A joint data assimilation system (Tan-Tracker) to simultaneously estimate surface CO2 fluxes and 3-D atmospheric CO2 concentrations from observations
Tian et al., 2013

Summary
This paper describes a data assimilation system that simultaneously estimates surface CO2 fluxes and atmospheric CO2 concentrations. Through Observing System Simulation Experiments (OSSEs), the authors demonstrate the performance of the system. It shows that simultaneously estimating surface CO2 fluxes and CO2 concentration performs better than only estimating CO2 fluxes. Estimating CO2 concentration along with CO2 fluxes was first proposed by Kang et al. (JGR, 2011), though they did not explicitly show that estimating CO2 concentration helped the CO2 flux inversion. Even though this paper did a nice job showing the importance of simultaneously estimating CO2 concentration and CO2 fluxes, some of the critical discussions are seriously flawed, the methodology description is not clear, and the English is hard to understand because of the incorrect grammar and usage of words. Therefore, I do not recommend publication of this paper at this stage on ACP. Below lists my detailed comments from major to minor.

1. Throughout the paper, the authors emphasize that one of the advantages of their system is the use of a persistent forecast operator as a dynamical model for the surface CO2 fluxes, so they regard the surface CO2 fluxes as prognostic variable instead of model parameters. They discuss in the abstract that the observations can propagate from one assimilation window to the next assimilation window because of this persistent operator. This discussion is seriously flawed. First of all, using of the persistent operator is not new in surface CO2 flux estimation, which was first used in Peters et al. (2005), and then in Kang et al. (2011, 2012). Second, this persistent operator cannot be regarded as a dynamical model, since there is no dynamics (some knowledge of the evolution of the fluxes as a function of physical variables (e.g., temperature)) built in. Therefore, this study still estimates the surface CO2 fluxes as model parameters as in previous studies. Third, the propagation of the observation information from current assimilation cycle to the next is not because of this so-called "dynamical" operator. Most likely, it is because how the truth and the prior flux were set up in this study. In this study, the prior flux is 1.8 times of the true flux over the whole year, while what they estimate is the scaling factor. You can imagine that the optimized scaling factor could be 1.5 after the current assimilation cycle, so it certainly helps when you start from 1.5 in the next assimilation cycle, since the optimal scaling factor is constant number for the whole year. However, it does not mean the persistent operator carries the observation information forward. On the other hand, estimation of CO2 concentration could remember the observation information in the past, since CO2 forecast is through a dynamical model. For the surface CO2 flux estimation, the CO2 observations observed in the past could not help the estimation, since the flux only affects the concentrations in the future.
2. In the experimental setup, the prior flux is a constant ratio (i.e., 1.8) of the true flux, and then the data assimilation propagates the optimized scaling factor from one assimilation window to the next. This makes the problem too simplistic. In Feng et al.,

(2009), they also use constant scaling factor, but they do not propagate the optimized scaling factor from one cycle to the next. Therefore, each assimilation cycle is independent, which makes it more realistic. I suggest the authors using a totally different flux as the prior flux, so the ratio between the prior flux and the true flux is not constant, but changes with time. The setup in this way will be more realistic. I do not think the authors can conclude the super performance of the system unless they can demonstrate the data assimilation can recover more realistic scaling factors that change with time.

3. One of the advantages of using EnKF is its ability of generating uncertainty estimation. I suggest the authors including the discussion on the uncertainty estimation from the data assimilation system.
4. The "observations" in the title "A joint data assimilation system (Tan-Tracker) to simultaneously estimate surface CO2 fluxes and 3-D atmospheric CO2 concentrations from observations" should be "CO2 observations".
5. The methodology described in Section 2.1 is not clear, since the terminology is not well defined. Many places in this section are confusing. I listed a few here:
    a. L24-L25 on P24760, Fs* is explained before it is actually used in any equation.
    b. L3 on P24762: what is the increment interval between i=1 and i=Lo? It only says that Lo is the length of the optimized window, which is one-week.
    c. L13-L15 on P24762: the description is confusing. Why there are so many different types of simulated observations, and how these different variables are used in the data assimilation.
    d. What fluxes are used in the background run and what fluxes are used in sampling run? And what are the differences in these two fluxes? And how are they generated? (section 2.1)
6. Equation 23: Covariance localization in EnKF is normally applied either on observation error covariance (e.g., Miyoshi et al., 2007) or the background error covariance (e.g., Houtekamer and Mitchell, 2001). The way this equation is written indicates that it localizes the observation increment, not the covariance. So it is not consistent with the section title. In the end, equation 23 can get to the similar result as the localization either on observation error covariance or background error covariance, since they all try to decrease the impact of the remote observations. However, the localization scale may not be comparable with the literature.
7. Where are the observation sites? What is this based on? I suggest including a map that shows the locations of observation sites.
8. Section 3.1 only briefly describes the fluxes used. I suggest including more details of the fluxes, such as the sources of each category of the fluxes, and whether there is diurnal cycle in the biosphere fluxes.
9. Use of the word "assimilation" or "assimilate" is not correct in many places. Here, I listed a few:
    a. In the abstract: "And thus both CO2 concentration and CFs are jointly assimilated by using the atmospheric observations". It should not be "assimilated", but should be "constrained" or "estimated".
    b. L20-21 on P 24767. Probably, the authors mean that the CO2 concentrations are not updated along with CFs in TT-S.

c. Line 23 on P24767 “assimilated” is not used correctly.
d. Line 15 on P24771: “simultaneous assimilation of CO2 concentrations and CFs in our current Tan-Tracker data assimilation system.”. “assimilation” should be “estimation”.

Throughout the paper, there are grammar issues, and incorrect use of words. I only listed two here:
1. Line 14 on P 24758: appear explicit should be “appear explicitly”.
2. L1 on P24760: “first guessed”, should be “first guess”