Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-54-AC2, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development Discussions
Open Access

Interactive comment on "A dual-pass carbon cycle data assimilation system to estimate surface CO_2 fluxes and 3D atmospheric CO_2 concentrations from spaceborne measurements of atmospheric CO_2" *by* Rui Han and Xiangjun Tian

Rui Han and Xiangjun Tian

hanrui@mail.iap.ac.cn

Received and published: 17 July 2019

Responses to Anonymous Referee #1

We would like to thank the reviewer for carefully reading our paper and for all the valuable questions/comments/suggestions. We have thoroughly studied the reports and made a minor revision to our paper by incorporating all suggestions given in the reports. Below we give our itemized responses to all the comments.

"- regarding the theory: I can not understand how the split of the inversion cycle into 2

steps ("passes") could be an improvement. It raises theoretical issues, at least regarding the assimilation of some data twice. And it should hamper the proper distinction between the errors from the prior initial conditions and from the prior fluxes. The need to use a shorter assimilation window for the first pass is an indication of this limitation. Controlling both the CO2 initial condition and fluxes together, using the transport model and the prior uncertainties to drive the balance between corrections to the initial condition and to the surface fluxes, should lead to more robust results and is much more satisfying in terms of theory. The "dual pass" could be seen as a pragmatic way of controlling "manually" this balance (by playing with the length of the assimilation window for the first "pass"), but refining the set-up of the prior uncertainties in the initial condition and in the fluxes is a much proper way for such a control."

Response

The dual-pass data assimilation strategy is well developed and used in earth science, (Moradkhani et al., 2005a, 2005b; Vrugt et al., 2005; Xu et al., 2015); It is well known that the dual-pass data assimilation strategy is absolutely not a reuse of observation because that the observations are used in the two assimilation passes to assimilate different variables separately. The proposed dual-pass assimilation strategy is really an improvement in carbon cycle data assimilation. This novel dual-pass strategy (with different length windows) is specially proposed to give the roper distinction between the errors from the initial conditions and from the background fluxes: For a certain assimilation cycle, the simulated CO2 concentration errors originated from both the initial CO2 and the background flux errors. These errors entangled with the model evolution, which is indeed difficult for us to optimize the CO2 concentrations and fluxes altogether. Fortunately, in the proposed dual-pass assimilation strategy, a shorter 3-days CO2 assimilation pass is first adopted to optimize the initial CO2 concentrations. The optimized initial CO2 concentrations are then used for the following flux assimilation pass, which is just a proper distinction between the errors from the initial conditions and the fluxes (For more details, please see section 2.1 and Fig.1 of our manuscript).

In the ensemble-based Tan-Tracker system, the uncertainties are described by the ensembles in the NLS-4DVar, which are further optimized with the ensembles update (see 2.3 Ensemble generation and update of the Tan-Tracker (v1) assimilation system). In a summary, dual-pass presents a proper way controlling both CO2 initial condition and flux successively, using observations and transport model driven by refined prior uncertain. This is more consistent with the theory and is indeed an improvement of the carbon cycle data assimilation community. To be clear, we add the above description in 4. Discussion. (see the supplement to this comment)

"- in practice: my understanding is that the comparison between Tan-Tracker(v1) and Tan-Tracker(v0) in section 3 is completely biased. For the direct comparison in section 3.3.1 and 3.3.2, TTv1 uses 14-day inversion cycles while TTv0 uses 7-day inversion cycles, and, more critically, TTv1 uses 3 iterations for the minimization of the cost function, while TTv0 uses 1 iteration only (they also use different localization radius and in practice, different systems asking for different parameters). Therefore, there is no reason to think that this comparison says something about the "dual pass" approach itself. Actually, TTv0 seems to provide results that are extremely similar to that of TTv1 using 1 iteration and 14-day inversion cycles (see Figure 5b vs. Figure 9b) ! One could even assume that it provides better results than TTv1 using 1 iteration and 7-day inversion cycles (since results are better with 14-day cycles than with 7-day cycles for TTv1), not speaking about using 2000 km localization radius. My understanding is thus that the authors have misinterpreted their experiments and results."

Response

Actually, the comparisons between Tan-Tracker (v0) and Tan-Tracker (v1) are performed with their respective optimal parameters. The same assimilation parameters have different effects and sensitivities in different systems due to different assimilation framework, assimilation algorithm, and different ensemble generation and update methods. To be clear, we add "Note that the comparisons between TT_v1 and TT_v0 are performed with their respective optimal parameters." on page 11 line 7. As com-

plementary, we conducted another group of experiments with the 7-day cycles for the maximum NLS-4DVar iteration number Imax=1 and Imax=3 (see the following figure fig. 1): the TT_v0 performance is not improved with the increasing Imax and the TT_v1 with Imax=3 is the best one among all the experiments.

