signed biased averaged, or the magnitude?

Author's response: As shown in Eq. (17), the BIAS is calculated as the average of summed differences between experiment results and truth. Thus, the differences have signs, and those signed values at model grid points are summed and averaged.

The BIAS values of ADD show many positive and negative values and they are cancelled out when summed over model grid points. In contrast, the BIAS values of SS and ECOSS experiments show dominant specific signs. In this case, the BIAS values are not cancelled out and show large values with a certain sign (positive or negative). In Fig. 7, the SS shows large positive BIAS on June 7. In this case, the SS shows positive BIAS on most of the grid points, thus they are added up to have large positive BIAS. In contrast, the ADD shows relatively small BIAS, but the BIAS values of ADD on that day on the grid points are not small in magnitude with different signs, thus they are cancelled out when summed.

In contrast, the RMSD considers the magnitude of BIAS as in Eq. (18). Thus, in terms of the magnitude of the error, we have to look at the RMSD instead of the BIAS. Although some BIAS values of SS and ECOSS show large magnitude on specific days, the average values of BIAS and RMSD of SS and ECOSS are smaller than that of ADD.

Therefore, the BIAS in Fig. 7 and Table 6 are consistent.

35. Figs 3m and 6 - the gap in observing sites in Figs 3m and 6 over the Himalayas is presumably due to elevation and therefore practicality of an observing site? Is this worth mentioning?

Author's response: We have omitted locations that are 2000 m above the mean sea level considering the maintenance of the observing sites. We have added the reason for the gap in observing sites over the Tibetan Plateau as follows. The added parts are underlined.

"In addition, the observation networks that have observation sites at every 2° intervals on the land (Fig. 3m, ALL experiment) are suggested as the reference to examine the maximum possible impact of additional observation sites. In ALL experiment, the observation locations that are located 2000 m above the mean sea level over the Tibetan Plateau are not included due to difficult accessibility and maintenance as practical observing sites."

References

Kim, H., Kim, H. M., Kim, J. and Cho, C.-H.: Effect of data assimilation parameters on the optimized surface CO_2 flux in Asia, Asia-Pacific J. Atmos. Sci., 54(1), 1–17, doi:10.1007/s13143-017-0049-9, 2018b.

Kim, J., Kim, H. M. and Cho, C. H.: Influence of CO_2 observations on the optimized CO_2 flux in an ensemble Kalman filter, Atmos. Chem. Phys., 14(24), 13515–13530, doi:10.5194/acp-14-13515-2014, 2014a.

Kim, J., Kim, H. M., Cho, C. H., Boo, K. O., Jacobson, A. R., Sasakawa, M., Machida, T., Arshinov, M. and Fedoseev, N.: Impact of Siberian observations on the optimization of surface CO_2 flux, Atmos. Chem. Phys., 17(4), 2881–2899, doi:10.5194/acp-17-2881-2017, 2017.

Nickless, A., Ziehn, T., Rayner, P. J., Scholes, R. J. and Engelbrecht, F.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling - Part 2: Sensitivity analyses and South African test case, Atmos. Chem. Phys., 15(4), 2051–2069, doi:10.5194/acp-15-2051-2015, 2015.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L. and Tans, P. P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res. Atmos., 110(D24), D24304, doi:10.1029/2005JD006157, 2005.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C. and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proc. Natl. Acad. Sci., 104(48), 18925–18930, doi:10.1073/pnas.0708986104, 2007.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., StröM, J., Haszpra, L., Meijer, H. A. J., van Der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M. and Tans, P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Glob. Chang. Biol., 16(4), 1317–1337, doi:10.1111/j.1365-2486.2009.02078.x, 2010.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R. and Tao, S.: Potential of European ^{14} CO_2 observation network to estimate the fossil fuel CO_2 emissions via atmospheric inversions, Atmos. Chem. Phys, 185194(July), 4229–4250, doi:10.5194/acp-18-4229-2018, 2018.

Yang, E.-G., H. M. Kim, J. Kim, and Kay J. K.: Effect of observation network design on meteorological forecasts of Asian dust events, Monthly Weather Review, 142, 4679-4695, doi:10.1175/MWR-D-14-00080.1, 2014.

ACP-2019-241 (Editor – Rachel Law)

Response to Reviewer 2

The authors thank the reviewer 2 for a thoughtful review of the manuscript. The responses for the reviewer's specific comments are as follows.

