Response to Reviewer #1's comments:

We would like to sincerely thank the referee for taking the time and effort in reviewing our manuscript. Their thoughtful comments (in bold style below) have helped us improve the manuscript greatly. Especially following his suggestion, we designed another OSSE to illustrate the limitation by only using the smoothing operator as the persistence dynamical model to generate all future scaling factors, which made the improvement of our work more obviously. The changes listed below have been incorporated into a final version of the manuscript.

General Comments:

A regional ensemble-based data assimilation system was developed to estimate CO2 surface fluxes and CO2 concentrations from atmospheric trace gas observations. Because of a lack of a suitable dynamical model to couple forecasted CO2 fluxes and analyzed CO2 fluxes, a new smoothing operator is proposed to estimate forecasted CO2 fluxes at finer scales. However, authors did not compare this new operator directly with the one used in the Carbon Tracker (Peters et al, 2007) to show its impact. The assimilation system needs to be described more clearly and the evaluations only in the OSSE context without using real observations are simply not enough. Therefore, a successful major revision is needed for this paper to be published.

The difference between our dynamical model and the one used in CarbonTracker (Peters et al, 2007) is in the way to set values for $\lambda^{\mathrm{p}}_{i,t|t-1}$. In CarbonTracker, all $\lambda^{\mathrm{p}}_{i,t}$ are set to 1. So the persistence dynamical model is only the smoothing operator. In our study, the CO_2 ensemble forecasts of the atmospheric transport model are used to calculate the values for $\lambda^{\mathrm{p}}_{i,t|t-1}$. So the persistence dynamical model in our study is associated the smoothing operator with the atmospheric transport model. We have discussed this difference briefly in Line 4 to Line 14, Page 10. Besides, we designed another OSSE to illustrate the limitation by only using the smoothing operator as the persistence dynamical model to generate all future scaling factors in Line 9, Page 19 to Line 6, Page 20. Then we discussed the assimilated results in Line 22, Page 26 to Line 3, Page 28. Please see details in the revised version manuscript.

This is the first time of introducing our regional carbon data assimilation system, CFI-CMAQ, so we focus mainly on introducing the methodology. We developed a persistence dynamical model to forecast the surface CO_2 flux scaling factors by associating the smoothing operator and the atmospheric transport model in CFI-CMAQ, so that the surface CO_2 flux scaling factors can be forecasted at grid scale without random noise. And finally, CFI-CMAQ can optimize surface CO_2 fluxes at gird scale. We tried to illustrate this ability of CFI-CMAQ though a set of OSSEs in this manuscript and the results demonstrated that CFI-CMAQ could in general reproduce true fluxes at grid scale with acceptable bias. For another thing, carbon data assimilation remains in its infancy and there are still many challenging scientific problems such as the large inaccuracies in chemical transport models, the sparseness of observation data, and so on. So most published works on optimization surface CO_2 flux through the use of data assimilation technique are still only in the OSSEs. To the author's knowledge, there are only a few works to assimilate real ground-based measurements (eg., Peters et al., 2007; Zhang et al., 2014a, 2014b) and there is no work to use real satellite retrievals. The reason that we did not use real ground-based measurements is because of the sparseness and heterogeneity of ground-based measurements. There are no more than 20 surface CO2 concentration observation stations in our model domain and most are located in Japan and Korea (Zhang et al., 2014a, 2014b). The reason that we did not use the real satellite retrievals is because of large inaccuracies of the chemical transport models and the satellite observations. Further work is needed to optimize surface CO_2 fluxes by assimilating real satellite retrievals. Therefore, using OSSEs is the best way to illustrate the ability of CFI-CMAQ at our first step. Nevertheless, we are trying to assimilate GOAST retrievals to constrain the surface CO_2 flux in the future.

1) The introduction section seems too long as compared with the remaining other sections.

We abbreviated the introduction by deleting statement that every one knows, such as the first sentence in the first draft of our manuscript, or by changing the structure of the sentences (eg. Paragraph 2, Page 6). Ultimately, the introduction section is in less 5 pages. Please see details in Lines 6, Page 3 to Lines 14, Page 7.