"The use of very short inversion cycles (here 14-days) exacerbates the problem of the corrections to initial conditions. For CO2 inversions, the use of short inversion cycles can hardly be seen as an advantage. Most of state of the art systems, especially global ones, use very long inversion cycles (1 year and more) to avoid breaking the link between uncertainties in the fluxes in an area and the errors at a remote location a long time later (which have to be solved for in the initial condition when using short inversion cycles while the target of inversion is a better estimate of the fluxes). I guess that the ensemble approach explains the need for short inversion cycles and maybe why results with 14-day inversion cycles are better than with 30-day inversion cycles. However, the manuscript does not attempt at explaining it."

Response

14-days window length is very close to those adopted by other published inversions systems, such as the one week length of Carbon-Tracker (Peters et al., 2007), the 8-days length of Chevallier et al. (2011) and the one-month length of Basu et al. (2013) also the 7-days length of Tan-Tracker (v0) (Tian et al., 2014). Similarly, we also utilize the sensitivity experiments (shown in Fig.8) to determine the optimal assimilation window-length. To be clear, we add "Note that, 14-days flux assimilation pass window length is close to those adopted by some other published inversion systems, such as the one week length of Carbon-Tracker (Peters et al., 2007), the one-month length of Basu et al. (2013) also the 7-days length of Tan-Tracker (v0) (Tian et al., 2014)." on page 14 line 3.

"The actual inversion system (i.e. the NLS-4DVar system, over which lies the TanTracker (v1) framework, and which is the actual code proposed in the "code and data availability" section) has already been detailed in past publications involving the second author. The section 2.2 is just the duplication of material from Tian et al. (2018), Zhang and Tian (2017), Tian and Feng (2015) and even in Tian et al. (2011). It thus cannot be a strong topic of this new manuscript, nor the overall changes from Tan-Tracker v0 to Tan- Tracker v1 i.e. from POD-4DVar with a single "pass" to NLS-4DVar with a "dual pass"."

Response

The NLS-4DVar is an indispensable part of the Tan-Tracker system and some necessary descriptions of this method is needed for the self-consistency of Tan-Tracker, which is also required by the our topical editor: "Additionally, I would suggest to the authors to slightly improve the description of the "assimilation method" as some details seem to be missing for a complete evaluation of the method without reading the two papers, Tian and Feng (2015) and Tian et al. (2018)". As stated above, the proposed dual-pass assimilation strategy is really an improvement in carbon cycle data assimilation. This novel dual-pass strategy (with different length windows) is specially proposed to give the roper distinction between the errors from the initial conditions and the prior fluxes. In additional, the NLS-4DVar algorithm with efficient localization scheme used in Tan-Tracker leads to high assimilation precision with lower computational costs, which is certainly a great improvement to a data assimilation system. âĂČ

"Regarding the specific analysis of this paper on this NLS-4DVar system, I see in Figure 9 that the optimal number of iterations found for the minimization of the cost function, when testing 1,2 and 3 iterations is 3. But 3 is still much smaller to the typical number of iterations usually used for such minimizations, and the experiments and analysis of this paper do not show that the minimization has converged after 3 iterations (Figure 9b even imply the opposite)."

Response

Sensitivity experiments (the following figure Fig. 2) told us that we really need 7∼8

iterations to completely reach the minimization of the cost function. Meanwhile, we also found that the results of three iterations are not much different from the results of more iterations. As such, limited by computational cost, we choose Imax=1, 2 and 3 as the maximum NLS-4DVar iteration number in this sensitivity experiments and Imax=3 performs best as given on last paragraph of 3.3.3 Sensitivity experiments. The opposite performance in Fig. 9b of our manuscript mainly because of spin-up process of CO2 and flux as describe in second paragraph of 3.3.2 Flux.

"It is important to note that the authors forget to say that their system is a global inversion system (and even to say that it inverts natural CO2 fluxes, until the details of the equations clarifies it) which should strongly influence the way the problem of the initial condition should be tackled, the choice of the inversion cycles and of the data assimilation windows, the size of the ensembles and the number of iterations used for the inversions... in a more general way, the authors ignore the influence of the specific framework of their inversions -domain, resolution, data assimilated- on their results and on their choices of values for the inversion parameters."

