Authors are grateful to the anonymous referee for helpful and thoughtful comments. Each comment is addressed individually below. The referee comments are indicated in green, and our responses are described in black.

Interactive comment on "Bias corrections of GOSAT SWIR XCO2 and XCH4 with TCCON data and their evaluation using aircraft measurement data" by M. Inoue et al.

Anonymous Referee #1

Received and published: 3 March 2016

This manuscript presents the derivation and application of an empirical bias correction for GOSAT XCO2 and XCH4 retrievals, and provides some evaluation of this bias correction through the use of independent aircraft measurements.

This paper was very clearly written and well structured, which made it quite pleasant to read and review. Nonetheless, I have some concerns that I would like to see the authors address.

Regarding the stratospheric completion of the methane profiles discussed at the top of page 7: I appreciate the effort to include the trend rather than just using a fixed climatology for the stratospheric extension of the methane profile above the tropopause, but I wonder if using a fixed growth rate is the best choice. The atmospheric growth rate as measured at the surface has been highly variable in the past years, ranging from 4.67 ppb/year in 2012 to 12.36 ppb/year in 2014 (see http://www.esrl.noaa.gov/gmd/ccgg/trendsch4=globalgrowth): This growth rate is of course lagged somewhat in the stratosphere

In this study, the annual trend "6 ppb / year" of CH_4 profiles above the tropopause was determined based on the recent report from the World Meteorological Organization (WMO). As your comment, the CH_4 growth rate at the surface has been variable in the past years. We show the CH_4 growth rate of each year reported by NOAA/ESRL webpage in Table R1-1.

We calculated aircraft-based XCH_4 using stratospheric and mesospheric profiles at a growth rate of 6ppb / year and at a growth rate of each year (Table R1-1) and compared them in Tsukuba and Park Falls (Table R1-2). Average of the differences between them over Tsukuba was -0.2 ppb with a standard deviation of 0.5 ppb (n=11). In 2014, the XCH_4 difference was as small as 0.7 ppb. Average of the differences between them over

Park Falls was -0.3 ppb with a standard deviation of 0.2 ppb (n=33). The amount of CH_4 above the tropopause was small, and consequently did not have a large effect on the aircraft-based XCH_4 calculation at the two observation sites. Therefore, we decided to use a fixed growth rate "6 ppb / year" in this study.

Table R1-1. CH_4 growth rate of each year reported by NOAA/ESRL webpage (http://www.esrl.noaa.gov/gmd/ccgg/trends_ch4/#global_data).

	Growth rate of CH_4 [ppb/year]
2009	4.74
2010	5.04
2011	5.38
2012	4.67
2013	5.88
2014	12.36

Table R1-2. The average and standard deviation of (i) aircraft-based XCH_4 at a growth rate of 6 ppb per year above the tropopause and (ii) aircraft-based XCH_4 at a growth rate of each year above the tropopause in (a) Tsukuba and (b) Park Falls. The average and standard deviation of the differences between (i) and (ii) are also shown in the rightmost column.

(a) TKB (Tsukuba)

Tsukuba	Number	(i) Aircraft-based XCH_4 at a growth rate of 6 ppb/year [ppb]		(ii) Aircraft-based XCH_4 at a growth rate of each year [ppb]		(ii)- (i) [ppb]	
		Average	SD	Average	SD	Average	SD
2010	3	1785.9	3.5	1785.7	3.5	-0.2	0.0
2011	3	1788.5	10.5	1788.1	10.5	-0.4	0.0
2012	1	1800.9	—	1800.2	—	-0.6	—
2013	2	1806.5	1.7	1806.1	2.0	-0.5	0.3
2014	2	1805.9	7.3	1806.6	7.4	0.7	0.0
All	11	1795.4	11.0	1795.2	11.2	-0.2	0.5

(b) LEF (Park Falls)

Park Falls	Number	(i) Aircraft-based XCH_4 at a growth rate of 6 ppb/year [ppb]		(ii) Aircraft-based XCH_4 at a growth rate of each year [ppb]		(ii)- (i) [ppb]	
		Average	SD	Average	SD	Average	SD
2009	4	1781.8	6.3	1781.8	6.3	0.0	0.0
2010	6	1785.5	6.9	1785.4	6.9	-0.1	0.0
2011	7	1791.6	6.2	1791.3	6.2	-0.3	0.1
2012	11	1790.3	7.5	1789.8	7.5	-0.5	0.1
2013	5	1799.6	13.8	1799.1	13.9	-0.4	0.2
All	33	1790.1	9.3	1789.7	9.3	-0.3	0.2

Regarding the colocation criteria used here: Have you considered using a more sophisticated approach such as that used in the Guerlet et al. 2013 JGR paper, which takes into account the impact of both transport and flux variability on the colocation, rather than simply the geographic limits? This makes it more likely that you are really comparing similar air masses, while simultaneously expanding the potential match area in space. This could be quite useful for getting TCCON colocations with M-gain regions, for instance. At very least a more thorough discussion of the potential drawbacks of the 5 degree x 5 degree approach (and alternatives proposed in the literature, such as that of Nguyen et al., AMT, 2014) should be discussed.