2) P.20352 Line 16: use F0 to be consistent with formula (1)

The superscripts/notations used in the first draft of our manuscript were not all consistent really. In the revised version, they are standard. $F^*(x,y,z,t)$ (refer to as $\boldsymbol{F}_t^*$) was served as the prescribed net CO_2 surface flux in formula (1) in Page 7 and the corresponding symbol has been changed. In this study, it was generated by formula (25) (Page 17). In addition, the superscript p, f, and a are standard.

Among them, the superscript p refers to the prior. It was used in the following variables:

① $F^{\mathrm{p}}(x,y,z,t)$ (refer to as $\boldsymbol{F}_t^{\mathrm{p}}$): the prior surface CO_2 flux. It was generated by Eq. (24) (Page 16) in this study. In all the OSSEs in this study, F_t^{p} was assumed as the true surface CO_2 flux.

② $\boldsymbol{\lambda}_{i,t|t-1}^{\mathrm{p}}$: the prior values of the linear scaling factors. We have addressed the way to generate $\boldsymbol{\lambda}_{i,t|t-1}^{\mathrm{p}}$ in Line 16, Page 9 to Line 3, Page 10.

③ $C^{\mathrm{p}}(x,y,z,t)$ (refer to as $\boldsymbol{C}_t^{\mathrm{p}}$): the artificial true CO_2 concentration fields. Forced by $\boldsymbol{F}_t^{\mathrm{p}}$, the RAMS-CMAQ model was run to produce the artificial true CO_2 concentration fields $\boldsymbol{C}_t^{\mathrm{p}}$ from 1 January 2010 to 30 March 2010. It was addressed in Line 18 to 20, Page 16.

④ X_{CO2}^{p} or $\boldsymbol{y}_{\mathrm{t}}^{\mathrm{obs}}$: the artificial GOSAT observations, which were generated by substituting $\boldsymbol{C}_t^{\mathrm{p}}$ into Eq. (19). It was addressed in Line 20 to 21, Page 16.

The superscript f refers to the forecast or the background. It was used in the following variables:

② $\hat{C}_i^{\mathrm{f}}(x,y,z,t)$ (referred to as $\hat{\boldsymbol{C}}_{i,t}^{\mathrm{f}}$): which was generated by applying CMAQ to integrate from time $t-1$ to t forced by $\boldsymbol{F}_t^{*}$ with $C_i^{\mathrm{a}}(x,y,z,t-1)$ as initial conditions. It was used to generate $\boldsymbol{\lambda}_{i,t}^{\mathrm{p}}$. It was addressed in Line 17 to 20, Page 9.

② $\overline{\hat{\boldsymbol{C}}_{i,t}^{\mathrm{f}}}$: $\overline{\hat{\boldsymbol{C}}_{i,t}^{\mathrm{f}}} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{C}}_{i,t}^{\mathrm{f}}$

③ $C_i^{\mathrm{f}}(x,y,z,t)$ (refer to as $\boldsymbol{C}_{i,t}^{\mathrm{f}}$): the *i*th ensemble member of the background concentration fields. CMAQ integrates from time $t-1$ to t forced by $\boldsymbol{F}_{i,t|t-1}^{\mathrm{a}}$ with $C_i^{\mathrm{a}}(x,y,z,t-1)$ as initial conditions. It was addressed in Line 21 to 22, Page 10.

④ $\overline{\boldsymbol{C}_t^{\mathrm{f}}}$: the ensemble mean of $\boldsymbol{C}_{i,t}^{\mathrm{f}}$. $\overline{\boldsymbol{C}_t^{\mathrm{f}}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{C}_{i,t}^{\mathrm{f}}$.

⑤ $C^{\mathrm{f}}(x,y,z,t)$ (refer to as $\boldsymbol{C}_t^{\mathrm{f}}$): the background (wrong) CO_2 concentration fields. Forced by $\boldsymbol{F}_t^{*}$, the RAMS-CMAQ model was run to produce these CO_2 concentration fields from 1 January 2010 to 30 March 2010. That was addressed in Line 17 to 19, Page 17.