Response

We added it in the first sentence of abstract and the title, see page 1 line 12: "Here we introduce a new version of the global carbon cycle data assimilation system". Tan-Tracker (v1) was based on GEOS-Chem, so it has same domain and resolution of it, as described in 3.1 Model settings and observations, and observation information can also be found there. Considering the selection of parameters, we show two main sensitivity experiments only to evaluate flux assimilation pass window length and maximum NLS-4DVar iteration number shown in fig.8 and fig.9 of our manuscript. The sensitivity experiments of the flux assimilation pass localization radius were used to select a localization radius of 1000 km (denoted as Loc-1k), 2000 km (denoted as Loc-2k), or 4000 km (denoted as Loc-4k); the other TT_v1 parameters remained unchanged. The flux and concentration results are shown in following figure Fig. 3. The time series of the monthly total flux (Fig. 3a) and the CO2 concentration (Fig. 3c) results showed that the assimilation results is better with 2000 km localization radius. It is reasonable to be longer than 900km in Carbon-Tracker (Peters et al., 2005) and Tan-Tracker (v0) because of a shorter model integration time meaning lower error from remote location. To be clear, we add this experiments on page 14 line 15 and Fig. 10 on page 31. The selection of localization parameters is based on the limit of computational cost (Zhang and Tian 2018).

"I add, without entering into too much details about it, that the quality of the text is not sufficient for a scientific publication. The abstract already gives a good illustration of the confusing way with which this manuscript is written. The authors insert a lot of technical jargon from the data assimilation community, but they actually misuse many of the corresponding terms (few examples: "surface flux inversion measurement", "flux assimilation" to speak about flux inversion assimilating CO2 data, the alternative use of "background" or "prior" to speak about the same thing, "one-step iteration", "atmospheric chemical transmission mode", "ensemble-based hybrid assimilation algorithm"...), leading to meaningless of confusing sentences. The abusive use of words and mathematical notations that seem more complicated than needed (or that they just forget to define, such as Py in eq 15) severely hampers the clarity of the paper."

Response

a. "surface flux inversion measurement" on page 2 line 14, 20; page 3 line 8 , 14 and Page 4 line 12 was changed to "surface flux inversion". And "CO2 concentration measurement" on page 15 line 10 was changed to "CO2 concentration".

b. Change "flux assimilation" to "flux assimilation pass" in our manuscript.

c. "Background flux" is different to "prior flux" in our manuscript as shown in Eq. 1. "background flux" served as the assimilation background field and where "prior flux" means prior flux data sets. To be clear, we add an explanation on page 4 line 21: "Note that for one certain assimilation cycle, "background flux" is different to "prior flux" as shown in Eq. 1; "background flux" served as the assimilation background field where

"prior flux" means prior flux data sets." To be clear, "prior (background) flux" on page 11 line 5 was changed to "prior flux"; "background results" on page 11 line 21 was changed to "prior results"; "direct background simulations" on page 15 line 15 was changed to "prior simulations".

d. "one-step iteration" on page 1 line 24 and on page 2 line 27 were changed to "one iteration".

e. "atmospheric chemical transmission mode" on page 3, line 12 was changed to "atmospheric chemical transport model".

f. "ensemble-based hybrid assimilation algorithm" on page 4, line 15 was changed to "ensemble-based assimilation algorithm".

g. Description of Py can be find on page 7 line 2, and now moved below Eq. 15.

Sorry for the inconvenience. After looking through our manuscript we changed some other errors, Listed below:

h. On page 2, line 9: "The surface carbon flux inversion method, especially carbon cycle data assimilation, obtained by combining model and atmospheric CO2 information, has made great progress"

i. On page 2, line 13: "Many attempts have been made, assimilating atmospheric CO2 measurements, to optimize surface carbon flux"

j. On page 2, line 25: "The same window lengths make uncertainties of CO2 and surface flux indistinguishable."

k. On page 2, line 30: "Several unconventional data assimilation attempts have been explored."

l. On page 3, line 19: "Specifically, the first performed CO2 assimilation pass uses a shorter window of 3 days to reduce the influence of background flux on the initial CO2"

m. On page 4, line 5: “Based on the NLS-4DVar assimilation method, assimilating satellite column-average CO2 concentration measurements of XCO2”

n. On page 4, line 11: “As such, the evolution of the CO2 concentration in the assimilation window is dominated by the background flux to improve the accuracy of surface flux inversion.”

o. On page 5, line 3: “Considering computational cost, we chose N = 36.”

p. On page 12, line 24: “However, in OSSEs, we can quantitatively analyze the prior flux, optimized flux and real flux to give the most direct judgment.”

“Most of the introduction sounds like a random sampling of references to past inversions, with meaningless comments (like "the surface carbon flux inversion method, obtained by combining model and atmospheric CO2 information, has made great progress in carbon cycle data assimilation", "For example, CarbonTracker is a well-designed carbon assimilation system", "Basu et al. showed that satellite data provided an effective constraint for surface carbon source-sink inversion"...). It hardly provides clues about the specific topic of this paper. The analyzes in section 3.3 lack of depth and of hindsight on the significance and scope of the OSSEs and of the conclusions.”

