Review of “In situ measurement of CO_2 and CH_4 from aircraft over northeast China and comparison with OCO-2 data” by Xiaoyu Sun et al.

General comments:

The paper by Sun et al. reports from airborne in-situ measurements over North-East China in August 2018. The in-situ profiles derived on three different days are compared with profiles from OCO-2 and a carbon cycle data assimilation data system (Tan-tracker). The topic of the manuscript is of high importance since high-quality observations are needed to enable a better analysis of the global carbon cycle. Specifically, in-situ measurements are highly valuable to study local phenomena in detail and to allow for an evaluation of satellite products. This is especially true and important for regions were observations are rare and the variability of the atmospheric greenhouses are not well constrained, because emission amounts are not well known. Therefore I strongly encourage the authors to continue their work in this field because the gained data sets are highly valuable to the carbon community. However, the manuscript lacks on a detailed description and discussion to support the conclusions drawn by the authors. Personally, I also doubt the quality of the aircraft-borne in-situ measurements and therefore suggest publication of the manuscript only after my main (specific) comments are carefully addressed.

Specific comments:

I suggest to re-structure the manuscript and to expand the section “instrumentation” to “methods” by including a subsection on Tan-Tracker and OCO-2 (including a thorough description of the model products and the derivation of the OCO-2 data product).

Page 2, L41: all-weather?

P3, L92: Which one? AIMMS-20?

P3, L94: Why did you use a CVI inlet? Where there other measurement (aerosol) systems onboard? Please be also more specific w.r.t. to the airborne set-up. Did you need to use an external pump to achieve the large gas flow? How long was the inlet (from the tip to the cell)? I am not aware of a publication which reports the airborne deployment of this kind of analyzer, so I suggest to include a schematic which shows the set-up and the periphery to control cell/inlet pressure, temperature and volume or mass flows.

P3, L96: SL/min? Which kind of Mass Flow controller?

P3, L97: The given values are from the manufacturer and might be valid for controlled laboratory conditions. Usually, the performance on a mobile platform is highly affected by variations of pressure, temperature and/or mechanical vibrations. I assume that this specific instrument is even more sensitive since it is not especially designed for use aboard research aircraft. Did you cross-check the theoretical precision values yourself in the laboratory, e.g. by supplying the system with sample gas of constant CO_2 and CH_4 mixing ratios? Did you check the sensitivity of your instrument to changes in pressure and temperature? Did you check the short and long-term drift of your instrument´s sensitivity (i.e. over one flight and over a couple of days, respectively)? Did you check the repeatability of your measurements?

P3, L98: What do you mean with response time in this case? Is this the response time of the system to a change in atmospheric concentrations (due to e.g. the residence time in the inlet)? Is it the averaging time to achieve the given precision (in theory)? O is it the flush time of the cell and thus, gives the best achievable time resolution?

P3, L98: Please specify: Where was the pressure controller installed? I assume in front of the instrument? How constant was the pressure during the flight?

P3, L99: Which temperature? The cell temperature? A range of more than 6 degrees seems very huge to me and should impact the sensitivity of the instrument. Did you check this in the lab (see also above)?

P3, L99: Please provide more details on the Standards. How many standards did you use? At which concentrations? How did you calibrate the system? Did you (or some of the other institutions cross-calibrate the standards in a way that they are traceable to the typically used NIST standards?

How stable was the system? How reproducible were the standard measuremenst before and after the fligh?

P4, L107, Figure 1: What do you want to show with this figure? I suggest to zoom in and to include at least the flight patterns of all 3 flights conducted in August 2017. You might also show a series of three figures with all three flight paths plotted over a weather map.

Section 3:

This section is especially weak. I suggest to include information about the flight strategy, as well.