General Comment:

This paper describes an impact of observation network against carbon cycle estimation by using CarbonTracker (CT). An important aspect is that the authors showed realistic solution this means that we have the potential to realize this observation network in the future. This viewpoint is very important, and I think it is necessary to advance research in this field in the future. I'd like to comment from a different perspective than Reviewer 1. I think authors need to do some additional experiments to take advantage of the excellent features of this paper. One important issue is that the authors show that root mean square error increases in many experiments in summer, but the reason is not well specified. Authors should consider this reason and suggest ways to reduce large summer uncertainty, if possible, without using ALL observations. The other issue is that authors should use observation sites registered in NOAA ObsPack but not assimilated in CT. This is because these stations are in operation and can be a precondition to be considered when considering future network expansion. The last issue is that this paper focuses on only ground observation network. Although it is considered unrealistic to use all observation points (ALL), the OSSE should be implemented in consideration of the observable area of satellite which can supply much more observation area than ground observation network even if the observation accuracy is inferior, if the authors want to evaluate the construction of a more realistic carbon cycle observation network.

Author's response: We have added discussions for large summer uncertainties. The specific discussions can be found in the responses to the specific comments 3 and 5 below.

We have added the Section 3.6 that considers many observation sites registered in NOAA ObsPack. The results based on the observation sites registered in NOAA ObsPack are very similar to those based on 7 observation sites used for CT2013B.

The purpose of this study is introducing the selection strategy (i.e., self-sensitivity and ecoregion information) for potential observation sites. Thus, we have examined these strategies using the ground observation sites. For the satellite observations, the self-sensitivity for the satellite observations are not well known yet. We have to know characteristics, spatial, and temporal distributions of self-sensitivities for satellite observations and how the self-sensitivities of satellite observations are different from those of the ground observations. Thus, the observation network design using both the ground observations and satellite observations will be studied in the future after the self-sensitivities for satellite observations are fully studied. In Section 4, we have added texts (underlined) considering this issue as follows.

"Although the simulation results showed an improvement in performance, the results also suggested that adding 10 extra observation sites in Asia may not be sufficient to fully optimize surface CO_2 fluxes, and more observation sites are required. Reliable observation data from some satellite sensors could supplement the model simulations on the basis of continuous surface observation sites. As the quality of satellite observation data increases, the observation network design for both surface and satellite observation data using the strategies (i.e., normalized self-sensitivity and ecoregion information) of this study will be investigated in the future."

Specific Comments:

1. Page 3, line 20: As we can expect an increase in satellite observation data and quality improvement in the future, so it is necessary to consider the mixed use of ground observation data and satellite observation data.

Author's response: The purpose of this study is showing the validity of the selection strategy (i.e., self-sensitivity and ecoregion information) for potential observation sites. Thus, we have examined these strategies using the ground observation sites. For the satellite observations, the self-sensitivity for the satellite observations are not well known yet. We have to know characteristics, spatial, and temporal distributions of self-sensitivities for satellite observations and how the self-sensitivities of satellite observations are different from those of the ground observations. Thus, the observation network design using both the ground observations and satellite observations will be studied in the future after the self-sensitivities for satellite observations are fully studied. In Section 4, we have added texts (underlined) considering this issue as follows.

"Although the simulation results showed an improvement in performance, the results also suggested that adding 10 extra observation sites in Asia may not be sufficient to fully optimize surface CO_2 fluxes, and more observation sites are required. Reliable observation data from some satellite sensors could supplement the model simulations on the basis of continuous surface observation sites. As the quality of satellite observation data increases, the observation network design for both surface and satellite observation data using the strategies (i.e., normalized self-sensitivity and ecoregion information) of this study will be investigated in the future."

2. Page 3, line 25: The authors should refer Patra et al., 2003 as this paper showed global CO2 observation network design.

Author's response: We have referred Patra et al. 2003 as follows. The added text is underlined.

"Observation system simulation experiments (OSSEs), using simulated observation data, provide an opportunity to evaluate the impact of observation data from the current and potential observation sites on the performance of the modeling system (<u>Patra et al., 2003</u>; Yang et al., 2014; Byrne et al., 2017; Wang et al., 2018)."