⑥ X_{CO2}^{f}: the column-averaged concentrations of $\boldsymbol{C}_t^{\mathrm{f}}$ at the GOSAT X_{CO2} locations, which were generated by substituting $\boldsymbol{C}_t^{\mathrm{f}}$ into Eq. (18). It was addressed in Line 19 to 20, Page 17.

The superscript a refers to the analysis. It was used in the following variables:

① $\boldsymbol{\lambda}_{i,j|t-1}^{\mathrm{a}}$: analyzed quantities from the previous assimilation cycle at time *j*, $|t-1$ means that these factors have been optimized by using observations at time $t-1$.

② $\boldsymbol{F}_{i,j|t-1}^{\mathrm{a}}$: analyzed fluxes from the previous assimilation cycle at time *j*.

③ $\overline{\boldsymbol{F}_t^{\mathrm{a}}}$: the ensemble mean values of the assimilated fluxes, which are before the next smoother window and will not be updated by the succeeding observations.

We regarded them as the final optimized fluxes. It was addressed in Line 11 to 13, Page 12.

④ $\boldsymbol{C}_{i,t}^{\mathrm{a}}$: the *i*th member of the assimilated CO_2 concentrations fields.

⑤ $\overline{\boldsymbol{C}_{t}^{\mathrm{a}}}$: the ensemble mean values of the assimilated CO_2 concentrations fields, which is regarded as the final analyzing concentration field.

⑥ X_{CO2}^{a} : the column-averaged concentrations of $\overline{\boldsymbol{C}_{t}^{\mathrm{a}}}$ at the GOSAT X_{CO2} locations, which were generated by substituting $\overline{\boldsymbol{C}_{t}^{\mathrm{a}}}$ into Eq. (18).

3) P.20352, the way the prior scaling factor $\lambda_{i,t}^{\mathrm{p}}$ is updated is associated with the atmospheric transport model, which should be considered as an important scientific improvement over the one used in Carbon Tracker (Peters et al, 2007). Direct comparison is needed here to show this new smoothing operator, as authors mentioned in the paper, could avoid the "signal-to noise" problem and estimate the surface CO2 fluxes at the grid scale.

We have discussed this difference between our dynamical model and the one used in CarbonTracker in detail in Line 4 to Line 14, Page 10. Besides, we designed another OSSE to illustrate the limitation by only using the smoothing operator as the persistence dynamical model to generate all future scaling factors in Line 9, Page 19 to Line 6, Page 20. Then we discussed the assimilated results in Line 22, Page 26 to Line 3, Page 28. Please see details in the revised version manuscript.

4) P.20353, lines16-20, the formulas seem confusing. In (3), j should start from t-M+1 and end at t. In (4), $\mathbf{S}_{j,t|t-1}^{\mathrm{e}}$ and $\mathbf{P}_{t,t|t-1}^{\mathrm{e}}$ should be the identical, and both should be defined at j since the integration of transport model from j=t-M+1 to j=t is involved. In formula (7), different symbol should be used to represent smoothing operator expressed by formula (2) because M has been used in (2) to denote the lag-window size. Also, it can be seen in (2), smoothing

operator is a function of all $\lambda^{a}_{i,j|t-1}$ in the window.

The symbols used in the first draft of our manuscript were really nonstandard. Even there were some mistakes. In the revised version, they are standard. Also, the mistakes have been revised. In formula (3), *j* starts from *t*-*M* and end at *t*. (see Fig. 1). In formula (4), $\mathbf{S}^{e}_{j,t|t-1}$ is the background error cross-covariance between the state vector $\lambda^{a}_{i,j|t-1}$ and $\lambda^{a}_{i,t|t-1}$, so it is defined at *j*. $\mathbf{P}^{e}_{t,t|t-1}$ is the background error covariance of the state vector $\lambda^{a}_{i,t|t-1}$ and is not related to *j*. $\mathbf{K}^{e}_{j,t|t-1}$ is related to time *j*, so it is defined at *j*. In formula (7), $\varphi(\bullet)$ is used to represent the atmospheric transport model. $M+1$ is used as the lag-window size. We have changed these in the revised manuscript, please see details in Line 2 Page 11, to Line 2, Page 12.