Response

We try to present in the introduction the development of carbon cycle data assimilation based on different observations, from in situ to aircraft to satellite measurements. After that, we introduce the assimilation system Tan-Tracker (v1). We added a discussion section in our manuscript, please see 4. Discussion.

Reference

Moradkhani H, Hsu K L, Gupta H, et al. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter[J]. Water resources research, 2005a, 41(5). Moradkhani H, Sorooshian S, Gupta H V, et al. Dual state–parameter estimation of hydrological models using ensemble Kalman filter[J].

Advances in water resources, 2005b, 28(2): 135-147. Vrugt J A, Diks C G H, Gupta H V, et al. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation[J]. Water resources research, 2005, 41(1). Xu T R, Liu S M, Xu Z W, et al. A dual-pass data assimilation scheme for estimating surface fluxes with FY3A-VIRR land surface temperature[J]. Science China Earth Sciences, 2015, 58(2): 211-230.

Please also note the supplement to this comment:
https://www.geosci-model-dev-discuss.net/gmd-2019-54/gmd-2019-54-AC2-supplement.pdf

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-54, 2019.

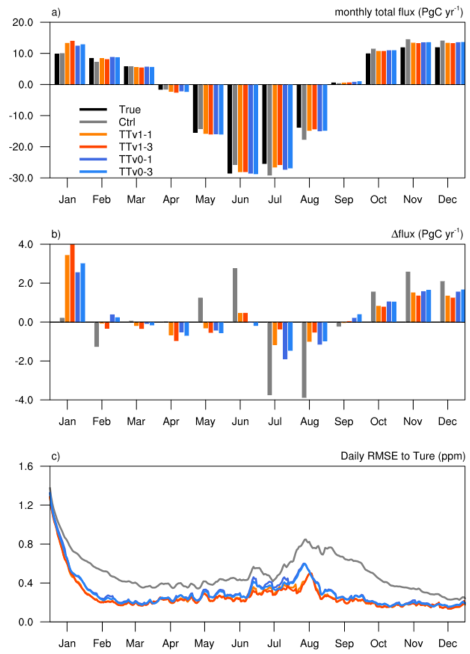

Fig. 1. Sensitivity experiment results: a. monthly total flux; b. monthly total flux deviation; c. daily root-mean-square error (RMSE) of CO2 concentration between the simulation/assimilation results and True

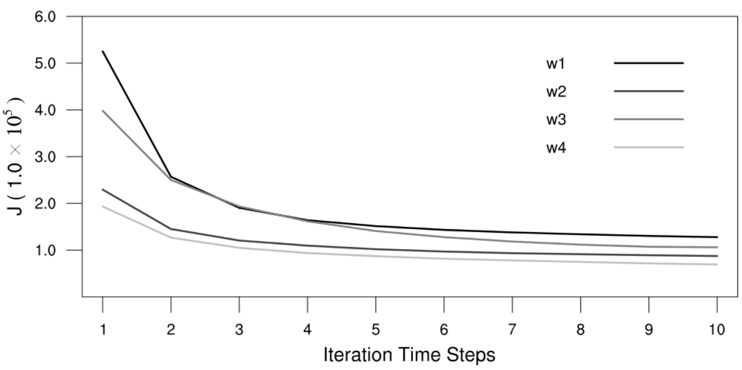

Fig. 2. Iteration time series of cost function value. (TT_v1 assimilation experiments with 10 iterations on flux assimilation pass performed on four consecutive windows, w1 to w4, from June to AugustïijĽ

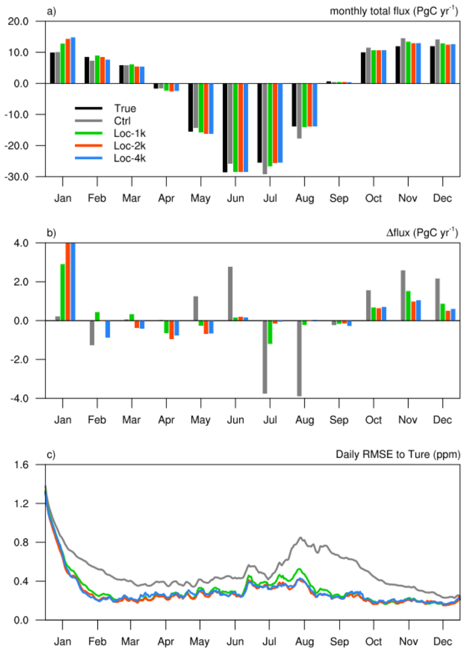

Fig. 3. Localization radii Sensitivity experiment results. ïijĹsame to Fig. 1ïijĽ