3. Page 9, Line 14 -15: The difference between hypothetical observations (TRUE) and real observation (OBS) is large in the summer, and this seems to be a cause of the increase in summer RMSD in each subsequent experiment. In order to analyze the cause of the increase in summer RMSD, another observation data that is close to actual observation should be used additionally.

Author's response: As the reviewer mentioned, the difference between hypothetical observations (TRUE) and real observations (OBS) is large in the summer, which causes the increase in summer RMSD in OSSE experiments.

We made hypothetical TRUE CO_2 mole fraction data different from OBS because we liked to produce hypothetical true data close to real data but not the same. The estimated model CO_2 mole fractions may represent or similar to the real observed CO_2 mole fractions, but they are constrained much by the real observation network. Thus, when we choose observation sites using several strategies, the experiment using the current observation network (i.e., CNTL in this study) has more benefits compared to other network designs. To be fairly compared the results from several network configurations, we have made hypothetical true CO_2 fraction data that is somewhat similar to the real feature but still hypothetical.

The large RMSD in summer is caused by the sensitivity of CO_2 flux estimated by inversions using different set of observations. In Fig. 6 of Kim et al. (2017) below, it is shown that the largest difference in surface CO_2 flux between the two experiments (i.e., inversion experiments with and without Siberian JR station observations) occurs in June and July, which represent the active season of the terrestrial ecosystem with a large surface CO_2 flux uncertainty. This feature is also shown in Fig. 6 of Kim et al. (2018b) below. The optimized biosphere fluxes that are weekly cumulated for EB (Eurasian Boreal), ET (Eurasian Temperate), and TA (Tropical Asia) averaged over 2007-2009 show that the differences between experiments become greater from the summer, which implies that the absorption of vegetation in summer has a large impact on the results of each experiment.

Thus, we have revised 2nd paragraph of Section 2.3 and associated texts of Section 3.1 as follows. The added parts are underlined.

"Figure 2 shows the station-averaged time series of CO_2 mole fractions from real observations (OBS), EXTASI, SF1, and an average (i.e., simulated hypothetical observations: TRUE, hereafter) of EXTASI and SF1. The time series of EXTASI is the closest to that of OBS, whereas that of SF1 with a static scaling factor (i.e., 1) differs from OBS, particularly in summer. Kim et al. (2017, 2018b) have shown that the largest difference in surface CO_2 flux estimation between experiments with different settings appears in summer, which is associated with more sensitive response of inversion results to the inversion model configurations for the active season of the terrestrial ecosystem."

"Regarding the BIAS, the three experiments have common variations that increase and decrease around zero, and have high amplitudes in summer compared to other seasons (Fig. 4b), which is associated with large uncertainties in the CO_2 mole fraction observations in summer shown in Fig. 2. In particular, CNTL_MOD (CNTL) shows the maximum positive BIAS of 23.74 (16.43) in early June. In contrast, the BIAS of REDIST is approximately 10.28 at the same time and maintains its value closest to zero among the three experiments. Considering the impact of BIAS on steady simulations of the model, the time series of BIAS also supports that the observation network of REDIST can perform more reliably in optimizing surface CO_2 fluxes in Asia compared to that of CNTL.

The RMSDs of all three experiments increase much in summer (Fig. 4c), which may be caused by large uncertainties in the CO_2 mole fraction observations in summer shown in Fig. 2."