As you suggested, other sophisticated methods may lead to obtaining TCCON co-locations with M-Gain regions. However, we need to be cautious about using other sophisticated methods you pointed out. The methods of Wunch et al. (2011b) and Nguyen et al. (2014) are based on the fact that the distribution of potential temperature at 700 hPa is deeply related to that of CO_2 density in the Northern Hemisphere. As shown in Nguyen et al. (2014), it is hard to apply this method to defining the co-location criteria in the low latitudes over the Northern Hemisphere and in the Southern Hemisphere. In addition, this method is not applicable to XCH_4. Guerlet et al. (2013) and Lindqvist et al. (2015) were based on the distribution of XCO_2 simulated by atmospheric transport model (e.g., the region where there is a modeled XCO_2 value within ±0.5 ppm of standard deviation for the modeled value at the observation site). This can lead to much larger matched data and be applied to the entire globe including the Southern Hemisphere. However, reliable XCH_4 modeled data are hard to obtain, and the sophisticated method for XCH_4 remains to be established.

The purpose of this study is to correct the GOSAT SWIR V02.21 XCO_2 and XCH_4 data that cover five years after the GOSAT launch. Because the number of available TCCON site has rapidly increased after the GOSAT launch, we could obtain enough matched data by the geophysical co-location method (co-location criteria: within ±5° latitude/longitude boxes). As you suggested, we added descriptions on more sophisticated methods in Sect. 3.1 as follows.

"Here, we discuss the spatiotemporal co-location criteria for calculations of the regression analyses. The ideal co-location criteria should be measurements at the same place during the same time (Zhou et al., 2016). In general, geographical co-location defines a spatiotemporal neighborhood region near the location of interest, and collects summary statistics (hereafter referred to as "geophysical co-location method", Cogan et al., 2012; Nguyen et al., 2014; Zhou et al., 2016). A disadvantage of the geophysical co-location method is that the number of matched data can become small when the spatiotemporal criteria are somewhat small. Therefore, several sophisticated methods were devised to obtain a sufficient number of co-located data. Following Keppel-Aleks et al. (2011) who implied a relationship between meridional gradients of free-tropospheric potential temperature and CO_2 concentrations in mid-latitudes over the Northern Hemisphere, Wunch et al. (2011b) used the distribution of potential temperature at 700 hPa when defining the co-location criteria in the Northern Hemisphere. Expansively, Nguyen et al. (2014) utilized a modified Euclidean distance weighted average of distance, time, and temperature at 700 hPa. Since this method is based on the fact that the distribution of potential temperature at 700 hPa is deeply related to that of CO_2 density in the Northern Hemisphere, it is hard to apply this method to defining the co-location criteria in the low latitudes over the Northern Hemisphere and in the Southern Hemisphere. In addition, this method is not applicable to XCH_4. Guerlet et al. (2013) and Lindqvist et al. (2015) were based on the distribution of XCO_2 simulated by atmospheric transport model (e.g., the region where there is a modeled XCO_2 value within ±0.5 ppm of standard deviation for the modeled value at the observation site). This can lead to much larger matched data and be applied to the entire globe including the Southern Hemisphere. However, reliable XCH_4 modeled data are hard to obtain, and the sophisticated method for XCH_4 remains to be established. In this study, five years of GOSAT SWIR V02.21 XCO_2 and XCH_4 data are used for the validation and correction. Because the number of available TCCON site has rapidly increased after the GOSAT launch, we can obtain enough matched data by the geophysical co-location method."

As a reader, I questioned why the HIPPO profiles were used so sparingly. In Table 2f it shows that only 9 HIPPO profiles were considered for analysis, almost all in the southern hemisphere. What about the rest of the HIPPO campaigns and profiles? Presumably of these 9 campaigns, none of them were from July 2009, as these were not included in Figure 7. Perhaps it would be more useful to choose a month where there were HIPPO profiles available to show in a figure like Figure 7? Or is there a practical reason why these measurements could not be included?

Aircraft measurements by the HIPPO mission were conducted changing the locations of observation every day. In addition, they couldn't always obtain complete profile data (e.g., approximately 0.3 km to 8.5 km) needed to calculate aircraft-based XCO_2 and XCH_4. GOSAT returns to the same point in space every three days, and the available GOSAT SWIR data are limited to the regions under clear sky conditions. In ocean regions where HIPPO flies, the latitudes of retrieved GOSAT data tend to be confined to the narrow regions (spatial co-location criteria in this study: within ±5° latitude/longitude boxes). Consequently, we could obtain only nine matched data.

