Interactive comment on “A comparison of in-situ aircraft measurements of carbon dioxide to GOSAT data measured over Railroad Valley playa, Nevada, USA” by J. M. Tadić et al.

Anonymous Referee #1

Received and published: 6 September 2012

The manuscript describes a series of measurements of CO₂ aircraft measurements performed as part of a GOSAT validation project. While it is apparent that there is a need for such data, the authors fail to put their results into a broader perspective, and it remains unclear how these data will contribute to improve GOSAT satellite retrievals. I am aware that this is a technical paper, but still I would like to see some more scientific discussion of the data, especially since there is no technical advancement reported.

Besides CO₂, also CH₄ mixing ratios were measured, and the data are shown, but there is no discussion of these, which is blamed on a lack of satellite retrievals. Pre-
sumably these will become available soon, and the publication of the study should be postponed so that it can be extended to CH$_4$.

In large parts the presentation is poorly structured and difficult to follow. In its present form the manuscript is not suitable for publication in AMT.

**General comments**

Overall, more attention needs be given to the way the work is presented, in particular to figure quality and readability. Several figures and tables are not properly referred to in the main text. The structure of the text needs to be improved, as connected information is scattered throughout the different parts. Sections 2 and 3 are cut into several subsections and even subsubsections which do not connect well. The extremely brief section about the aircraft (2.2.) should be merged with the description of the site and the flight patterns, parts of it make more sense in section 2.1. The data analysis section is very short, and especially its introductory paragraphs contain experimental details that should go to the instrumental section.

The instrumental section itself is rather general, e.g. I could not find details about the time resolution and how it relates to the precision. Some important information is missing, such as the calibration scales that the data are reported on. What is the calibration procedure? Why is only one standard used for calibration and why was one chosen that has as high a CO$_2$ mixing ratio as 416.27 ppm, much higher than the values encountered during the aircraft measurements? It is mentioned that linearity is checked with three synthetic standards. Were these compared to the NOAA ESRL standard? 1 % relative to NIST is not a very high accuracy. Please quantify the degree of linearity of the instrument.

It is mentioned that H$_2$O mixing ratios were logged as well. What range of mixing ratios was encountered? Chen et al. mention that in more humid air CO$_2$ and CH$_4$ measure-
ment results were affected. How did you correct for the influence of water vapor, and how does it affect the measurement uncertainty? Did you account for that during calibration? In the appendix the built-in water-vapor correction function is mentioned, but there is no assessment of the uncertainty of this. Also in the appendix it is stated these data cannot be calibrated, that’s confusing, how would you record mixing ratios without calibration. Without a thorough discussion of this, I cannot agree with the statement that the accuracy of the CRDS measurement is limited by the primary NOAA ESRL standard.

The discussion of the data and the results needs to be extended. Too little weight is put on the presentation of the actual measurement results and their representativeness. The observational site was chosen because it is flat, and it is a desert site which is far from major sources and sinks. There is no discussion how the measurements at this particular site can be compared to other regions, e.g. with a more complex orography. The authors claim: "Agreement of the satellite and aircraft CO₂ mixing ratios as well as ground measurements fall within the uncertainties of the methods employed to acquire these numbers." However, there is no discussion how these data contribute to improve GOSAT retrievals in the future.

Two extrapolation methods to higher altitudes are used, one applying a fixed tropopause height. It should be discussed how varying tropopause heights - and the varying relative contribution of stratospheric air - influence the results. CO₂ and CH₄ behave differently across the tropopause and in the stratosphere, which is also worthwhile a brief discussion. How does the averaging kernel affect the results?

**Specific comments**

Use hPa for pressure values throughout the text (currently Torr and hPa are used alternately, that’s confusing)
Use consistently ppb for CH₄ mixing ratios in the text and the figures, they are sometimes given in ppb, sometimes in ppm.

Abstract

Line 3: remove instrument type in parentheses, too detailed for abstract

Line 4: remove "(a.m.s.l.)"; not used again in abstract

Line 8: remove ", also"

Introduction

P5643,L3: no comma here

P5643,L11–L14: Make clear that [Vay et al.] is only one regionally restricted example. There are many more measurements, some even on a regular basis and more systematic, e.g. the NOAA aircraft program, the Japanese CONTRAIL project, also a French program using a Licor instrument, or the recent HIPPO missions. Chen et al. is not an appropriate reference here, that paper discusses the instrument and a comparison with NDIR during aircraft flights, but it does not discuss vertical profiles or data application for satellite validation.