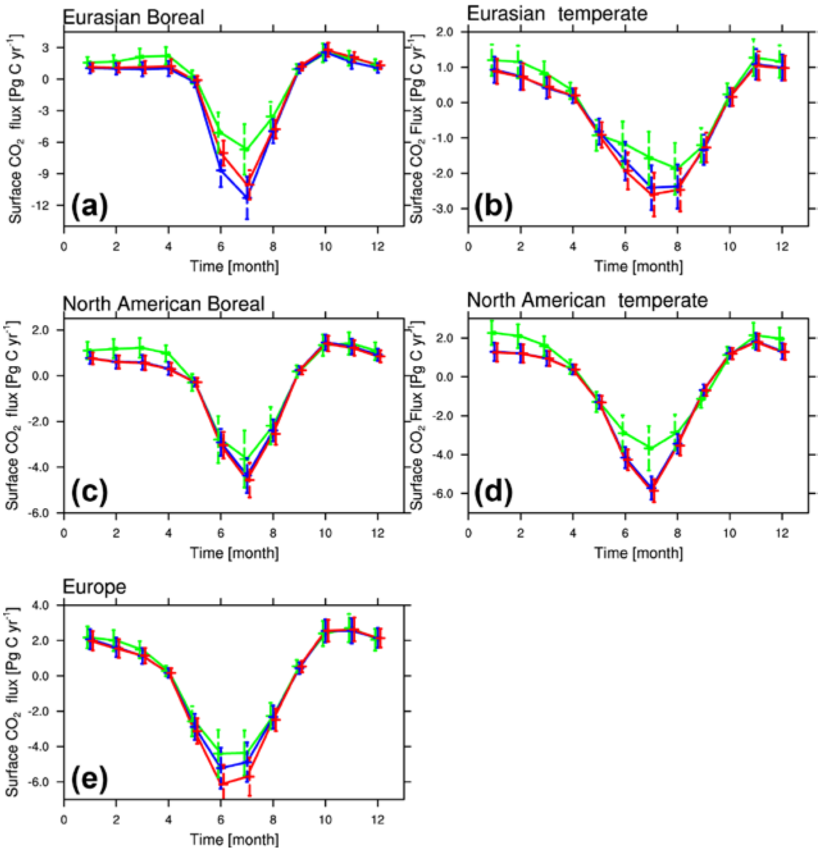


Figure 6. The monthly prior (green) and optimized biosphere fluxes averaged from 2002 to 2009 of the CNTL (blue) and JR (red) experiments with their uncertainties over the (a) Eurasian boreal, (b) Eurasian temperate, (c) North American boreal, (d) North American temperate, and (e) Europe. (Kim et al. 2017, ACP)

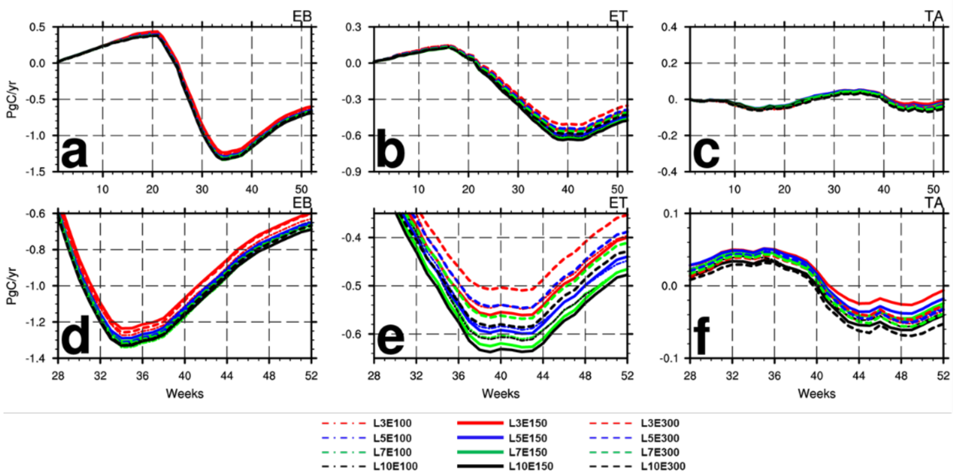


Figure 6. Weekly cumulative flux in: (a) EB, (b) ET, (c) TA region averaged over 2007-2009. (d), (e), and (f) are the magnifications of (a), (b), and (c) in the latter part of year, respectively. Note that EB, ET, and TA region use different scales in y-axis. (Kim et al. 2018b, APJAS)

4. Page 11, Line 24: In addition to the ALL observation network, XCO2 observations of already operated satellites (ex. GOSAT, OCO-2) should be discussed as well as the expansion of the ground observation network.

Author's response: The purpose of this study is showing the validity of the selection strategy (i.e., self-sensitivity and ecoregion information) for potential observation sites. Thus, we have examined these strategies using the ground observation sites. For the satellite observations, the self-sensitivity for the satellite observations are not well known yet. We have to know characteristics, spatial, and temporal distributions of self-sensitivities for satellite observations and how the self-sensitivities of satellite observations are different from those of the ground observations. Thus, the observation network design using both the ground observations and satellite observations will be studied in the future after the self-sensitivities for satellite observations are fully studied. In Section 4, we have added texts (underlined) considering this issue as follows.