So e.g., why did you fly in in the morning hours (during which the boundary layer develops)? Did you try to match the time of an OCO-2 overflight? Or was this due to logistical (ATC) reasons? Why did you choose to fly over a horizontal distance of 150 km (the swath is a couple of km´s only)? Did you adjust the flight track to measure along-track of OCO-2? Or is this Did you always follow the same flight strategy on the three days? What was the descending rate and the corresponding pressure variation during spiraling down?

P4, L110, Figure 2: From this figure it looks like that you did ~7 constant flight legs, is that correct? I don´t think you need this figure if you provide a horizontal map of the flight patterns as suggested for figure 1, which gives an idea about the flight dimensions in Lat/Lon direction. Instead, I suggest to include a simple time-series of in-situ measured CO2, CH4, and flight altitude for this particular flight.

Section 4: Please keep in mind that there are several ways how variable water vapor levels influence the CO2/CH4 measurements: 1) The dilution effect, 2) variation in the line broadening of the carbon dioxide and methane lines due to varying water vapor concentration, and 3) nonlinearity of the reported water vapor concentration due to self-broadening of the water vapor line.
Here you discuss the dilution effect which certainly is the most important one. However, the water concentration measurement must be highly accurate to allow for a meaningful accuracy in the derived dry gas concentrations. Therefore I´d like to see an in-depth error analysis for the approach used herein. Moreover, I suggest to include the time-series of measured relative humidity and derived water vapor (including error bars!), at least in the supplement.

P5, L 132: All data are recorded at 1s and then smoothed to remove errors because of the response time? As mentioned above, please be clear in the use of your wording w.r.t. response time. The residence time usually can be corrected for if e.g. volume flow and inlet pressure are known.

Section 5:

Figure 3: This figure shows that you did much more constant flight legs than it seems from figure 2. Do these dots represent 10s values? What is real variability and what is instrument precision? Variation in CO2 on each leg is large (maybe also because of the large horizontal distance), the vertical variability on Aug 10 seems larger than the horizontal variability on that day. Is this a real atmospheric feature and do you have any explanation for this? Also, the boundary layer variability on 9 Aug seems much larger than on other days. Is this a horizontal gradient? Please include a more in-depth discussion on these profiles. I´d like to also see the standard deviation or even better, median instead of average values including some percentiles. What was the boundary layer height? Please include also a vertical profile of met. variables, at least in the supplement.

P5, L 142: "was attributed to different weather conditions": Please provide more details on this hypothesis.

P6, L161: Please give a short introduction about the Model, the used a-priori information and the simulations – and the difference. Which data are assimilated? The OCO-2 data? Doesn´t look like. The aircraft data?

P6, L 163: "The variation of CO2 ….". I don´t understand this sentence. Do you talk about the aircraft measurements? What about uncertainty bars for the aircraft data?

P6, L165: Reproducing CO2 uptake from vegetation by a model is highly challenging, but I do not see any information from the model (neither a-priori or simulated). Is this what you mean with "Below 2km, CO2 is assumed to be vertically mixed…"?

P6, L166: OCO-2 data were averaged over what area and what time? Please provide a graphical explanation which OCO2-Data you used. How did you get the vertical information?

P6, L175: "…with large differences in values". So do you have any explanation? Apart from the quality of the in-situ data, one reason might be the comparison of measurements on different days. To get an impression about the day-to-day variability in that region, you might have a look at a longer time-series of OCO-2 data. You also might have a look at the weather conditions (low or high pressure systems, frontal crossings) and how these may have influenced the day-to-day variability.

P7, L194: Did you use all these observations (aircore, balloon, aircraft) for your specific case or is this a general description? Is this Tan-Tracker? Please be more specific.

P7, L197: This information come much too late (should be at the beginning of section 5.2).

P7, L199: You compare XCO2 values with the in-situ measured data. The variability of the latter is nearly 40 ppm, which is not at all captured by the OCO-2 average profile. In my opinion, you can´t compare column values and derive a bias (especially not with the accuracy given).

Table 2: Please provide details on the uncertainty analysis for the aircraft errors: accuracy (traceability to WMO scale) and precision.