As for Figure 7, the latitudinal distributions in July 2009, we would like to show representative monthly averaged CO_2 concentrations of respective latitudes. Therefore, the measurements by the HIPPO mission with one datum for a month are not shown in Fig. 7.

When reading about the improvement in the correlation coefficient from the uncorrected to the corrected version of the data over land (page 11), I questioned the significance of the improvement. Given the error bars, I doubt that an improvement from 0.70 to 0.71 (or even 0.86 to 0.88) is really significant. Likewise, the decrease in the correlation over ocean is likely not a cause for concern (although it should be not simply be ignored in the text, as is the case now). Taking the uncertainties of the individual measurements into account makes it possible to estimate uncertainties on the correlation coefficients as well, so this should be easily resolved.

We agree with you for the most part. It may be difficult to show the effectiveness of the bias correction by only the correlation coefficients. Before the correction, there already exists a high correlation between the present GOSAT data (V02.21) and TCCON data, and the difference between them showed non-physical relationships, probably the retrieval artifacts (Figs. 2 and 3). We would like to emphasize that our method led to noticeable reduction of GOSAT XCO_2 and XCH_4 biases and vanishing the artifacts with keeping this high correlation. Wunch et al. (2011b) conducted the bias corrections of two versions (V2.8 and V2.9) of ACOS-GOSAT XCO_2 data, where quality in V2.9 is considered to be better than V2.8. Comparisons showed that empirically-derived bias correction improves the agreement between GOSAT data and the TCCON data more remarkably in V2.8 XCO_2 data than V2.9 XCO_2 data.

I stumbled a bit in the interpretation of Figures 7 through 9. To begin with, it would be easier to interpret these figures if the logical colour scheme used previously (green for land, blue for ocean) had been maintained. Amending this is recommended.

We revised them (Figs. 7, 8, and 9) as you suggested. We hope that new Figs. 7, 8, and 9 look better.

In Figure 7, I am not entirely clear what the take-away message should be. I acknowledge that getting a feeling for the improvement globally is difficult, but I don't know why the satellite values are binned while the aircraft values are not. The only significant difference I see between the corrected and uncorrected values of XCO2 are that the land and the ocean seem to be further apart after correction, though still agreeing within uncertainty. The change in agreement with the aircraft measurements is difficult to discern.

In fact, GOSAT data are consistent with TCCON data and aircraft data in many latitudes after the correction (Fig. 7). For clarity, we show some XCO_2 and XCH_4 values in Sect. 3.3 as follows.

"Aircraft-based XCO_2 at the Honolulu site located around 20°N in July 2009 was 387.25 ppm. The uncorrected GOSAT XCO_2 and corrected GOSAT XCO_2 over the land regions around 20°N including Honolulu were 385.00 ppm and 386.56 ppm, respectively, whereas those over the ocean regions were 384.82 ppm and 387.07 ppm, respectively. By the correction, it was shown that GOSAT XCO_2 approached the aircraft-based XCO_2 value over both land and ocean. TCCON XCO_2 at Lauder was 383.71 ppm, and the uncorrected GOSAT XCO_2 and corrected GOSAT XCO_2 around 50°S including the Lauder sites were 381.18 ppm and 382.88 ppm, respectively. In the Southern Hemisphere, we found that GOSAT XCO_2 approached the TCCON value by the correction."

"For example, TCCON XCH_4 at the Ny Ålesund site was 1762.9 ppb, and the uncorrected GOSAT XCH_4 and corrected GOSAT XCH_4 around 80°N including Ny Ålesund were 1746.8 ppb and 1752.0 ppb, respectively. By the bias correction, GOSAT XCH_4 approached the TCCON value."

In Figures 8 and 9 (panels c and d) it can be seen that the difference between the satellite soundings averaged over one latitudinal band and one TCCON site is reduced after bias correction, but, as is stated in the text, "the seasonality in the difference remains". And so it should! There is no reason to think that the measurements of Lamont should be identical to the mean total column values over a 15 degree latitudinal band. This also made the interpretation of panels a and b of the same figures rather difficult. At first I was wondering why the Tsukuba XCH4 values for, say, autumn 2013 so much higher were, but then realized that it was likely just the effect of regional emissions or synoptic transport during this period, and there was no reason to interpret any offset from the zonal mean GOSAT XCH4 value as a problem in the GOSAT data. Thus I found the interpretation of these figures rather difficult. I am not sure how this could best be improved, but at least the limitations of the comparison should be discussed. At present I am not convinced that the figures add a lot to the reader's understanding of the work.