"Although the simulation results showed an improvement in performance, the results also suggested that adding 10 extra observation sites in Asia may not be sufficient to fully optimize surface CO_2 fluxes, and more observation sites are required. Reliable observation data from some satellite sensors could supplement the model simulations on the basis of continuous surface observation sites. <u>As the quality of satellite observation data increases, the observation network design for both surface and satellite observation data using the strategies (i.e., normalized self-sensitivity and ecoregion information) of this study will be investigated in the future.</u>"

5. Page 13, Line 16-17: Authors should clarify why RMSD grows in summer. Additional experiments using another hypothetical observation data closer to actual observation data may help. Other possible factors are meteorological conditions and rectifier effects.

Author's response: As the reviewer mentioned, the difference between hypothetical observations (TRUE) and real observations (OBS) is large in the summer, which causes the increase in summer RMSD in OSSE experiments.

We made hypothetical TRUE CO_2 mole fraction data different from OBS because we liked to produce hypothetical true data close to real data but not the same. The estimated model CO_2 mole fractions may represent or similar to the real observed CO_2 mole fractions, but they are constrained much by the real observation network. Thus, when we choose observation sites using several strategies, the experiment using the current observation network (i.e., CNTL in this study) has more benefits compared to other network designs. To be fairly compared the results from several network configurations, we have made hypothetical true CO_2 fraction data that is somewhat similar to the real feature but still hypothetical.

The large RMSD in summer is caused by the sensitivity of CO_2 flux estimated by inversions using different set of observations. In Fig. 6 of Kim et al. (2017) below, it is shown that the largest difference in surface CO_2 flux between the two experiments (i.e., inversion experiments with and without Siberian JR station observations) occurs in June and July, which represent the active season of the terrestrial ecosystem with a large surface CO_2 flux uncertainty. This feature is also shown in Fig. 6 of Kim et al. (2018b) below. The optimized biosphere fluxes that are weekly cumulated for EB (Eurasian Boreal), ET (Eurasian Temperate), and TA (Tropical Asia) averaged over 2007-2009 show that the differences between experiments become greater from the summer, which implies that the absorption of vegetation in summer has a large impact on the results of each experiment.

Thus, we have revised 2nd paragraph of Section 2.3 and associated texts of Section 3.1 as follows. The added parts are underlined.

"Figure 2 shows the station-averaged time series of CO_2 mole fractions from real observations (OBS), EXTASI, SF1, and an average (i.e., simulated hypothetical observations: TRUE, hereafter) of EXTASI and SF1. The time series of EXTASI is the closest to that of OBS, whereas that of SF1 with a static scaling factor (i.e., 1) differs from OBS, particularly in summer. Kim et al. (2017, 2018b) have shown that the largest difference in surface CO_2 flux estimation between experiments with different settings appears in summer, which is associated with more sensitive response of inversion results to the inversion model configurations for the active season of the terrestrial ecosystem."

"Regarding the BIAS, the three experiments have common variations that increase and decrease around zero, and have high amplitudes in summer compared to other seasons (Fig. 4b), which is associated with large uncertainties in the CO_2 mole fraction observations in summer shown in Fig. 2. In particular, CNTL_MOD (CNTL) shows the maximum positive BIAS of 23.74 (16.43) in early June. In contrast, the BIAS of REDIST is approximately 10.28 at the same time and maintains its value closest to zero among the three experiments. Considering the impact of BIAS on steady simulations of the model, the time series of BIAS also supports that the observation network of REDIST can perform more reliably in optimizing surface CO_2 fluxes in Asia compared to that of CNTL.

The RMSDs of all three experiments increase much in summer (Fig. 4c), which may be caused by large uncertainties in the CO_2 mole fraction observations in summer shown in Fig. 2."