5) P. 20354: Merging subsection 2.4 with subsection 2.3.

We have merged subsection 2.2 with subsection 2.3. Please see details in Line 3 to 13, Page 12.

6) P.20356: The prior scaling factor will be updated based on the inflated CO2 concentration forecast, so it has been inflated indirectly. Why does it need to be inflated again in (17)?

The prior scaling factors have been inflated indirectly though the inflated CO2 concentration forecast. However, the values of the ensemble spread of $\lambda^{p}_{i,t|t-1}$ before inflating are very small. So we have to inflate them again in Eq. (20) before using them into Eq. (2) . We explained that in detail in Line 20, Page 25, to Line 7, Page 26 and added Fig. 11 to illustrate.

7) P.20359 Line 9, specify the year of the OSSE experiments.

All the numerical experiments started on 1 January 2010 and ended on 30 March 2010. We have specified the year in the manuscript (see in Line 13, Page 18).

8) P20360 Line 1-3: Fig.5 should be mentioned here.

Fig. 5a and 5b are mentioned in Line 11, Page 21.

9) P20360 Line 15: "Fig.7" should be "Fig.6".

I have changed the mistakes in Line 2, Page 22.

10) P20362 Line 6: "Fig.9" should be "Fig.10".

I have changed the mistakes in Line 17, Page 24.

11) P20371 Fig.2: The flowchart seems confusing as it is not clear what next cycle should look like. Also, symbols used in the chart are inconsistent with those used in the text part of paper. For example, H in the text represents the whole observation operator including the atmospheric transport model, the bilinear interpolation and weighted CO2 column average. While in the flowchart, it represents everything except the atmospheric transport model.

In order to describe the procedure clearly, we revised Fig. 1 and Fig. 2 and their descriptions. Fig. 1 show that in the previous assimilation cycle *t*-1-*M*~*t*-1, we had the optimized scaling factors in the smoother window $(\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1-M|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M+1|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,j|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1|t-1})$ and the assimilated CO_2 concentrations fields at time $t-1$ ($\boldsymbol{C}^{\mathrm{a}}_{i,t-1}$). In the current assimilation cycle *t*-*M*~*t*, we should optimize the scaling factors in the smoother window $(\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M+1|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,j|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t|t-1})$, and update the forecast CO_2 concentrations fields at time t ($\boldsymbol{C}^{\mathrm{f}}_{i,t}$). We added these in Line 2 to 7, Page 9.

When the assimilation cycle moved on, the scaling factors in the smoother window and the CO_2 concentrations fields are optimized by applying the observations. Fig. 2 is the flowchart of every assimilation cycle. It shows that CFI-CMAQ includes the following four parts in turn at each optimization cycle (1) forecasting of the linear scaling factors at time t, $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t|t-1}$ (red arrows); (2) optimization of the scaling factors in the smoother window , $(\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-M+1|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,j|t-1}, \cdots, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1|t-1}, \boldsymbol{\lambda}^{\mathrm{a}}_{i,t1|t-1})$, by EnKS (blue arrows); (3) updating of the fluxes in the smoother window , $(\boldsymbol{F}^{\mathrm{a}}_{i,t-M|t-1}, \boldsymbol{F}^{\mathrm{a}}_{i,t-M+1|t-1}, \cdots, \boldsymbol{F}^{\mathrm{a}}_{i,j|t-1}, \cdots, \boldsymbol{F}^{\mathrm{a}}_{i,t-1|t-1}, \boldsymbol{F}^{\mathrm{a}}_{i,t|t-1})$ (green arrows); and (4) assimilation of the forecast CO_2 concentration fields at time t, $\boldsymbol{C}^{\mathrm{f}}_{i,t}$ by EnKF (black arrows). We address these in Line 8 to 19, Page 8.

In addition, in the revised manuscript, the observation operator $H(\bullet)$ in the updating equation of the EnKS (Eq. (3) to (7)) and the EnKF (Eq. (13) and (14)) is the same for convenience. It includes the bilinear interpolation and weighted CO2 column average. Please see detail in the manuscript.