The point of this study is, as I understand, not the discussion of vertical profiles (their shape does not get discussed at all in the result section), but the validation of the satellite retrievals. However, no background information on the validation issue is given, and there are no references to other validation studies studies. Also, what are the specific requirements for validation of satellite instruments using aircraft data?
Experimental

Page 5645, L 4/5: "The additional (...) filter". That filter was not mentioned before. What is its purpose?

Page 5645, L 5: What type of inlet did you use, how far out did it reach?

Page 5645, L 9: Explain "operating equilibrium temperature".

Page 5645, L 12: subsection title not numbered

Page 5645, L 14: Omit last digit of quoted mixing ratios, if the precision is not that high.

Page 5646, L 10–15: this should be part of section 2.1

Page 5646, L 27: reference to Table 1 is missing here

Page 5647, L 5: no new paragraph here

merge 2.3.1 and 2.3.2 into one section

Data Analysis

Page 5648, L 1: This information should be part of the instrumental section, this also applies to some facts from the previous paragraph.

Page 5648, L 26: Why was a different instrument type used for the ground measurements? What is the difference between the two instrument models? The second instrument was not mentioned in the instrument section.

Page 5649, L 1: What do you mean by "duration of the flight in the spiral"? Is that
the duration to complete one of the circles or only for the one at the lowest altitude? Explain and refer to Figure 2.

Page 5649, L 3: How good is that agreement?

Page 5649, L 17: Explain what you mean by "informed guess"

Page 5649, L 10: What does the acronym AJAX stand for? Fig. 5 has AJAX, AJAX 1 and AJAX 2 in the legend, the text only mentions the latter two.

Page 5649, L 17: Discuss Figure 6 in more detail.

Page 5649: There is no discussion of the uncertainty induced by the two different extrapolation procedures.

Page 5649, L 22: remove "easily"

Page 5649, L 24/25: is 25 hPa the resolution of the vertical grid?

Page 5650, L 2: correct "and and" to "and"

Page 5650, L 7: change "filled in" to "filled in". Better use "extrapolated".

Page 5650: the comparison is too qualitative and a discussion of uncertainties (which was projected on the previous page) is missing.

Page 5650, L10: The difference is up to 0.5 ppm, this is not "very close". Again, assessment of uncertainties of the extrapolation modes is necessary to interpret this.

Page 5650, L15: one cannot assess synoptic variability based on only two individual measurements.

Page 5650, L17: how did you derive the value of the bias?
AJAX 2 mixing ratios are below AJAX 1 mixing ratios on all three days, please discuss that.

Conclusions

It does not become clear what can be learned from these measurements, and how GOSAT retrievals can be improved using the result. The main conclusion seems to be that the current retrievals are not too far off. Which extrapolation method (both fairly simplistic) is to be preferred and for what reason? How does this study relate to previous validation measurements, how good an agreement would be desirable or achievable?

Appendix

This is relevant information. Move it to the main text, e.g. at Page 5648, line 18, and correspondingly incorporate Figure A1 or make it Figure 5b and discuss its content.

Figures

The font size used for axis labels and legends is too small in Figures 2, 3, 5, 6, A1. It needs to be enlarged.

Figures and tables are not always referred to in the main text. Some are referred to in parentheses, but they are not discussed in context. They should be incorporated into the discussion.

Figure 1: This Figure seems to be not referred to in the main text

Figure 2: the color scheme is not well-chosen, even larger variations are barely visible.
Use "flight path" or "flight track" rather than "trajectory" (more commonly used for air mass origin).

Figure 4: What’s the purpose of that figure? Aircraft photos are nice to be shown in presentations but hardly contain relevant information. The picture does not even illustrate the usage/integration of the instrument onboard the aircraft, therefore remove it. Besides, this Figure seems to be not referred to in the main text anyway.

Figure 5/A1: start y-axis from 0, axis label should be "altitude amsl" as it is used in the text.