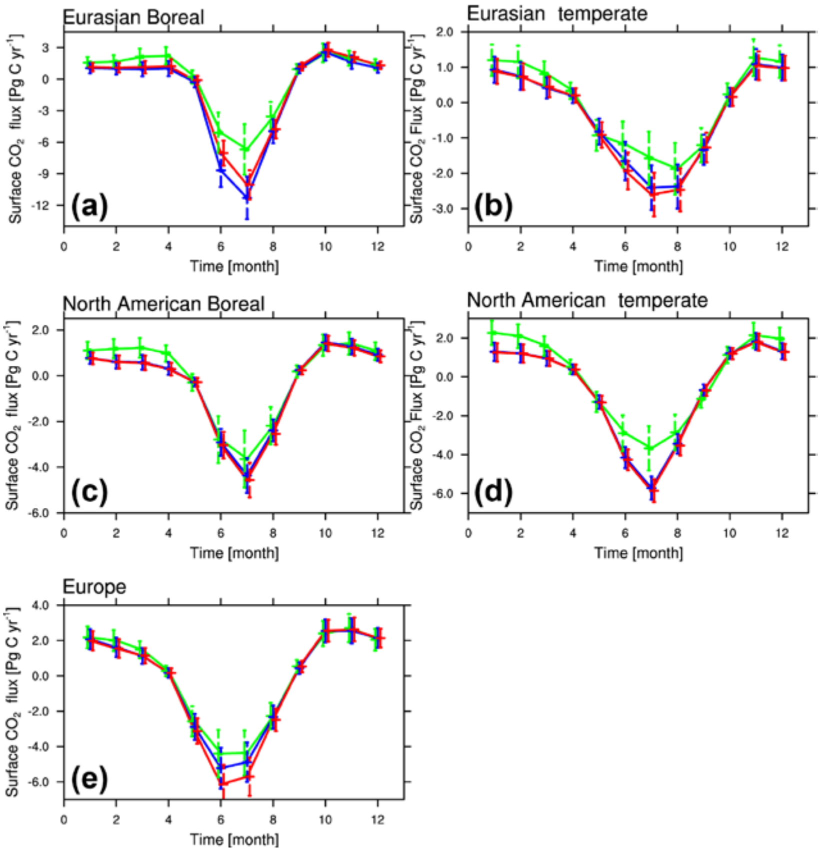


Figure 6. The monthly prior (green) and optimized biosphere fluxes averaged from 2002 to 2009 of the CNTL (blue) and JR (red) experiments with their uncertainties over the (a) Eurasian boreal, (b) Eurasian temperate, (c) North American boreal, (d) North American temperate, and (e) Europe. (Kim et al. 2017, ACP)

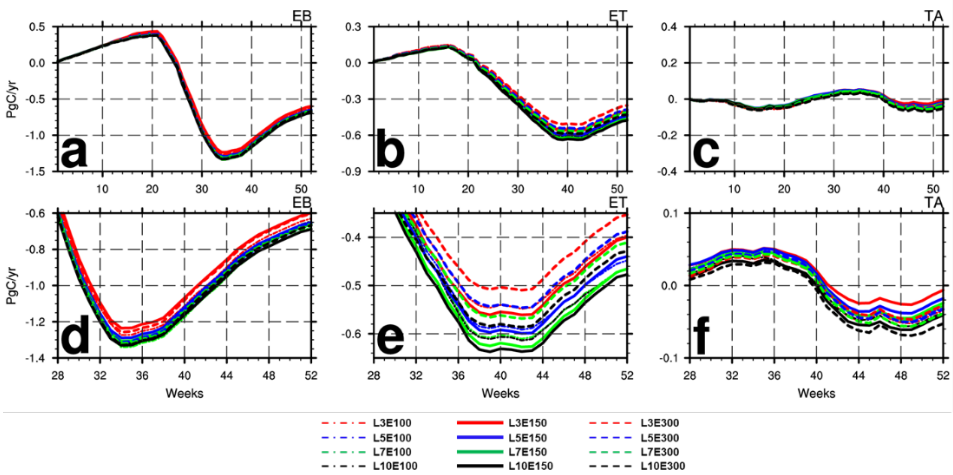


Figure 6. Weekly cumulative flux in: (a) EB, (b) ET, (c) TA region averaged over 2007-2009. (d), (e), and (f) are the magnifications of (a), (b), and (c) in the latter part of year, respectively. Note that EB, ET, and TA region use different scales in y-axis. (Kim et al. 2018b, APJAS)

6. Page 15, Section 3.2: As already mentioned, the authors should evaluate the observation sites that are included in ObsPack and not assimilated in CT, in the sense that they are most feasible.

Author's response: We have added the Section 3.6 that considers many observation sites registered in NOAA ObsPack. We also have added Figs. 11 and 12 in the Section 3.6. The results based on the observation sites registered in NOAA ObsPack are similar to those based on 7 observation sites used for CT2013B.