Based on your comments, we added the following sentences in Sect. 3.4.

"Note that the TCCON XCO_2 data at Lamont are values at a particular location, while GOSAT XCO_2 data are zonal averaged values."

"To clarify it, we show that monthly variations of the difference between uncorrected and corrected GOSAT XCO_2 and TCCON XCO_2 at Lamont (Figs. 8c and 8d)."

And finally, and perhaps most importantly, I agree with Referee 2 that the results should be discussed in the context of other related studies using empirical multivariate approaches to correct the bias in retrieved GOSAT products. Are the same variables found to be significant? Are the spatial patterns in correction consistent? A full comparison might be too much to undertake in this study (though it would be an interesting topic for a follow-up study), but at least a superficial comparison would be appropriate. If the authors can address these concerns, I think the manuscript is appropriate for publication in AMT.

We agree. The following sentences were added in Sect. 1.

"Following Wunch et al. (2011b), Cogan et al. (2012) performed bias correction of GOSAT XCO_2 data retrieved from the University of Leicester Full Physics (UoL-FP) retrieval algorithm using pseudo observations based on GEOS-Chem model calculations. Guerlet et al. (2013) used XCO_2 measurements from 12 TCCON sites around the world as a reference for correction of GOSAT XCO_2 data retrieved from the Netherlands Institute for Space Research/Karlsruhe Institute of Technology (SRON/KIT) Full Physics retrieval algorithm."

In addition, we revised the sentences in Sect. 1 as follows.

"Our method has three primary differences from Wunch et al. (2011b): (1) we explicitly use TCCON data from numerous sites throughout the world as reference values for the regression analysis; (2) the regression variables and coefficients for correction of GOSAT data are determined separately for observations made over land and those made over the ocean; and (3) we perform this analysis for both XCO_2 and XCH_4."
--->
"Similar to Guerlet et al. (2013), we explicitly use TCCON data from 22 sites throughout the world as reference values for the regression analysis. Our method has two primary differences from the previous bias correction studies: (1) the regression variables and coefficients for correction of GOSAT data are determined separately for observations made over land and those made over the ocean and (2) we perform this analysis for both XCO_2 and XCH_4."

Based on your comment, we performed a brief comparison of the bias correction obtained from our work with previous works. However, it was hard to discuss the details such as cause of biases in various different retrieval algorithms in this study. The following sentences are added in Sects. 3.2 and 3.3.

"We here compare our results to those by other XCO_2 bias correction study. Cogan et al. (2012) showed the annual mean global difference to be reduced by about half (-1.22 ppm to -0.68 ppm) and the correlation coefficients to increase from 0.61 to 0.74. Thus, our correction method is effective for removing the biases significantly."

"The differences between corrected XCO_2 and uncorrected XCO_2 were about 2-4 ppm and less than 2 ppm in western part and eastern part of North America, respectively (Fig. 6c). This larger spatial gradient over North America is consistent with the result of Guerlet et al. (2013), though the months analyzed in their study (August and September) and our present study (July) differ. This feature over North America may be due to the differences in the type and condition of vegetation which have a strong impact on the surface albedo. Finally, this can have an influence on the estimation of regional CO_2 fluxes over North America by inverse analysis."

Minor comments:

The previous commenter suggested (rightly) that the work be better put in the context of relevant literature. Might it be that the Guerlet et al. 2013 reference on page 4, line 5, is in fact referring to the paper this reviewer brought up?

As mentioned above, the details of Guerlet et al. (2013) were shown in text.

Figure 2 and 3 have very difficult to read axes, and rather poorly chosen axis limits.

We revised them as you suggested. We hope that new Figs. 2 and 3 look better.

Figure 4 b is interesting in that for ocean data the correlation coefficient is quite high and rather improved, while the slope is considerably further from the 1:1 line. (I agree with the other reviewer that information about the slope would be useful here.) What does this shift look like in general (and not just for July 2009)? Does it seem reasonable? Granted the ocean colocations to TCCON sites are rather limited in geographic extent, so this might be difficult to interpret, but I think it warrants discussion.

Fig. 4b shows scatter diagrams between corrected XCO_2 and TCCON XCO_2 during the entire period. We believe that the indication by Referee 2 should be on the statistics for the regression lines (slope) of Figs. 2 and 3. Since the atmosphere over ocean is generally cleaner than that over land because of the absence of polluted air and aerosols from urban areas, the bias features of GOSAT data retrieved over ocean may be different from those over land. However, it is difficult to discuss why the biases of GOSAT data are reduced over ocean remarkably.

P5 line 21: HeymanN spelled incorrectly in citation.

We corrected it.