7. Page 16, Section 3.3: As shown in general comments, authors should implement OSSE that assumes satellites in actual operation (data coverage, accuracy, etc.).

Author's response: The purpose of this study is showing the validity of the selection strategy (i.e., self-sensitivity and ecoregion information) for potential observation sites. Thus, we have examined these strategies using the ground observation sites. For the satellite observations, the self-sensitivity for the satellite observations are not well known yet. We have to know characteristics, spatial, and temporal distributions of self-sensitivities for satellite observations and how the self-sensitivities of satellite observations are different from those of the ground observations. Thus, the observation network design using both the ground observations and satellite observations will be studied in the future after the self-sensitivities for satellite observations are fully studied. In Section 4, we have added texts (underlined) considering this issue as follows.

"Although the simulation results showed an improvement in performance, the results also suggested that adding 10 extra observation sites in Asia may not be sufficient to fully optimize surface CO_2 fluxes, and more observation sites are required. Reliable observation data from some satellite sensors could supplement the model simulations on the basis of continuous surface observation sites. As the quality of satellite observation data increases, the observation network design for both surface and satellite observation data using the strategies (i.e., normalized self-sensitivity and ecoregion information) of this study will be investigated in the future."

8. Page 18, Line 1-2: Authors should consider and show the reason (There are other similar examples).

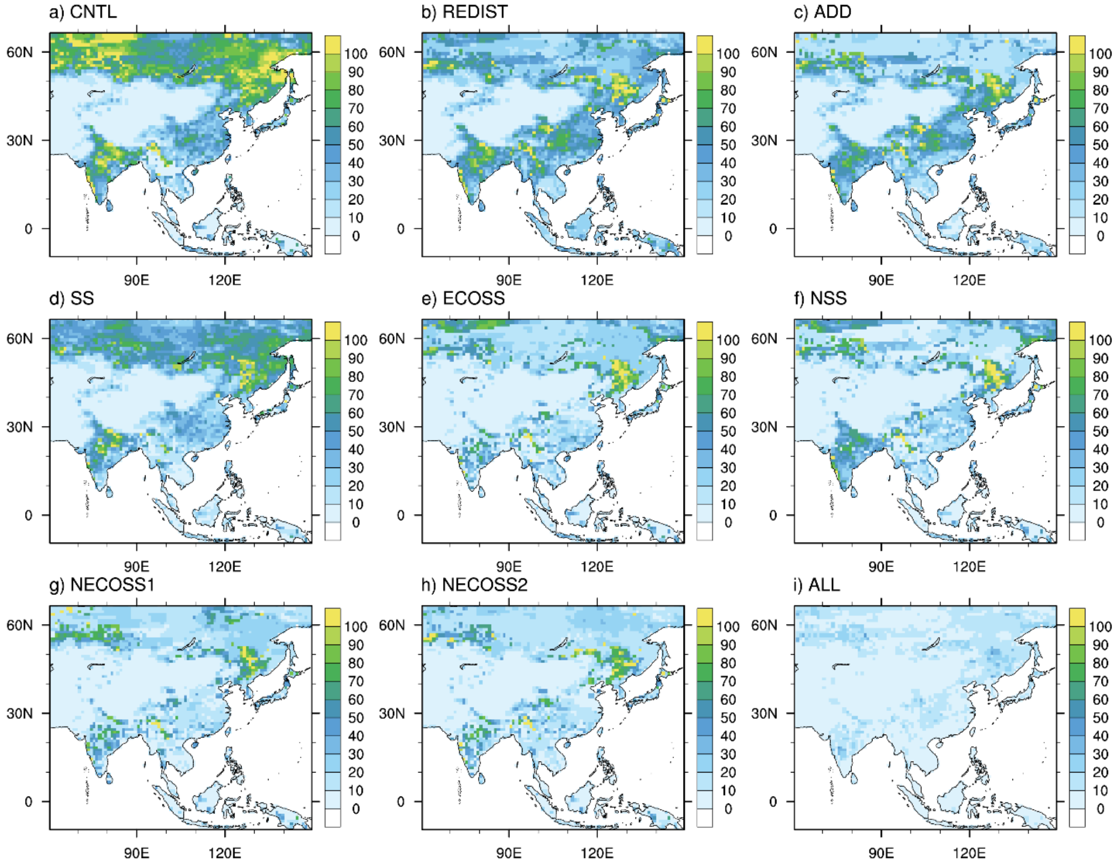
Author's response: The SS strategy determines the potential observation sites based on SS values. The SS values can be concentrated in relatively small area, or in certain ecoregions. This concentration of large SS values can cause few observation sites in other ecoregions since we have a limited number of observation sites as an addition constraint. Thus, the SS strategy can cause a larger bias in certain period. To clarify, we have revised the text as follows. The added and revised parts are underlined.

"However, the BIAS of SS shows a sudden increase in early June, with a maximum positive BIAS of 21.79 (Fig. 7b), which is associated with concentrated sites by large SS values in certain ecoregions that cause not enough DA in other ecoregions. Although the BIAS of ECOSS is generally closer to 0 than that of ADD, except in July, ECOSS shows the maximum negative BIAS of -15.78 in late July. These tendencies suggest that the DA method that optimizes parameters such as the scaling factor used in CT2013B may occasionally have trouble in optimizing surface CO_2 fluxes when using limited observation sites for a larger area."

9. Page 20, Section 3.5: Authors should show summer RMSD of surface CO_2 fluxes and discuss their features.

Author's response: Following the reviewer's suggestion, we investigated the summer (from June to August) RMSD of surface CO_2 fluxes (Fig_rev2). Although the magnitude of the summer RMSD is stronger than that of all year (Fig. 9 in the manuscript), the characteristics of the spatial distributions for all year and for three months in summer are very similar. Since the summer RMSD governs the RMSD of all year, they are similar. Thus, we have not included the Figure for summer RMSD in the revised manuscript. Instead, we have included the text below at the end of the first paragraph of Section 3.5.

"The spatial RMSD distribution during the summer from June to August (not shown) is also similar to that for whole year shown in Fig. 9."



Figure_rev2. The spatial distribution of the average of weekly RMSD of surface CO_2 fluxes (gC m^{-2} yr^{-1}) from June to August for a) the CNTL, b) the REDIST, c) the ADD, d) the SS, e) the ECOSS, f) the NSS, g) the NECOSS1, h) the NECOSS2, and i) the ALL experiments.

10. Table 3-6: Since Ecoregion Index is difficult to understand intuitively, authors should include the region number and vegetation type.

Author's response: Following the reviewer's suggestion, we have added Transcom region and ecosystem type information in Tables 3, 4, and 5.

11. Figure 1: The authors should specify Transcom region boundaries in Asia. If the vegetation type can be illustrated, it is still preferable.

Author's response: Following the reviewer's suggestion, we have included the map of Transcom regions and ecoregions for Asia in Fig. 1b. We also have added Transcom region information in Tables 3, 4, and 5.

References

Byrne, B., Jones, D. B. A., Strong, K., Zeng, Z. C., Deng, F. and Liu, J.: Sensitivity of CO2 surface flux constraints to observational coverage, J. Geophys. Res. Atmos., 122(12), 6672–6694, doi:10.1002/2016JD026164, 2017.

Kim, H., Kim, H. M., Kim, J. and Cho, C.-H.: Effect of data assimilation parameters on the optimized surface CO2 flux in Asia, Asia-Pacific J. Atmos. Sci., 54(1), 1–17, doi:10.1007/s13143-017-0049-9, 2018b.

Kim, J., Kim, H. M., Cho, C. H., Boo, K. O., Jacobson, A. R., Sasakawa, M., Machida, T., Arshinov, M. and Fedoseev, N.: Impact of Siberian observations on the optimization of surface CO2 flux, Atmos. Chem. Phys., 17(4), 2881–2899, doi:10.5194/acp-17-2881-2017, 2017.

Patra, P. K., Maksyutov, S. and TransCom-3 modelers: Sensitivity of optimal extension of observation networks to the model transport, Tellus B, 55, 498– 511, 2003.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R. and Tao, S.: Potential of European 14 CO2 observation network to estimate the fossil fuel CO2 emissions via atmospheric inversions, Atmos. Chem. Phys, 185194(July), 4229–4250, doi:10.5194/acp-18-4229-2018, 2018.

Yang, E.-G., H. M. Kim, J. Kim, and Kay J. K.: Effect of observation network design on meteorological forecasts of Asian dust events, Monthly Weather Review, 142, 4679-4695, doi:10.1175/MWR-D-14-00080.1, 2014.